# Contents

# Deep Dive – questions and answers about Lancaster chapters 43 and 44

Bill Celmaster

November 2025

As I've made my way through Lancaster chapters 43 and 44, I've encountered a number of assumptions, equations, ideas etc. that I found very murky. As time went on, I have concluded that the authors have done their best to construct a coherent and concise narrative of the subject matter, but ultimately have put too much faith in the reader's ability (or at least, my ability) to fill in gaps.

One might argue that the only way to properly address this, is to use a different text – inevitably a text that is much longer and much more thorough and undoubtedly much more work to follow. Instead, I've chosen the ChatGPT route. The good thing about that route, is that I can start from the various things that I do understand, and then prompt ChatGPT regarding the difficulties I have in the text – and keep prompting until we get to 'the bottom' of whatever was bothering me. It requires patience, but provides a service that no mentor would do (partly because the mentor would find it way too tedious if they even knew the answers, partly because the mentor probably would do a worse job than ChatGPT of understanding my issues, and partly because I'd be embarrassed to keep asking questions).

Anyway, what follows is an attempt to go through the various parts of these chapters that, for me, required a deeper dive. This may or may not be useful to anyone else – depending on whether the starting point of their knowledge is similar to what mine was when I began this process.

# 1 Starting with Lancaster chapter 43.1

Everything that follows in this document, is taken at (approximately) 0 temperature. We use this fact to justify our assumption – used almost always – that all particles occupy the states with lowest possible energies. So when we speak of $\mathcal{N}$ quasiparticles, this is approximately the same as saying the lowest energy $\mathcal{N}$ states.

In Lancaster chapter 43.1, Wick's theorem, normal-ordering, contractions, a general two-body interaction and mean field theory are introduced. These ideas come up again in the theory of superconductivity and implicitly in band theory, without a lot said about their connections.

## 1.1 General two-particle interaction and ground states

$$H = H_0 + V + V' \tag{1}$$

where

$$V = \frac{1}{2}\sum_{pqk} V_q a^\dagger_{p-q} a^\dagger_{k+q} a_k a_p \tag{2}$$

and

$$H_0 = \sum_p \frac{p^2}{2m} a^\dagger_p a_p \tag{3}$$

and $V'$ includes interactions between pairs of ions and also includes ion-electron interactions. For now, we put off discussion of $V'$.

HOWEVER, what is also important when discussing superconductivity, is a different way of decomposing H which I'll write as $H = H_{SC} + V_{SC} + V'$ where we define $H_{SC} = H_0 + \delta_{SC}$ and $V_{SC} = V - \delta_{SC} + V'$.

The analysis generally starts with solutions to the eigenvalue problems $H_0\psi_n = E_0^n\psi_n$. Then you use perturbation theory to find the perturbed eigenstates and eigenvalues.

Typically you are most interested in the lowest-energy state(s), which we call the *ground state(s)*. But sometimes we are also interested in the energy-gap, which is essentially the difference in energy between the ground state and the first excited state.

There are 3 separate ground states of interest in Lancaster chptrs 43 and 44:

- The Fock ground state a.k.a Fock vacuum denoted $|0\rangle$ and which is the state annihilated by all the $a_p$ (for the moment I'm suppressing spin indices). Another way of thinking of this ground state, is that it is the lowest-energy state of $H_0$.

- The $N$-body Fermi ground state which I'll denote $|N\rangle$. This is the state which minimizes $H_0$ for $N$ electrons. Because of the Pauli exclusion principle, this state consists of electrons that fill up all energy levels up to the ones whose energy $\epsilon_F$, is determined by $N$ and which is known as the Fermi energy. The corresponding momentum $p_F$, is the Fermi momentum. This $N$-body ground state can be defined by the condition $a_p^\dagger|p\rangle = 0$ for all $p < p_F$. Since all states are occupied below the Fermi energy, it isn't possible to add an electron to any of those states (because of the Pauli exclusion principle). If we define a new operator $f_p = a_p^\dagger$ then we can think of $f_p$ as the annihilation operator for $|N\rangle$.

- The BCS ground state $|BCS\rangle$. This is the state that minimizes the energy amongst states consisting of Cooper pairs and more precisely is the state minimizing $H_{SC}$. It is approximately the state that minimizes the full Hamiltonian $H$ for $N$ electrons. We will see later that it is annihilated by an operator of the form $\gamma_k = \alpha_k\, a_k + \beta_k\, a_{-k}^\dagger$.

  Note that all above ground states $|G\rangle$ are annihilated by a set of annihilation operators $d_{p\sigma} = u_p\, a_{p\sigma} \pm v_p\, a_{-p\sigma}^\dagger$ that obey the "anticommutation rules"[1]

$$\{d_{\sigma' p'}, d_{\sigma p}^\dagger\} = \delta(p - p')\delta_{\sigma\sigma'}$$
$$\{d_{\sigma p}, d_{\sigma' p'}\} = 0$$
$$\{d_{\sigma p}^\dagger, d_{\sigma' p'}^\dagger\} = 0$$

  where $\{A, B\} \equiv AB + BA$. In the above, I've re-enstated the spin index.

[1]Going forward, coherence factors like $u_{\mathbf{p}}$ and $v_{\mathbf{p}}$ will be taken to be real, because states (and therefore ladder operators) can always be redefined by multiplying them by a phase factor. Moreover, the coherence factors are only dependent on the magnitude of the momentum appearing in the index – so $u_{\mathbf{p}} \to u_p$ and $v_{\mathbf{p}} \to v_p$. I haven't thought too hard about why this is. I'm guessing it's an ansatz which is ultimately validated by self-consistency and/or symmetry considerations. Also, notice that in most of the equations that follow, I'll write operators that look like $a_p$ where I intend $a_{\mathbf{p}\sigma}$. I've done this for convenience and hopefully won't confuse anyone by this overloading of symbols.

## 1.2 Normal ordering and ground states

Usually, when we write a Hamiltonian in terms of operators like $a_p$ and $a_p^\dagger$ – where $a_p$ annihilates the Fock ground state – we define the normal ordering of a product of such operators, to be a product where creation operators are all to the left of annihilation operators. For example, for boson operators

$$: a_{p_1}^\dagger a_{p_2} a_{p_3}^\dagger a_{p_4} := a_{p_1}^\dagger a_{p_3}^\dagger a_{p_2} a_{p_4}$$

and for fermion operators

$$: a_{p_1}^\dagger a_{p_2} a_{p_3}^\dagger a_{p_4} := -a_{p_1}^\dagger a_{p_3}^\dagger a_{p_2} a_{p_4}$$

because for fermions, every permutation of operators is accompanied by a negative sign.

However, all of this is imprecise. When we specify normal ordering, we must also specify what ground state and system of annihilation and creation operators we are using. Let's define some notation.

- Let the state $|G\rangle$ be a ground state as defined in the previous section.
- Let $d_p$ be an operator that annihilates $|G\rangle$, i.e. $d_p|G\rangle = 0$, and also $d_p$ follows the anti-commutation rules. We say that $d_p$ is an annihilation operator and $d_p^\dagger$ is a creation operator, both with respect to the state $|G\rangle$.
- Let $A_p = \alpha_p d_p + \beta_p d_p^\dagger$ and $B_p = \alpha_{p'} d_{p'} + \beta_{p'} d_{p'}^\dagger$. Then define "the normal ordering $: AB :_G$ of $A$ and $B$ with respect to the $|G\rangle$ ground state" by $: AB :_G = \beta_p \beta_{p'} d_p^\dagger d_{p'}^\dagger - \alpha_p \beta_{p'} d_{p'}^\dagger d_p + \alpha_{p'} \beta_p d_p^\dagger d_{p'} + \alpha_p \alpha_{p'} d_p d_{p'}$. We see that for each term in the sum, the creation operators are in front of the annihilation operators.
- Define the "contraction $\langle AB \rangle_G$" of $A$ and $B$ with respect to $|G\rangle$, to be $AB - : AB :_G$. By simple calculation, we get $\langle AB \rangle_G = \alpha_p \beta_{p'} d_p d_{p'}^\dagger + \alpha_p \beta_{p'} d_{p'}^\dagger d_p = \alpha_p \beta_{p'} \{d_p, d_{p'}^\dagger\} = \alpha_p \beta_p \delta(p - p')$.
  - Note the following.

$$\begin{aligned}
\langle G|AB|G\rangle &= |G\rangle|\beta_p \beta_{p'} d_p^\dagger d_{p'}^\dagger + \alpha_p \beta_{p'} d_p d_{p'}^\dagger + \alpha_{p'} \beta_p d_p^\dagger d_{p'} + \alpha_p \alpha_{p'} d_p d_{p'}|G\rangle \\
&= \langle G|\alpha_p \beta_{p'} d_p d_{p'}^\dagger|G\rangle \\
&= \alpha_p \beta_p \delta(p - p')
\end{aligned}$$

where the second and third lines are consequences of the fact that the $d$ operators annihilate $|G\rangle$ and the $d^\dagger$ operators annihilate $\langle G|$. What this shows is that

$$\langle AB\rangle_G = \langle G|AB|G\rangle \tag{4}$$

## 1.3 Wick's theorem

Wick's theorem is stated and proved in Lancaster section 18.5. However, that chapter mostly concentrates on the time-ordering version of Wick's theorem. This is what particle theorists normally encounter and we normally encounter it in the context of the Fock vacuum. But in many-body physics, we are also interested in the time-independent version of Wick's theorem – which is more or less the same algebra.

Lancaster, in Eq. (43.3) applies Wick's theorem to the 2-body interaction $V$, expanding the interaction into a sum involving normal-ordered products and contractions. The expansion is for an arbitrary ground-state (as mentioned above, we really can take any arbitrary state, provided we can identify the associated creation and annihilation operators) and Lancaster doesn't show a subscript $G$ for either the normal-ordering or contraction operator. Unfortunately, some of the remarks on page 381 are unclear because the specific ground-state isn't mentioned. For completeness, I'll rewrite Lancaster Eq. (43.4) using the above-defined notations.

$$\begin{aligned}
& a^\dagger_{p-q}a^\dagger_{k+q}a_k a_p =: a^\dagger_{p-q}a^\dagger_{k+q}a_k a_p :_G \\
+\ & \langle a^\dagger_{p-q}a^\dagger_{k+q}\rangle_G : a_k a_p :_G + : a^\dagger_{p-q}a^\dagger_{k+q} :_G \langle a_k a_p\rangle_G \\
+\ & \langle a^\dagger_{p-q}a_p\rangle_G : a^\dagger_{k+q}a_k :_G + : a^\dagger_{p-q}a_p :_G \langle a^\dagger_{k+q}a_k\rangle_G \\
-\ & \langle a^\dagger_{p-q}a_k\rangle_G : a^\dagger_{k+q}a_p :_G - : a^\dagger_{p-q}a_k :_G \langle a^\dagger_{k+q}a_p\rangle_G \\
+\ & \langle a^\dagger_{p-q}a^\dagger_{k+q}\rangle_G\langle a_k a_p\rangle_G + \langle a^\dagger_{p-q}a_p\rangle_G\langle a^\dagger_{k+q}a_k\rangle_G - \langle a^\dagger_{p-q}a_k\rangle_G\langle a^\dagger_{k+q}a_p\rangle_G
\end{aligned} \tag{5}$$

## 1.4 Perturbation theory around the N-body Fermi ground state

If we consider a system of $N$ particles, then its $H_0$ ground state is the Fermi ground state $|N\rangle$ where all electron levels are filled up to the Fermi level

for N particles (at the Fermi level, the electrons have momentum $p_F$ with $E_F = \frac{p_F^2}{2m_e}$). The annihilation and creation operators for this ground state are $f_p$ and $f_p^\dagger$. So to proceed, we should expand the operators $a_p$ and $a^\dagger$ in terms of the $f$ operators – the way we expanded $A$ and $B$ in terms of the $a_p$ operators.

This is particularly simple: Notice that we can't add any electrons that have $|p| <= |p_F|$, since those states are already filled. Thus, for $|p| <= |p_F|$, we have $a_p^\dagger|N\rangle = 0$ hence $f_p = a_p^\dagger$. On the other hand, when $|p| > |p_F|$, we can fill states as usual, so $f_p = a_p$.

We're ready to do perturbation theory around the $H_0$ Hamiltonian. In first order, all eigenvalues $E_p$ of $H_0$ are modified to $E_0 + \langle N|V|N\rangle$. This is just the Born approximation. (Note that Lancaster also refers to the technique in section 43.1 as a "mean field approximation" but in first order perturbation theory, I don't think there's a difference.)

The first thing to notice is that $\langle N| : AB :_N |N\rangle = 0$ as long as $A$ and $B$ are linear functions of $a_k$ and $a_{k'}^\dagger$ for all values of $k$ and $k'$. This takes a bit of analysis. Suppose at least one operator is an $f$ (annihilation operators). Then it is on the right and acts to the right on $|N\rangle$, giving 0. But suppose at least one operator is an $f^\dagger$. Then it is on the left and acts to the left on $\langle N|$. Again this gives 0.

Exactly the same argument applies to any normal-ordered product of operators that are linear functions of annihilation and creation operators.

Therefore the only potentially non-zero terms in $\langle N|V|N\rangle$ are the last three in Eq. (5). These are just c-numbers (times the identity). Notice, however, that the first of these – known as the Cooper term[2] $C_0\ \langle a_{p-q}^\dagger a_{k+q}^\dagger\rangle_N \langle a_k a_p\rangle_N$ is 0 because $\langle a_k a_p\rangle_N = 0$. To see this, remember that $\langle a_k a_p\rangle_N = \langle N|a_k a_p|N\rangle$ and consider the following cases.

- Suppose $|k|, |p| < |p_F|$. Then $a_k = f_k^\dagger$ and $a_p = f_p^\dagger$ and $a_k a_p =: f_k^\dagger f_p^\dagger :_N$ (since normal-ordering won't change the order of a product of two creation operators). Then as shown above, this gives an expectation value of 0 in the $N$ ground state.
- A similar argument shows that when both momenta are above (I mean,

[2]Strictly speaking, this is the Cooper term only when the ground state is $|BCS\rangle$, i.e. when the subscript $G$ in Eq. (5) is set to $BCS$.

of course, the absolute values) the Fermi momentum, the expectation value is 0.

- What about $|k| < |p_F|$ and $|p| \geq |p_F|$?. Then $a_k = f_k^\dagger$ and $a_p = f_p$. The product is in normal order, and therefore by our normal-ordering argument above the expectation value is 0.

- Finally consider $|k| \geq |p_F|$ and $|p| < |p_F|$. We then end up with $f_k f_p^\dagger$. Although this isn't in normal order, the order doesn't matter because the two operators anti-commute since they have different values of momentum. So again, we show that these give the same expectation value (up to a sign) as the normal-ordered product which again is 0.

We have finally shown that the Born approximation is a sum of the last two terms in Eq. (5). Those are the Hartree direct term $D_0$ and the Fock exchange term $F_0$. Lancaster goes on in Section 43.1 and 43.2 to evaluate those terms. Lancaster (or at least my posted notes) also includes the Born approximation for $V'$, in the case that the ions are treated as jellium.

# 2 Superconductivity: Lancaster chapter 44

Before tackling superconductivity, we extended the $V'$ model considered above by Lancaster. Instead of treating the ions as jellium, we treated them as a periodic lattice. That led to the Bloch theory and bands. In terms of the original Hamiltonian of Eq. (1), this simply is a change to $V'$.

Then, as we move into the theory of superconductivity, we simplify things so that the above theory of metals, based on perturbations of the solutions to $H_0$, is now based on perturbations to a new starting Hamiltonian, known as the mean-field approximation of superconductivity. The theory should look roughly as follows:

- The above-mentioned theory of metals – including the Bloch theory – applies to electrons outside of an energy band $E_F - \omega_D < E < E_F + \omega_D$ where $\omega_D$ is known as the Debye frequency, and is approximately the maximum energy of phonons (this isn't something we've discussed up to this point, so simply take $\omega_D$ as an assumption to be covered later.)

- Within the Debye frequency band, electrons are described by the mean-field Hamiltonian to be introduced shortly in these notes, plus a perturbative term $V''$ representing the difference between the mean-field Hamiltonian and the full Hamiltonian of Eq. (1).
- The mean-field Hamiltonian is one that we can solve exactly so it will become the basis for a new perturbative expansion.
- What I'm still murky about is how the two theories – inside and outside the Debye band – are formally coupled together. I think that is covered properly in more detailed texts. In Lancaster Chapter 43 (and many other treatments) the Debye band – which includes the perturbative term $V''$ – is treated independently.

## 2.1 The Reduced CP Hamiltonian

What follows is the treatment given by Lancaster, but also found in Annett and many other texts. This treatment is fairly straightforward and calculationally easy, but unnecessary. Fetter and Walecka, in their chapter on canonical transformations (chapter 37 in my edition of their text), start with the full effective Hamiltonian (the most general two-body interaction term as shown in Eq. (1)) and arrive at the same end-point. The key observation in both treatments, is that Wick's theorem (see later) allows the interaction term to be written as a sum of quadratic monomials plus an irreducible 4-body term which can be argued – then shown – to be perturbatively small relative to the other terms. The quadratic expression can be solved exactly and is what leads to the BCS ground state. I recommend Fetter and Walecka for a thorough treatment of the theory – including the precise inclusion of the interaction coherence factor which I've summarized in Eq. (1))as $V_q$. In the meantime, I'll begin with the Lancaster approach.

The idea is this: When electrons are within the Debye band, they attract one another and form bound pairs – the Cooper pairs. We will treat the superconducting theory as though it were factored away from the "normal" theory outside the Debye band. We therefore focus on a theory entirely consisting of electrons in the Debye band. Instead of referring to the $a_p$ annihilation operators, we'll use operators $c_p$ which are the same – except that their action is strictly on the Debye "sector".

The intuition for an appropriate starting Hamiltonian, is that it should only

contain the momentum combinations appropriate to Cooper pairs with 0 total momentum. In particular, the Cooper-pair sector of the Hamiltonian will look like

$$H_{CP} = H_0 + \delta_S C = \sum_{p\sigma} \epsilon_p c^\dagger_{p\sigma} c_{p\sigma} - \kappa^2 \sum_{pk} c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} \tag{6}$$

where $\epsilon_p$ is the Bloch-state energy minus the chemical potential (i.e. $\epsilon_p = \epsilon^B_p - \mu$)[3] and **the sum is over all momenta in the Debye band.** Notice that the factor $c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}$ is what we've been calling the pair creation operator $P^\dagger_k$. Also notice that the interaction strength $-\kappa^2$ is shown as a negative constant, rather than as a more general $V_{kp}$. This simplification ultimately is a manifestation of the BCS analysis of electromagnetic interactions between electrons as mediated by phonons.

### 2.1.1 The $|CP\rangle$ states

We will consider the many-pair states that can be constructed as $|CP\rangle = \prod_k (u_k + v_k P^\dagger_k)|0\rangle$. Later, I will argue that for the mean-field Hamiltonian, these are the most general many-pair states that need be considered for finding the ground state of that Hamiltonian. For now, let's show the family of CP creation and annihilation operators associated with each $|CP\rangle$ state. Consider the $\mathbf{k}$-sector of the $|CP\rangle$ state:

$$|\mathrm{CP}\rangle_{\mathbf{k}} = \left(u_{\mathbf{k}} + v_{\mathbf{k}}\, c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow}\right)|0\rangle\,, \qquad |u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1.$$

The proposed annihilation and creation operators are

$$\gamma_{\mathbf{k}\uparrow} = u_{\mathbf{k}}\, c_{\mathbf{k}\uparrow} - v_{\mathbf{k}}\, c^\dagger_{-\mathbf{k}\downarrow}, \tag{7}$$

$$\gamma_{-\mathbf{k}\downarrow} = u_{\mathbf{k}}\, c_{-\mathbf{k}\downarrow} + v_{\mathbf{k}}\, c^\dagger_{\mathbf{k}\uparrow}. \tag{8}$$

We now show that $\gamma_{\mathbf{k}\uparrow}$ and $\gamma_{-\mathbf{k}\downarrow}$ are indeed annihilation operators for $|CP\rangle$, namely

$$\gamma_{\mathbf{k}\uparrow}\,|\mathrm{CP}\rangle_{\mathbf{k}} = 0, \qquad \gamma_{-\mathbf{k}\downarrow}\,|\mathrm{CP}\rangle_{\mathbf{k}} = 0.$$

We will also show that $\gamma_{\mathbf{k}\uparrow}$ and $\gamma_{-\mathbf{k}\downarrow}$ and their adjoints follow the standard anticommutation rules for annihilation and creation operators.

First the proof that $\gamma_{\mathbf{k}\uparrow}$ annihilates $|CP\rangle$.

[3] My notation differs from everyone else's. They usually write $\epsilon_p - \mu$ where their definition of $\epsilon$ corresponds to my $\epsilon^B$. Recall that the chemical potential is included in the Hamiltonian so that in a many-body theory, it acts as a Lagrange multiplier used to constrain the particle-number to $N$.

- We compute:

$$\gamma_{\mathbf{k}\uparrow}\,|\mathrm{CP}\rangle_{\mathbf{k}} = \left(u\,c_{\mathbf{k}\uparrow} - v\,c^{\dagger}_{-\mathbf{k}\downarrow}\right)\left(u + v\,c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\right)|0\rangle$$
$$= u^2\,c_{\mathbf{k}\uparrow}\,|0\rangle + uv\,c_{\mathbf{k}\uparrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle - uv\,c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle - v^2\,c^{\dagger}_{-\mathbf{k}\downarrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle\,.$$

- Using $c_{\mathbf{k}\uparrow}\,|0\rangle = 0$, the first term vanishes.
- Also,

$$c_{\mathbf{k}\uparrow}c^{\dagger}_{\mathbf{k}\uparrow} = 1 - c^{\dagger}_{\mathbf{k}\uparrow}c_{\mathbf{k}\uparrow}.$$

  Acting on $c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle$, the second term becomes

$$c_{\mathbf{k}\uparrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle = c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle\,.$$

  Thus the second term is $+uv\,c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle$, which cancels exactly with the third term $-uv\,c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle$.
- The fourth term vanishes since $(c^{\dagger}_{-\mathbf{k}\downarrow})^2 = 0$. Therefore, $\gamma_{\mathbf{k}\uparrow}\,|\mathrm{BCS}\rangle_{\mathbf{k}} = 0$.

Second, the proof for $\gamma_{\mathbf{k}\uparrow}$.

- Similarly,

$$\gamma_{-\mathbf{k}\downarrow}\,|\mathrm{CP}\rangle_{\mathbf{k}} = \left(u\,c_{-\mathbf{k}\downarrow} + v\,c^{\dagger}_{\mathbf{k}\uparrow}\right)\left(u + v\,c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\right)|0\rangle$$
$$= u^2\,c_{-\mathbf{k}\downarrow}\,|0\rangle + uv\,c_{-\mathbf{k}\downarrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle + uv\,c^{\dagger}_{\mathbf{k}\uparrow}\,|0\rangle + v^2\,c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle\,.$$

- The first term vanishes because $c_{-\mathbf{k}\downarrow}\,|0\rangle = 0$.
- For the second term, anticommuting $c_{-\mathbf{k}\downarrow}$ past $c^{\dagger}_{\mathbf{k}\uparrow}$ gives

$$c_{-\mathbf{k}\downarrow}c^{\dagger}_{\mathbf{k}\uparrow} = -c^{\dagger}_{\mathbf{k}\uparrow}c_{-\mathbf{k}\downarrow}.$$

  Then $c_{-\mathbf{k}\downarrow}c^{\dagger}_{-\mathbf{k}\downarrow} = 1 - c^{\dagger}_{-\mathbf{k}\downarrow}c_{-\mathbf{k}\downarrow}$, and acting on $|0\rangle$ leaves

$$c_{-\mathbf{k}\downarrow}c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}\,|0\rangle = -\,c^{\dagger}_{\mathbf{k}\uparrow}\,|0\rangle\,.$$

  Thus the second term is $-uv\,c^{\dagger}_{\mathbf{k}\uparrow}\,|0\rangle$, which cancels the third term $+uv\,c^{\dagger}_{\mathbf{k}\uparrow}\,|0\rangle$.
- The last term vanishes because $(c^{\dagger}_{\mathbf{k}\uparrow})^2 = 0$. Therefore, $\gamma_{-\mathbf{k}\downarrow}\,|\mathrm{CP}\rangle_{\mathbf{k}} = 0$.

Finally, we extend the proofs to the full CP state $|CP\rangle = \prod_{\mathbf{q}} \left( u_{\mathbf{q}} + v_{\mathbf{q}}\, c^{\dagger}_{\mathbf{q}\uparrow} c^{\dagger}_{-\mathbf{q}\downarrow} \right) |0\rangle$. Fermionic operators on different $\mathbf{q} \neq \mathbf{k}$ anticommute, so $\gamma_{\mathbf{k}\sigma}$ can be moved through the product to act only on its sector, where the above proof shows the result is zero.

Therefore,

$$\gamma_{\mathbf{k}\sigma} |\mathrm{CP}\rangle = 0 \quad \forall\, \mathbf{k}, \sigma.$$

We now show the anticommutation rules.

- $\{\gamma_{\mathbf{p}\uparrow}, \gamma_{\mathbf{p}'\uparrow}\} = \tilde{u}_{\mathbf{p}} \tilde{u}_{\mathbf{p}'} \{c_{\mathbf{p}\uparrow}, c_{\mathbf{p}'\uparrow}\} + \tilde{v}_{\mathbf{p}} \tilde{v}_{\mathbf{p}'} \{c^{\dagger}_{-\mathbf{p}\uparrow}, c^{\dagger}_{-\mathbf{p}'\uparrow}\} + \tilde{u}_{\mathbf{p}} \tilde{v}_{\mathbf{p}'} \{c_{\mathbf{p}\uparrow}, c^{\dagger}_{-\mathbf{p}'\downarrow}\} - \tilde{v}_{\mathbf{p}} \tilde{u}_{\mathbf{p}'} \{c^{\dagger}_{-\mathbf{p}\downarrow}, c_{\mathbf{p}\uparrow}\}$. All the anticommutators on the right are 0. Those with both annihilation or both creation operators are 0 because similar-type operators always anti-commute. In the others, the operators have different spins from one another ($\uparrow$ and $\downarrow$). Hence $\{\gamma_{\mathbf{p}\uparrow}, \gamma_{\mathbf{p}'\uparrow}\} = 0$.
- A very similar argument can be used to show that $\{\gamma^{\dagger}_{\mathbf{p}\uparrow}, \gamma^{\dagger}_{\mathbf{p}'\uparrow}\} = 0$.
- Now consider $\{\gamma_{\mathbf{p}\uparrow}, \gamma^{\dagger}_{\mathbf{p}'\uparrow}\}$. When we expand in terms of the $c$ and $c^{\dagger}$ operators we can again ignore anticommutators between operators of opposite spin. This leaves $\{\gamma_{\mathbf{p}\uparrow}, \gamma^{\dagger}_{\mathbf{p}'\uparrow}\} = \tilde{u}_{\mathbf{p}} \tilde{u}^{*}_{\mathbf{p}'} \{c_{p\uparrow}, c^{\dagger}_{\mathbf{p}'\uparrow}\} + \tilde{v}_{\mathbf{p}} \tilde{v}^{*}_{\mathbf{p}'} \{c^{\dagger}_{\mathbf{p}\uparrow}, c_{\mathbf{p}'\uparrow}\} = (|u_{\mathbf{p}}|^2 + |v_{\mathbf{p}}|^2)\, \delta(\mathbf{p} - \mathbf{p}') = \delta(\mathbf{p} - \mathbf{p}')$ since $(|u_{\mathbf{p}}|^2 + |v_{\mathbf{p}}|^2) = 1$.

Finally, since we will need this in the next section, invert Eqs. (7) and (8) to obtain

$$c_{\mathbf{k}\uparrow} = u_{\mathbf{k}} \gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}} \gamma^{\dagger}_{-\mathbf{k}\downarrow}$$
$$c_{\mathbf{k}\downarrow} = u_{\mathbf{k}} \gamma_{\mathbf{k}\downarrow} - v_{\mathbf{k}} \gamma^{\dagger}_{-\mathbf{k}\uparrow}$$

### 2.1.2 Wick expansion with respect to $|CP\rangle$ states

If we use the Wick expansion of Eq. (6) we get

$$\begin{aligned}
H_{CP} = \sum_{p\sigma} \epsilon_p c^\dagger_{p\sigma} c_{p\sigma} - \kappa^2 \sum_{pk} [&: c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} :_{CP} \\
&+ \langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} \rangle_{CP} : c_{-p\downarrow} c_{p\uparrow} :_{CP} + : c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} :_{CP} \langle c_{-p\downarrow} c_{p\uparrow} \rangle_{CP} \\
&+ \langle c^\dagger_{k\uparrow} c_{p\uparrow} \rangle_{CP} : c^\dagger_{-k\downarrow} c_{-p\downarrow} :_{CP} + : c^\dagger_{k\uparrow} c_{p\uparrow} :_{CP} \langle c^\dagger_{-k\downarrow} c_{-p\downarrow} \rangle_{CP} \\
&- \langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} : c^\dagger_{-k\downarrow} c_{p\uparrow} :_{CP} - : c^\dagger_{k\uparrow} c_{-p\downarrow} :_{CP} \langle c^\dagger_{-k\downarrow} c_{p\uparrow} \rangle_{CP} \\
&+ \langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} \rangle_{CP} \langle c_{-p\downarrow} c_{p\uparrow} \rangle_{CP} + \langle c^\dagger_{k\uparrow} c_{p\uparrow} \rangle_{CP} \langle c^\dagger_{-k\downarrow} c_{-p\downarrow} \rangle_{CP} - \langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} \langle c^\dagger_{-k\downarrow} c_{p\uparrow} \rangle_{CP}]
\end{aligned} \tag{9}$$

It's straightforward to see that if $p_1 \neq p_2$ then $\langle CP | c^\dagger_{p_1\uparrow} c_{p_2\uparrow} | CP \rangle = 0$. Therefore the third line above can be written as a sum over $p, \sigma$ and combined into the kinetic energy term[4] , where it modifies $\epsilon_p$.

$$\begin{aligned}
H_{CP} = \sum_{p\sigma} \left( \epsilon_p - \kappa^2 \langle c^\dagger_{-p\sigma} c_{-p\sigma} \rangle_{CP} \right) c^\dagger_{p\sigma} c_{p\sigma} - \sum_{pk} [&: c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} :_{CP} \\
&+ \langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} \rangle_{CP} : c_{-p\downarrow} c_{p\uparrow} :_{CP} + : c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} :_{CP} \langle c_{-p\downarrow} c_{p\uparrow} \rangle_{CP} \\
&- \langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} : c^\dagger_{-k\downarrow} c_{p\uparrow} :_{CP} - : c^\dagger_{k\uparrow} c_{-p\downarrow} :_{CP} \langle c^\dagger_{-k\downarrow} c_{p\uparrow} \rangle_{CP} \\
&+ \langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} \rangle_{CP} \langle c_{-p\downarrow} c_{p\uparrow} \rangle_{CP} + \langle c^\dagger_{k\uparrow} c_{p\uparrow} \rangle_{CP} \langle c^\dagger_{-k\downarrow} c_{-p\downarrow} \rangle_{CP} - \langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} \langle c^\dagger_{-k\downarrow} c_{p\uparrow} \rangle_{CP}].
\end{aligned}$$

Next, we show that $\langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} = 0$. The simplest way to do this is to expand $c^\dagger_{k\uparrow} c_{-p\downarrow}$ in terms of the $\gamma$ operators.

$$c^\dagger_{k\uparrow} c_{-p\downarrow} = u_k u_p \gamma^\dagger_{k\uparrow} \gamma_{-p\downarrow} - v_k v_p \gamma_{-k\downarrow} \gamma^\dagger_{p\uparrow} - u_k v_p \gamma^\dagger_{k\uparrow} \gamma^\dagger_{p\uparrow} + u_k v_p \gamma_{-k\downarrow} \gamma_{-p\downarrow}. \tag{10}$$

When $p \neq k$, each monomial, $\mathcal{M}$ has a 0 contraction $\langle s | \mathcal{M} | s \rangle$ because we can anticommute the two ladder operators so that either, or both, of the 'bra' and 'ket' are annihilated. When $p = k$ the only monomial whose contraction could be non-zero is $v_k v_p \gamma_{-k\downarrow} \gamma^\dagger_{p\uparrow}$. However, that contraction is also 0 because the two operators have different spin, so anticommute and thus $\langle s | \gamma_{-p\downarrow} \gamma^\dagger_{p\uparrow} | s \rangle = 0$. This establishes that $\langle c^\dagger_{k\uparrow} c_{-p\downarrow} \rangle_{CP} = 0$ and similarly $\langle c^\dagger_{-p\downarrow} c_{p\uparrow} \rangle_{CP} = 0$. We use

[4] To be precise, we need to turn our momenta sums into integrals $\sum_{\mathbf{kp}} \rightarrow g^2(\epsilon_F) \int_{|\mathbf{p}|-|p_F|,|\mathbf{k}|-|k_F|,<\omega_D} d^3p d^3k$ where $g(\epsilon_F)$ is explained in section (3.2),

this to further simplify Eq. (9).

$$\begin{aligned} H_{CP} = &\sum_{p\sigma}\left(\epsilon_p - \kappa^2\langle c^\dagger_{-p\sigma}c_{-p\sigma}\rangle_{CP}\right)c^\dagger_{p\sigma}c_{p\sigma} - \kappa^2\sum_{pk}[:c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}c_{-p\downarrow}c_{p\uparrow}:_{CP} \\ &+ \langle c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}\rangle_{CP} : c_{-p\downarrow}c_{p\uparrow}:_{CP} + : c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}:_{CP}\langle c_{-p\downarrow}c_{p\uparrow}\rangle_{CP} \\ &+ \langle c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}\rangle_{CP}\langle c_{-p\downarrow}c_{p\uparrow}\rangle_{CP} + \langle c^\dagger_{k\uparrow}c_{p\uparrow}\rangle_{CP}\langle c^\dagger_{-k\downarrow}c_{-p\downarrow}\rangle_{CP}] \end{aligned}$$

This isn't quite the mean-field Hamiltonian but it's close. We have an expression that is based on a general $|CP\rangle$ state. We define the reduced Hamiltonian by dropping the quartic normal-ordered term (assumed – then later shown – to be perturbatively small) and the $c$-numbers (they have no effect on energy differences). We'll also redefine (a.k.a. "renormalize") the energies $\epsilon_p$ to $\epsilon'_p = \epsilon_p - \kappa^2\langle c^\dagger_{-p\sigma}c_{-p\sigma}\rangle_{BCS}$. We can compute the contraction by expanding $c^\dagger_{-p\sigma}c_{-p\sigma}$ similarly to what we did with Eq. (10).

$$\begin{aligned} c^\dagger_{-p\uparrow}c_{-p\uparrow} &= u_pu_p\gamma^\dagger_{-p\uparrow}\gamma_{-p\uparrow} + v_pv_p\gamma_{p\downarrow}\gamma^\dagger_{p\downarrow} + u_pv_p\gamma^\dagger_{-p\uparrow}\gamma^\dagger_{p\downarrow} + u_pv_p\gamma_{p\downarrow}\gamma_{-p\uparrow} \\ c^\dagger_{p\downarrow}c_{p\downarrow} &= u_pu_p\gamma^\dagger_{p\downarrow}\gamma_{p\downarrow} + v_pv_p\gamma_{-p\uparrow}\gamma^\dagger_{-p\uparrow} - u_pv_p\gamma^\dagger_{p\downarrow}\gamma^\dagger_{-p\uparrow} - u_pv_p\gamma_{-p\uparrow}\gamma_{p\downarrow} \end{aligned} \tag{11}$$

When we compute the contractions of the above quadratic terms, the only surviving contribution is $\langle s|v_pv_p\gamma_{p\sigma}\gamma^\dagger_{p\sigma}|s\rangle = v_p^2$, so $\epsilon'_p = \epsilon_p - \kappa^2v_p^2$. Later, we will estimate the size of $\kappa^2v_p^2$ relative to $\epsilon_p$ and will see that it is small.

We'll also define the sum of the pure contractions to be a constant $E_0 = \sum_{pk}\langle c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}\rangle_{BCS}\langle c_{-p\downarrow}c_{p\uparrow}\rangle_{BCS} + \langle c^\dagger_{k\uparrow}c_{p\uparrow}\rangle_{BCS}\langle c^\dagger_{-k\downarrow}c_{-p\downarrow}\rangle_{BCS}$.

Then we end up with the reduced Hamiltonian.

$$\begin{aligned} H_R = &\sum_{p\sigma}\epsilon'_pc^\dagger_{p\sigma}c_{p\sigma} \\ &+ -\kappa^2\sum_{pk}[\langle c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}\rangle_{CP} : c_{-p\downarrow}c_{p\uparrow}:_{CP} + : c^\dagger_{k\uparrow}c^\dagger_{-k\downarrow}:_{CP}\langle c_{-p\downarrow}c_{p\uparrow}\rangle_{CP}] + E_0 \end{aligned} \tag{12}$$

## 2.2 The ground state $|R\rangle$ of $H_R$ is a $|CP\rangle$ state

In this section, I will show that for $H_R$, the lowest-energy state is a $|CP\rangle$ state. Remember, the most general Cooper-pair state isn't necessarily a $|CP\rangle$ state, so this demonstration is required. The subsections below were

written by ChatGPT and describe a system of single particles rather than pairs. However, the principle is the same.

Importantly, when we add a quartic term in the Hamiltonian, the argument which follows does not show that the ground state is of the type $|CP\rangle$.

**Setup**

Consider a fermionic system with Fock vacuum $|0\rangle$ and two possible momentum states $k_1, k_2$. The corresponding one-particle states are

$$|k_1\rangle = a^\dagger_{k_1}|0\rangle, \qquad |k_2\rangle = a^\dagger_{k_2}|0\rangle.$$

A general normalized state has the form

$$|\psi\rangle = \alpha|0\rangle + \beta|k_1\rangle + \gamma|k_2\rangle,$$

with $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$.

$$H_R = H_1 + H_2,$$

where $H_1$ acts only on the $\{|0\rangle, |k_1\rangle\}$ sector, and $H_2$ acts only on the $\{|0\rangle, |k_2\rangle\}$ sector.[5] That is, $[H_1, H_2] = 0$ and each $H_i$ preserves the corresponding sector.

**Step 1: Eigenstates of each block**

Let the normalized ground state of $H_1$ be

$$H_1|g_1\rangle = e_1|g_1\rangle,$$

where $|g_1\rangle \in \text{span}\{|0\rangle, |k_1\rangle\}$.

Similarly, let the normalized ground state of $H_2$ be

$$H_2|g_2\rangle = e_2|g_2\rangle,$$

with $|g_2\rangle \in \text{span}\{|0\rangle, |k_2\rangle\}$.

[5]Since the CP annihilation and creation operators can be written as linear combinations of $c$ and $c^\dagger$, all with the same momentum index, then the normal ordered quadratic product of $c$ and/or $c^\dagger$, is a sum of quadratic combinations of the CP annihilation and creation operators in the same momentum sector.

**Step 2: Tensor product structure**

Since $H_1$ and $H_2$ commute and act on distinct subspaces, the Hilbert space decomposes as

$$\mathcal{H} = \mathcal{H}_{k_1} \otimes \mathcal{H}_{k_2}.$$

An arbitrary eigenstate of $H$ must then be of the form

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle,$$

with eigenvalue $E = E_1 + E_2$ if

$$H_1|\psi_1\rangle = E_1|\psi_1\rangle, \quad H_2|\psi_2\rangle = E_2|\psi_2\rangle.$$

**Step 3: Ground state minimization**

To minimize the energy, we must choose $|\psi_1\rangle = |g_1\rangle$ and $|\psi_2\rangle = |g_2\rangle$. Thus the ground state of the full Hamiltonian is

$$|\Psi_0\rangle = |g_1\rangle \otimes |g_2\rangle,$$

with ground-state energy

$$E_0 = e_1 + e_2.$$

**Conclusion**

For quadratic Hamiltonians that decompose into independent commuting block terms $H_i$, the ground state always factorizes into the product of block ground states:

$$|\Psi_0\rangle = \prod_i |g_i\rangle.$$

# 3 The mean-field Hamiltonian

Next we will define the mean-field Hamiltonian $H_{BCS}$ by setting the $|CP\rangle$ state used for normal-ordering $H_R$, to be $H_R$'s ground state $|R\rangle$, which at this point we don't yet know. However, we can find $|R\rangle$ by noting that it

satisfies a self-consistency condition.[6] Although we could pose the $|R\rangle$ self-consistency condition as an equation to be solved, it turns out to be easier to find a solution by ansatz. Incidentally, part of the self-consistency for $H_{BCS}$ is the assumption that the quartic term is negligible when $|CP\rangle$ is chosen to be $|R\rangle$. Later we will redefine $|R\rangle$ to be $|BCS\rangle$.

## 3.1 Eigenstates including $|R\rangle$, of $H_R$

Now follow Annett. First, we know that the $H_R$ (ground) state is a Cooper pair state that we will write[7] as $|R\rangle = \prod_k \left(\bar{u}_k + \bar{v}_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}\right)$. The normalization condition imposes the equality $|\bar{u}_k|^2 + |\bar{v}_k|^2 = 1$.

Define $\Delta = \kappa^2 \sum_p \langle c_{-p\downarrow} c_{p\uparrow}\rangle_{CP}$. This is a function of the parameters $u_p$ and $v_p$ that appear in the definition (above) of $|CP\rangle$. Then $\Delta = \kappa^2 \sum_p u^*_p v_p$. We can then rewrite $H_R$ as

$$H_R = \sum_{p\sigma} \epsilon'_p c^\dagger_{p\sigma} c_{p\sigma} - \sum_p [\Delta^* : c_{-p\downarrow} c_{p\uparrow} :_{CP} + \Delta : c^\dagger_{p\uparrow} c^\dagger_{-p\downarrow} :_{CP}] + E_0$$

**We call this the mean-field Hamiltonian for the BCS model.** In matrix form, we can write this as

$$H_R = \sum_p : \begin{pmatrix} c^\dagger_{p\uparrow} & c_{-p\downarrow} \end{pmatrix} \begin{pmatrix} \epsilon'_p & -\Delta \\ -\Delta^* & -\epsilon'_p \end{pmatrix} \begin{pmatrix} c_{p\uparrow} \\ c^\dagger_{-p\downarrow} \end{pmatrix} :_{CP} + E_0$$

We then diagonalize this matrix – that's the Bogoliubov procedure. We end

[6] This business of self-consistency is, I believe, ill-defined. Here are two examples of equations. Equation A is $x = x$. Equation B is $x^2 = -3x - 2$. Both equations are solvable. The first – a tautology – is solved by **any** $x$. The second is solved by two potentially different values of $x$. Both equations are examples of self-consistency conditions. But a "self-consistency" method of solution proceeds differently than the usual algebraic method which we used in deriving the above results. In the "self-consistency" method, you start with an ansatz, and then you either verify that the ansatz is true (i.e.., the equation is solved)– or if the ansatz involves some free parameters, you adjust the parameters until the ansatz is true. For Equation A, you could take the ansatz "$x = 3$". That is true, so we have a self-consistent solution. For Equation B, you could take the ansatz $x = -2$. That is also true, so is a self-consistent solution. However, in both cases using the ansatz method, the solutions aren't unique. With the "usual algebraic" method, you would find all the solutions. With the "ansatz" method you might miss solutions.

[7] as usual in this section, I'm suppressing spin and also not indicating that **k** is a vector.

up with

$$H_R = \sum_p : \begin{pmatrix} b^\dagger_{p\uparrow} & b_{-p\downarrow} \end{pmatrix} \begin{pmatrix} E_p & 0 \\ 0 & -E_p \end{pmatrix} \begin{pmatrix} b_{p\uparrow} \\ b^\dagger_{-p\downarrow} \end{pmatrix} :_R + E_0'$$

where $b_{p\uparrow} = \tilde{u}_p c_{p\uparrow} - \tilde{v}_p c^\dagger_{-p\downarrow}$, $b^\dagger_{-p\downarrow} = \tilde{v}^*_p c_{p\uparrow} + \tilde{u}^*_p c^\dagger_{-p\downarrow}$, and $|\tilde{u}_p|^2 = \frac{1}{2}\left(1 + \frac{\epsilon'_p}{E_p}\right), |\tilde{v}_p|^2 = \frac{1}{2}\left(1 - \frac{\epsilon'_p}{E_p}\right), \tilde{u}^*_p \tilde{v}_p = \frac{\Delta}{2E_p}$ where $E_p = \sqrt{\epsilon'^2_p + |\Delta|^2}$. Notice that $\epsilon'_p$ is a function of $v_p$ so in principle, we have to solve a quartic equation in order to obtain $v_p$ as a function of $\epsilon_p$ and $\Delta$. However, as shown earlier, the correction to $\epsilon_p$ is small, so we can approximate $\epsilon'_p$ with $\epsilon_p$.

One observation turns out to be very useful in making our approximations later. When $\epsilon'_p \gg \Delta$, $v_p \approx 0$. Also when $-\epsilon'_p \ll -\Delta$, $u_p \approx 0$. We will use this fact to bound sums of terms with coherence factors $u_p v_p$.

Also, the constant $E_0$ was modified to $E'_0$ to account for further changes of operator-order in diagonalizing the matrix. $E'_0$ is known as the ***condensation energy***. We notice at this time that the $b$ and $b^\dagger$ operators are defined in terms of the $c$ and $c^\dagger$ operators using exactly the linear combinations of Eqs. (7) and (8), if we substitute $\tilde{u}$ and $\tilde{v}$ for $u$ and $v$. Therefore the $b$ and $b^\dagger$ operators satisfy the standard anticommutation rules for annihilation and creation operators, and also the $b_{\mathbf{p}\uparrow}$ and $b_{-\mathbf{p}\downarrow}$ operators annihilate the state $|\tilde{R}\rangle = \prod_k \left(\tilde{u}_k + \tilde{v}_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}\right)$.

Noting that $: b_{-p\downarrow} b^\dagger_{-p\downarrow} := - : b^\dagger_{-p\downarrow} b_{-p\downarrow} :$, and that sums over $p$ are the same as sums over $-p$, we end up with

$$H_R = \sum_p E_p \left( : b^\dagger_{p\uparrow} b_{p\uparrow} :_{CP} + : b^\dagger_{p\downarrow} b_{p\downarrow} :_{CP} \right) + E'_0. \tag{13}$$

Now we're ready to impose the self-consistency condition. So far, $\tilde{u}$ and $\tilde{v}$ are functions of $u, v$. In particular, normal-ordering with respect to the CP state (defined by $u, v$) is not necessarily the same as normal-ordering with respect to $|\tilde{R}\rangle$ (defined by the variables $\tilde{u}, \tilde{v}$ both of which are functions of $u, v$).

The self-consistency condition for the ground state of $H_R$, is that it is annihilated by $b_k$'s where $\tilde{u} = u$ and $\tilde{v} = v$ (that is our ansatz). To see that such a state is a ground state, note that normal ordering is with respect to the

$\tilde{u}, \tilde{v}$ operators and therefore, since those are the $b$ operators, we can dispense with the normal-ordering on the RHS of Eq. (13), leaving us with

$$H_R = \sum_{|\mathbf{p}|-|p_F|<\omega_D} E_p \left( b^\dagger_{p\uparrow} b_{p\uparrow} + b^\dagger_{p\downarrow} b_{p\downarrow} \right) + E'_0. \tag{14}$$

This is the standard form for the Hamiltonian of free fermions with energy[8] $E_p$. Since $|R\rangle$ is annihilated by $b_{\mathbf{p}\uparrow}$ and $b_{-\mathbf{p}\downarrow}$, it is a ground-state eigenvector of $H_R$. Moreover, any other state in the Hilbert space is obtained by consecutive applications of the creation operator $b^\dagger$ and as usual for free particles, the average energy of such a state would be larger than the ground state energy. So if we can find a solution to the self-consistency condition, it is unique.

In summary, if we take the ansatz $\tilde{u} = u$ and $\tilde{v} = v$, then the Cooper-pair state defined by the parameters $u, v$ is the ground state of $H_R$. We can therefore write $H_R$ using normal-ordering with respect to that state. We will refer to this Hamiltonian as the mean-field Hamiltonian $H_{BCS}$ and this ground state as $|BCS\rangle$.

## 3.2 Solving for the ansatz

We can now solve for $\tilde{u}$ and $\tilde{v}$ following Lancaster section 44.3. Recall that $\Delta = \kappa^2 \sum_{|\mathbf{p}|-|p_F|<\omega_D} u^*_p v_p = \kappa^2 \sum_{|\mathbf{p}|-|p_F|<\omega_D} \tilde{u}^*_p \tilde{v}_p$ where we've employed our ansatz, and also have re-enstated the momenta-bounds in the summation. Also recall from the Bogoliubov solution, that $\tilde{u}^*_p \tilde{v}_p = \frac{\Delta}{2E_p}$. Substituting into above, we end up with

$$\Delta = \kappa^2 \sum_{|\mathbf{p}|-|p_F|<\omega_D} \frac{\Delta}{2E_p}$$

This is Lancaster's Eq. (44.24). We can then solve this as shown in Lancaster Eqs. (44.25) - (44.27) to obtain $\sinh^{-1} \frac{\omega_D}{\Delta} = \frac{1}{\Lambda}$, where $\Lambda = \kappa^2 g(\epsilon_F)$, and $g(\epsilon')$ is the density of states at energy $\epsilon'$ (assumed to be approximately constant when $\epsilon'$ lies in the Debye band around $\epsilon_F$.) When $|\Delta| << \omega_D$, this becomes

$$|\Delta| \approx 2\omega_D e^{-\frac{1}{\Lambda}}.$$

[8] As far as I can tell, there's nothing that would have stopped us from taking $E_p$ to be negative (remember, we derived that $E_p$ was a square-root). However, I believe that the net effect of doing this would be to exchange the definitions of creation and annihilation operators – or equivalently exchange definitions of holes and electrons. This may need to be double-checked.

With this, we obtain all the parameters defining both the BCS mean-field Hamiltonian ground state, and the eigen-energies of the Cooper-pair states of the theory.

$$E_p = \sqrt{\epsilon_p'^2 + |\Delta|^2}$$
$$|\tilde{u}_p| = \sqrt{\frac{1}{2}\left(1 + \frac{\epsilon_p'}{E_p}\right)}$$
$$|\tilde{v}_p| = \sqrt{\frac{1}{2}\left(1 - \frac{\epsilon_p'}{E_p}\right)}$$
$$\tilde{u}_p^* \tilde{v}_p = \frac{\Delta}{2E_p}$$

Note that the $u$ and $v$ parameters are defined only up to a phase. To see this, write $u_p = e^{i\alpha_p}|u_p|$ and $v_p = e^{i\beta_p}|v_p|$. Then $u_p v_p^* = e^{i(\alpha_p - \beta_p)}|u_p||v_p|$. This phase is uniquely given by the above equation $\tilde{u}_p^* \tilde{v}_p = \frac{\Delta}{2E_p}$ and therefore the overall phase ambiguity simply amounts to the fact that the BCS ground state is only defined up to a phase.

One final topic regarding the derivation of the mean field Hamiltonian. In section 44.3, Lancaster derives the BCS ground state without all the artillery of Wick ordering and so on. In that section, he is only able to obtain the ground state and ground state energy, but doesn't have a mechanism for obtaining the excited states, nor the energy gap. The derivation of section 44.3 starts with the Hamiltonian $H_{CP}$ shown in Eq. (6) above. Ultimately, apart from renormalization of $\epsilon_p$ and the introduction of $E_0$, this is the same as our BCS Hamiltonian **except for** the term $V_{NO} \equiv -\kappa^2 \sum_{pk} :\, c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger c_{-p\downarrow} c_{p\uparrow} \,:_{BCS}$. What is noteworthy, is that the minimization procedure of Lancaster section 44.3 leads to the same ground state as what we obtained above when ignoring the normal-ordered term $V_{NO}$. That result is good to all orders in $\kappa$ but we will also show it later in lowest-order perturbation theory.

## 3.3 The superconducting Schrodinger equation

Before beginning our study of the energy gap and quasi-particle spectrum, I want to briefly revisit the superconducting Schrodinger equation, which we used to explain some of the key features of superconductivity including 0 resistivity and the Meissner effect. The summary of the Schrodinger-equation approach is this: We write a classical Schrodinger equation describing a par-

ticle in the presence of an electromagnetic field. We then use this equation to develop the time evolution of a system of $N$ particles. In order to justify this, I've said that we can make an identification between fields and wavefunctions and that this identification is based on coherent states. It was my intention to clarify how this is done when we're dealing with fermion-pairs.

I've previously discussed how this is done with bosonic systems. In those systems, we defined coherent states $|\alpha\rangle$ that have the property $a_k|\alpha\rangle = \alpha(k)|\alpha\rangle$ and which are also eigenstates of the field operators. The Euler-Lagrange equations for the field operators then become equations for complex fields. Lancaster argues that the operator $P_k^\dagger = c_{\boldsymbol{k}\uparrow}^\dagger c_{-k\downarrow}^\dagger$ sort of behaves like a boson creation operator. But then Lancaster goes on to show that this pair operator doesn't actually satisfy the bosonic commutation relation $\left[P_k, P_{k'}^\dagger\right] = \delta(\boldsymbol{k} - \boldsymbol{k'})$. Nevertheless, Lancaster defines a state, which I'll call $|CP\rangle$, that resembles the definition of a bosonic coherent state, namely $|CP\rangle = \prod_p C_p e^{\alpha_p P_p^\dagger|0\rangle}$, where $C_p$ is a normalization constant. As we'll see shortly, this state is not an eigenvector of the pair creation operator. So what gives?

First notice that, since $P$ is composed of fermion operators, that $P_k^\dagger P_k^\dagger = 0$. Then, when we Taylor expand the exponential operator that appears in the definition of $|CP\rangle$, the only terms in the expansion that survive are either $0^{th}$ or $1^{st}$ order in $P_k^\dagger$. We obtain $|CP\rangle = \prod_k (u_k + v_k P_k^\dagger)|0\rangle$ where $\alpha_k = \frac{v_k}{u_k}$ and the coefficients $u_k$ and $v_k$ are set so that the state is normalized. We use this representation of $|CP\rangle$ in the next sections.

What we would like to see, if $|CP\rangle$ is truly like a bosonic coherent state, is that $P_k|CP\rangle = \alpha_k|CP\rangle$. That turns out not to be correct. However, it's not a necessary requirement for developing the Schrodinger-like theory of superconducting. I've spent a great deal of time (weeks?) trying to massage this idea into something rigorous. The best I can do still has significant holes and I've decided – reluctantly – to abandon my approach, realizing that there are systematic treatments in the literature. In particular, Fetter and Walecka, in Chapters 51 and 52, are particularly careful and appear to fill in the many open questions which I wasn't able to answer. I'll admit that, as of this moment (November 2025), I haven't studied these chapters in any depth. Perhaps if I did, I would be able to connect back to what I'd been attempting to do – which was to show that the Schrodinger equation could be justified formally. Unfortunately, with the Fetter and Walecka treatment, it's easy to lose sight of the simple boson-like Schrodinger ap-

proach (which I borrowed mostly from Feynman). Presumably, one should be able to justify the Schrodinger approach by some approximations which correspond intuitively to a theory that looks like a boson theory. I think I've gotten myself into trouble by attempting to make this correspondence without clearly understanding how to go about the correspondence between formalism and intuition. The intuition presumably applies to any system where fermions strongly couple together to form integer-spin systems. In that intuition, the strong-binding has the effect at distances far from the bound system (i.e. long wavelength), of washing out any fermionic aspects of the bound state. So if I want to make the connection with the formalism, I would need to somehow separate the large-frequency from the low-frequency behaviors. The problem should probably be formulated by starting with a "pair-field" $\hat{\Psi}_\uparrow(\boldsymbol{r})\hat{\Psi}_\downarrow(\boldsymbol{r}')$, then rewrite both position vectors and momentum vectors in terms of center-of-mass and relative-coordinates. ChatGPT has offered a lot of analysis of this problem, but perhaps I misguided ChatGPT by suggesting a starting point to be approximated. Here's a concrete issue I've run into. In connecting the Schrodinger equation to our basic superconducting field theory, I've assumed that the "wavefunction" must somehow correspond to a field that is a Fourier-transform of pair-creation operators. Then the question becomes "what complex-valued function should we take to represent that field?" One possibility – the one that ChatGPT has analyzed – is that we take the BCS-state expectation value of all field operators. That seems reasonable, since at 0 temperature, the state of the system is the BCS state (the lowest-energy state). However, the BCS state consists of a condensate of pairs each with 0 COM (center-of-mass) momentum. Then moving pairs will always have 0 expectation value in the BCS state. But when we speak of currents, we are explicitly interested in moving pairs, so the BCS vacuum must be the wrong state in which to take expectation values. In the treatment by Fetter and Walecka, they look at the ensemble average of operators and they also set up a self-consistency equation to go along with their Schrodinger-like differential equation. I don't know if that self-consistency equation ultimately "selects" the correct ensemble state but perhaps it does – and if so, could fix the problem caused by starting with the BCS state.

Another thing that I haven't tried, is to take the Hamiltonian considered earlier in this document (the one with the contact potential between particle pairs) and add to it, the EM interaction. Since the EM field is here treated as an external field, the relevant terms of the Hamiltonian are all quadratic in the fermion fields, so it should be possible to find the eigenstates for the corresponding mean-field theory (i.e., one where we ignore the normal-ordered potential term). I don't know what would happen. It doesn't sound

hard to do and maybe the Schrodinger equation would emerge from that. But if it did, it seems to me that this approach would have appeared in one of the references I've been using.

So in summary, the Schrodinger equation approach seems reasonable – and seems like it leads to the correct superconductivity physics – if we follow the intuitive idea that the Cooper-pair field should be approximately bosonic, and that there is a coherent-state structure with similar properties to bosonic coherent states. For a more complete derivation of superconductivity physics and the Ginsburg-Landau theory, we may need a different starting point which uses the Euler-Lagrange equations (which resemble the Schrodinger equation). Fetter and Walecka provide such a derivation.

## 3.4 Perturbation theory

The objective of this entire section on perturbation theory, is to show that the mean-field Hamiltonian in Lancaster and elsewhere, gives a reliable approximation to the Cooper-pair spectrum. We approach this by taking the mean-field Hamiltonian as the $0^{\text{th}}$ order term in a perturbative expansion of the full Hamiltonian, and then comparing the size of perturbations to the values of the energies they are modifying. It should suffice to consider only the first-order perturbations, but we will also examine the second-order perturbations to convince ourselves that consecutive terms in the series indeed are suppressed by powers of the perturbative parameter (which we'll introduce later as $\Lambda$). Also, some of the perturbative terms have coefficients inversely proportional to the total number of particles in the Debye band. Since that is a very large number (only a few orders of magnitude smaller than Avogadro's number), such perturbative terms can be regarded as essentially 0 (but we'll write $\mathcal{O}(\frac{1}{\mathcal{N}})$). These terms are often described by saying that they are phase-space diluted.

Although I've looked in all sorts of references, and used ChatGPT, I've found that most of these do only a cursory job of perturbation theory. The one text that goes into the appropriate level of detail, is Fetter and Walecka Quantum Theory of Many-Particle Systems. I will refer to this book from time to time.

The mean-field Hamiltonian is an approximation to the two-body interaction Hamiltonian of Eq. (1). Here are the various terms that have been left out

of the approximation:

- The normal-ordered quartic term of Eq. (9), $V_{NO} = -\kappa^2 \sum_{pk} : c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} :_{BCS}$.
- The $q$ dependence of $V$ (by which I will generally also mean "the general spin dependence of $V$"), namely $V'_Q = -\frac{1}{2} \sum'_{\substack{pqk \\ \sigma_1=\sigma_3\sigma_2=\sigma_4}} V_q c^\dagger_{(p-q)\sigma_1} c^\dagger_{(k+q)\sigma_2} c_{\boldsymbol{k}\sigma_3} c_{p\sigma_4}$, where the symbol $\sum'$ means that we should exclude momenta combinations already counted in $H_{CP}$. The mean-field Hamiltonian sets $q = 0$ and $p = -k$. When we explore this contribution later, we will set $V_q = \kappa^2$. For a more accurate perturbative analysis, we would need to analyze the momentum-dependence of $V_q$ – and that would require a proper treatment of the phonon-electron electromagnetic interactions. We won't consider this further in this note.

  The $q$ dependence also should extend $V_{NO}$ to $V_{QNO} = -\frac{1}{2} \sum_{\substack{pqk \\ \sigma_1=\sigma_3\sigma_2=\sigma_4}} V_q : c^\dagger_{(p-q)\sigma_1} c^\dagger_{(k+q)\sigma_2} c_{k\sigma_3} c_{p\sigma_4} :_{BCS}$. We recover $V_{NO}$ by restricting the sum to $k = -p$, setting $\sigma_1 = \sigma_4 = \uparrow$, $\sigma_2 = \sigma_3 = \downarrow$ and then relabeling the summation indices. Notice that in the definition of $V_{QNO}$, I have not excluded any momenta from the sum. If we wanted to show how the $q$-dependent normal-ordered quartic term perturbed the correspond term from $H_{CP}$, we would need to exclude certain momenta in order to avoid double-counting. However, it will be easier for us to simply deal directly with the full (no momenta excluded) quartic term. That's what we'll do later.
- $V_{ei}$ – the electron-ion interactions – other than for the implicit dependence captured in the four-body interaction. We will not consider $V_{ei}$ for now. See Fetter and Walecka for further information.
- The separation of normal and Cooper-pair states for electrons in the Debye band. We also won't consider this separation for now, nor the value of the Debye frequency.

The key parameters of a perturbation expansion are $\Lambda = \kappa^2 g(\epsilon_F)$ and $\frac{1}{N}$. As a reminder, in general for QM, when $H = H_0 + H'$ (when "$H'$" is small), the eigenvalues of $H$ are obtained to the first two orders of a perturbation expansion in $H'$ by

- First-order correction:

$$E^{(1)}_n = \langle n^{(0)} | H' | n^{(0)} \rangle$$

This is known as the Born term.

- Second-order correction:

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left|\langle m^{(0)}|H'|n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}}$$

where $H'$ is the perturbing Hamiltonian, $|n^{(0)}\rangle$ and $E_n^{(0)}$ are the unperturbed eigenstates and energies.

When we compute the various inner products, we will use identities of the following sort:

- Use the fact that $|BCS\rangle$ is the 'vacuum' state annihilated by the $b$ annihilation operators

$$b_{r\sigma}|BCS\rangle = \langle BCS|b_{r\sigma}^\dagger = 0$$

- Use canonical anticommutation relations to simplify expressions involving several operators – including expressions where excited states are described as products of creation operators acting on $|BCS\rangle$. So, as an example, consider a state $|s\rangle = b_{p\sigma}^\dagger|BCS\rangle$ and operate on it with the operator $b_{p'\sigma'}$.

$$b_{p'\sigma'}b_{p\sigma}^\dagger|BCS\rangle = -b_{p\sigma}^\dagger b_{p'\sigma'}|BCS\rangle + \delta_{p,p'}\delta_{\sigma,\sigma'}|BCS\rangle = \delta_{p,p'}\delta_{\sigma,\sigma'}|BCS\rangle,$$

where we've used the fact that $b_{p'\sigma'}$ annihilates the BCS vacuum. We generically refer to the kind of step used in the above equality, as a "contraction". Notice also in this example and elsewhere, that we are treating the momenta as discrete quantities. If we were to treat the momenta as continuous quantities we would need to employ integrals and a density function. Examples are given elsewhere of how this is done.

Final overall comment: As we'll quickly see, perturbation theory becomes unwieldy fairly quickly when expressing all calculations using annihilation and creation operators. For accurate and easier calculations, it's necessary to restructure and re-use intermediate results. The most systematic way (that I know of) is to schematically represent the various steps using Feynman diagrams. Not all calculations can be most effectively done in this way. See Fetter and Walecka for a comprehensive treatment of all this.

### 3.4.1 Why BCS doesn't pop out of free-electron perturbation theory

One might ask the question "why should we begin with the mean-field Hamiltonian and then use perturbation theory to compute the corrections to that Hamiltonian?" Couldn't we have started with the free Hamiltonian (no Cooper pairs or BCS vacuum) and treated all interactions (including the interactions of the mean-field Hamiltonian) as perturbations?

I believe this question is covered in Fetter and Walecka but I haven't examined that section. Instead I'll state my opinion based on experience. The BCS theory has the feature that if we start with a free Hamiltonian ($H_0 = \sum_{p\sigma} \frac{p^2}{2m} a^\dagger_{p\sigma} a_{p\sigma}$) and add the interaction potential, then if we do perturbation theory around that free Hamiltonian, the terms either blow up or the theory becomes unstable. I haven't checked that this is true, but I think there are similarities with the Higgs mechanism, where one is led to the conclusion that the correct starting Hamiltonian (corresponding to a "true" ground state) is based on a shifted starting point (a point on the rim of the "Mexican hat"). In the case of BCS, the correct starting point for our mean-field Hamiltonian is based on Cooper pairs. The $0^{th}$-order energy eigenvalues turn out to be a non-perturbative function of $\Lambda$ (by which I mean a function without a convergent perturbation series around $\Lambda = 0$). In particular, the energy gap (which appears in all the energy eigenvalues) $\Delta$ satisfies $|\Delta| \approx 2\omega_D e^{-\frac{1}{\Lambda}}$. An alternate approach might be that we start with the original "divergent" perturbation series and extract pieces at each order, which add up in a systematic way, as a geometric series with a sum such as $\frac{1}{1-\lambda^2}$ – a function with a pole at $\lambda = 1$, but which can be extended to larger values of $\lambda$ beyond the radius of convergence of the original theory. I'm not sure this actually happens in super-conductivity but see Fetter and Walecka to find out.

### 3.4.2 Proper normalization of sums

This section addresses an important intuition that seems to be evasive in the notation I've been using. To set the stage for this discussion I'll focus on the quartic term $V_{NO} = -\kappa^2 \sum_{pk} a^\dagger_p a^\dagger_k a_k a_p$. This is a double-sum which suggests that in general, there are $\mathcal{N}^2$ terms, where $\mathcal{N}$ is the total number of single-particle states (aka the single-particle phase space) in the Debye band. $V_{NO}$ therefore looks as though it should be $\mathcal{O}(\kappa^2 \mathcal{N}^2)$. However, as we'll see

shortly, when taking $\langle n^{(0)}|V_{NO}|n^{(0)}\rangle$, the double-sum is reduced to the order of the number of Cooper pairs in the state $|n^{(0)}\rangle$ So this suggests that the Born term is suppressed by a factor of $\mathcal{N}$ relative to the $0^{\text{th}}$-order energies.

What follows is an attempt to quantify this idea. I'll begin by admitting that this question plagued me for over a week and that I devoted hours and hours chasing it down in ChatGPT, then textbooks and papers. I'm not especially confident in my conclusions, mostly because sources are sloppy about the precise meaning of terms like "density" and "interaction strength". Nevertheless, I will present my current best understanding of the situation.

Of critical importance is the realization that some parameters are physically meaningless because they can't be directly measured. This is the insight that is key to understanding renormalization techniques. In the literature, parameters that can't be directly measured are known as "bare parameters". In superconducting field theory (i.e., a theory with the occupation number – or ladder operator – representation), parameters like $\kappa^2$ and $g(\epsilon_F)$ are examples of bare parameters. (For example, when we say that $g$ is a density, we need to resolve the question "density with respect to what?".) Regardless of how they are presented, the two of them can't be dis-entangled from one another. In what follows, I'll discuss the notion of explicitly exposing the total number of particles (alternatively I could have chosen the total volume). Ultimately, measurable results won't depend on this, so if I made an inappropriate definition, that won't matter in the end. However, for guessing in advance the order-of-magnitude of certain quantities, it's convenient to pick our best guess for introducing the number of particles as a "renormalizing" factor. One way in which condensed-matter physics differs from elementary particle physics, is that the total number of particles is, in fact, finite and measurable. So in principle, one might be able to measure some of the bare parameters, although the experiments of most interest, are ones which combine parameters into a measured quantity which should be independent of $N$. If our theoretical instincts are right, then measurements will confirm that those measured parameters are independent of system-size.

With all this in mind, we turn to the question of *intensive* versus *extensive* quantities. We want the total system energy to be extensive – that is, the energy should be proportional to either $\mathcal{V}$ (the volume) or $\mathcal{N}$ (the total number of particles). Different authors choose different conventions (and some authors don't bother mentioning either one). Since the double-sum of $V_{NO}$ has $\mathcal{N}^2$ terms, the coefficient should – in order for the whole thing to scale as $\mathcal{N}$ – be of the form $\frac{\lambda_0}{\mathcal{N}}$ where $\lambda_0$ is an "intensive" interaction

strength (characteristic, for example, of a single two-electron interaction). I will therefore set $\kappa^2 = \frac{\lambda_0}{\mathcal{N}}$. When we compare this to the Cooper-pair energies – all of which exceed $\Delta$ – immediately see that the perturbation ratio is $-\frac{\lambda_0}{\mathcal{N}\Delta}$ and since $\mathcal{N}$ is huge, the ratio is very small – as hoped.

One last – but very important – remark about our conventions. We set the dimensionless coupling $\Lambda = \frac{\lambda_0}{\mathcal{N}} g(\epsilon_F) = \kappa^2 g(\epsilon_F)$ where the density $g(\epsilon_F)$ is defined to be the total number of electrons per unit energy. Since $g$ is an extensive quantity (proportional to $\mathcal{N}$) then $\Lambda$ is intensive (independent of number of particles). But as mentioned in the previous paragraph, things are actually the other way around. $\Lambda$ is observed (and even predicted from first principles – but that's outside of the scope of the current information we've introduced about $\kappa^2$) to be intensive and from there, we are motivated to make the assignments of the renormalized parameters.

We could instead use $\Lambda$ as the physically relevant interaction strength. Then we write $\kappa^2 = \frac{\Lambda}{g(\epsilon_F)}$. We next observe that the density $g(\epsilon_F)$ is roughly the number of electrons in the Debye band, divided by the energy range of the Debye band, i.e., $\frac{2\mathcal{N}}{\omega}$. Then $\kappa^2 = \frac{2\Lambda\omega}{\mathcal{N}}$.

### 3.4.3 The difference between $\epsilon'$ and $\epsilon$

We can estimate the size of the modification to $\epsilon_p$ as follows: $-\kappa^2 \langle c^\dagger_{-p\sigma} c_{-p\sigma} \rangle_{BCS} = -\kappa^2 v_p^2 \approx -\frac{2\Lambda\omega v_p^2}{\mathcal{N}}$. Notice that $|\epsilon_p| \gtrsim \frac{\omega}{\mathcal{N}}$. So comparing the correction term to $\epsilon_p$, we find a ratio $\lesssim \Lambda v_p^2$. A typical experimental value for $\Lambda$ is $\Lambda = .2$. Furthermore $v^2 < 1$. So the correction is small.

### 3.4.4 Some caveats about the perturbation theory that follows

During the work that I've laid out in the many sections below, I came to realize a few important points that I missed along the way. This section is being written after all the analysis which follows, and I don't have the desire or energy to make the corrections which should be made based on the two caveats to be raised here in this section. However, I don't believe that any of my qualitative conclusions depend on these. If we were to attempt an accurate calculation, we'd have to take into account the points about to be discussed, but there are many other matters that would also have to be taken into account and are completely outside the scope of my work.

- First Caveat: Throughout the perturbative analysis, we encounter sums of monomials consisting of products of ladder operators. For example, we will find operators that look like $u_p v_k u_{k-q} v_{p+q} \sum_{pkq} b^\dagger_{p\uparrow} b^\dagger_{k\uparrow} b^\dagger_{(k-q)\downarrow} b^\dagger_{(p+q)\uparrow} b^\dagger_{r\uparrow} b^\dagger_{r'\downarrow}$. Then we will take these operators and operate on states that are defined as ladder operators acting on the ground state. In our example, we might operate on a state of the form $\langle S|$ which is composed as $\langle BCS|b_{\alpha_1\sigma_1}...b_{\alpha_6\sigma_6}$. When the creation operators match the annihilation operators, we are left with a non-zero result – typically $\pm 1$ multiplied by the coherence factor monomial $u_p v_k u_{k-q} v_{p+q}$ for the particular values of $p, q, k$ that match the annihilation operators that define our state. Ultimately, we sum over the $\alpha_i$'s and $\beta_i$'s. However, there is an exception which – for the most part – I missed. There are situations that can occur, when two of the creation operators are precisely the same. So for example, we might have $p + q = k$ and (in our example) since both have the same spin, the creation operators are identical. But for fermion operators, that always results in 0. So after we do all our state-matching and summing, we end up with sums of monomials in the coherence factors for example $u_p v_k u_{k-q} v_{p+q}$ **except** that those sums need to exclude the cases where two operators are equal (in this example, the case when $p + q = k$). I have taken care of that situation in certain cases involving sums encountered during perturbations of the ground state (specifically, when computing energy-perturbations involving monomials of only creation operators acting on the right), the sums are the same as those encountered for the ground state **except** that we have to leave out individual terms that match the quasiparticle energ(ies) of the state being perturbed. Typically, the effect is to reduce the dimensionality of the sum – for example, from a sum like $\sum_{pqk}$ might become $\sum_{pk}$. As we will explain later, single sums typically suppress perturbative results $\mathcal{O}(\mathcal{N})$ so if we have found below a perturbative result with a double sum – which is therefore not suppressed – then if we were to correct that sum by eliminating terms that have two equal operators, we would be left with a single sum which is suppressed by $\mathcal{O}(\mathcal{N})$. The reason we were particularly careful with the case where all potential operators in the monomial are creation operators, is that this is a triple-sum which would normally 'diverge' as $\mathcal{O}(\mathcal{N})$. However, that triple-sum must be subtracted from all energy eigenvalues in order to make the ground state have 0 energy. For states other than the vacuum, we need to introduce one or more creation operators (those that put quasiparticles into the state) and when any of those are equal to a creation operator that appears in the triple sum, we have a 0-result that differs from the situation with the ground state. The total of

all such differences is a double-sum, and since we have subtracted the troublesome triple-sum, we have a nice convergent result (the double sum).

- Second Caveat: This is a lot easier to explain. Our analysis is done entirely based on the idea that the potential term in the Hamiltonian has a coefficient $\kappa^2$ which multiplies all the 4-ladder-monomials. However, that doesn't accurately capture the nature of the 2-body interactions in the theory. Fetter and Walecka are careful to include those interactions by writing the potential term as

$$\mathcal{V} = -\frac{1}{2} \sum_{\substack{k_1+k_2=k_3+k_4 \\ \lambda_1\lambda_2\lambda_3\lambda_4}} \langle \mathbf{k}_1\lambda_1, \mathbf{k}_2\lambda_2 | V | \mathbf{k}_3\lambda_3, \mathbf{k}_4\lambda_4 \rangle \times a^\dagger_{\mathbf{k}_1\lambda_1} a^\dagger_{\mathbf{k}_2\lambda_2} a_{\mathbf{k}_4\lambda_4} a_{\mathbf{k}_3\lambda_3}.$$

  They then do all of their analysis using the matrix elements of $V$, and explaining what approximations are used. By including these terms, we could in principle avoid questions of how to deal with the Debye boundaries, since that information would be part of the model for $V$. In particular, in calculations later where we need to be careful about limits of integration, the inclusion of these $V$ matrix elements would, I think, make it easier to think about what the appropriate boundaries would be for integrals that involve $p + q$, for example. To be a bit more explicit, notice that Fetter and Walecka hypothesize that $V$ is spin-independent and can be approximated by $\langle \mathbf{k}, -\mathbf{k} | V | \mathbf{k}', -\mathbf{k}' \rangle = \kappa^2 \Theta(\omega - |\epsilon'_{|\mathbf{k}|}|)\Theta(\omega - |\epsilon_{|\mathbf{k}'|}|)$. Then, below when we have integrands involving, for example $k$ and $p$, I implement this approximation by putting bounds $-\omega$ and $\omega$ on each of the $k$ and $p$ integrals. However, for integrands involving for example, $p$,$q$ and $p - q$, I have continued to use the same integration bounds but what I should have done (which would follow from inclusion of the $V$ matrix element) is had $\Theta$ functions corresponding to each of $p$, $q$ and $p - q$. This might have helped prove that all coherence factors ultimately are 'cut off' by $\Delta$ (that's far from obvious and perhaps completely untrue, since the $\omega$-cutoff doesn't appear to be directly related to any $\Delta$-energy suppression.) I haven't carried through with all that. For my purposes in the analysis that follows, I've mostly been able to fudge over those nuances. However, I'm concerned that in situations where we really have to carefully do integrals over 6 or even 9 dimensions (3 each for $k$, $q$ and $p$), and then split the integrals into constant energy hyperplanes etc., it might be necessary to pay more attention to the $V$ matrix elements – otherwise risk integrating over regions that should be suppressed. But like I said above, it's far from obvious how that will ultimately help.

### 3.4.5 The residual potential – expansion in terms of $b$ operators

The normal-ordered quartic potential $V_{NO}$ is expressed using the $c$ operators and can therefore be expanded as a sum of quartic terms using the $b$ operators.

Introduce $u, u', v, v'$ in the expansions

$$c^\dagger_{\boldsymbol{k}\uparrow} = ub^\dagger_{k\uparrow} + vb_{-k\downarrow} \quad c^\dagger_{p\uparrow} = u'b_{p\uparrow} - v'b^\dagger_{-p\downarrow}$$

with

$$u \equiv u_k, \quad v \equiv v_k,$$
$$u' \equiv u_p, \quad v' \equiv v_p,$$

The expanded and normal-ordered result is

$$\boxed{V_{NO} = -\kappa^2 \sum_{i=1}^{16} T_i}$$

with the $T_i$ given below.

$$
\begin{aligned}
T_1 &= u^2 u'^2 \; b^\dagger_{k\uparrow} b^\dagger_{-k\downarrow} b_{-p\downarrow} b_{p\uparrow}, \\
T_2 &= -u^2 u' v' \; b^\dagger_{k\uparrow} b^\dagger_{-k\downarrow} b^\dagger_{-p\downarrow} b_{-p\downarrow}, \\
T_3 &= -u^2 u' v' \; b^\dagger_{k\uparrow} b^\dagger_{-k\downarrow} b^\dagger_{p\uparrow} b_{p\uparrow}, \\
T_4 &= -u^2 v'^2 \; b^\dagger_{k\uparrow} b^\dagger_{-k\downarrow} b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow}, \\
T_5 &= -uvu'^2 \; b^\dagger_{k\uparrow} b_{k\uparrow} b_{-p\downarrow} b_{p\uparrow}, \\
T_6 &= -uvu'v' \; b^\dagger_{k\uparrow} b^\dagger_{-p\downarrow} b_{k\uparrow} b_{-p\downarrow}, \\
T_7 &= -uvv'u' \; b^\dagger_{k\uparrow} b^\dagger_{p\uparrow} b_{k\uparrow} b_{p\uparrow}, \\
T_8 &= +uvv'^2 \; b^\dagger_{k\uparrow} b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b_{k\uparrow}, \\
T_9 &= -vuu'^2 \; b^\dagger_{-k\downarrow} b_{-k\downarrow} b_{-p\downarrow} b_{p\uparrow}, \\
T_{10} &= -vuu'v' \; b^\dagger_{-k\downarrow} b^\dagger_{-p\downarrow} b_{-k\downarrow} b_{-p\downarrow}, \\
T_{11} &= +vuv'u' \; b^\dagger_{-k\downarrow} b^\dagger_{p\uparrow} b_{-k\downarrow} b_{p\uparrow}, \\
T_{12} &= -vuv'^2 \; b^\dagger_{-k\downarrow} b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b_{-k\downarrow}, \\
T_{13} &= -v^2 u'^2 \; b_{-k\downarrow} b_{k\uparrow} b_{-p\downarrow} b_{p\uparrow}, \\
T_{14} &= +v^2 u'v' \; b^\dagger_{-p\downarrow} b_{-k\downarrow} b_{k\uparrow} b_{-p\downarrow}, \\
T_{15} &= +v^2 v'u' \; b^\dagger_{p\uparrow} b_{-k\downarrow} b_{k\uparrow} b_{p\uparrow}, \\
T_{16} &= +v^2 v'^2 \; b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b_{-k\downarrow} b_{k\uparrow}.
\end{aligned}
$$

In the remainder of this section on the residual potential, these $T_i$ will be used. We've been a bit cavalier in our definitions of $T_i$ because they are, in fact, functions of $p$ and $k$. We could show that dependence as $T_i(p,k)$ so that we would write $V_{NO} = -\kappa^2 \sum_{pk} \sum_{i=1}^{16} T_i(p,k)$. In what follows, we usually suppress the $p, k$ arguments of $T_i$.

### 3.4.6 The residual potential – first order changes to eigenvalues

We'll start with the first-order correction to energy eigenvalues. Each state $|n^{(0)}\rangle$ is of the form $\prod_{(k_i)} b^\dagger_{k_i}|0\rangle_{BCS}$ or ${}_{BCS}\langle 0|\prod_{(k_i)} b_{k_i}$ for some collection $(k_i)$. The annihilation operators (in terms of the $b$ ladder operators) of $V_{NO}$ operate on the right and the creation operators act on the left.

First consider an excited state consisting of exactly one quasiparticle, e.g. $b^\dagger_{r\uparrow}|BCS\rangle$. The first-order energy correction is $\langle BCS|b_{r\uparrow}V_{NO}b^\dagger_{r\uparrow}|BCS\rangle$. The only possible non-zero contributions would require a monomial in $V_{NO}$ which includes $b^\dagger_{r\uparrow}$ and $b_{r\uparrow}$. These would be required in order to pair off with the excited quasiparticle for both the 'bra' and 'ket'. There remain two unpaired operators in that monomial (since the monomial is a quartic) and because they are normal-ordered, their $|BCS\rangle$ expectation value is 0.

What about states with two quasiparticles? First consider a state of the form $b^\dagger_{r\uparrow}b^\dagger_{r'\uparrow}|BCS\rangle$, where both quasiparticles have the same spin. The only contributing monomials of $V_{NO}$ would be ones where all 4 operators have that spin. Moreover, there would need to be two annihilation operators and two creation operators. The only candidate is term $T_7 = -uvv'u'\ b^\dagger_{k\uparrow}\, b^\dagger_{p\uparrow}\, b_{k\uparrow}\, b_{p\uparrow}$. By matching annihilation and creation operators, we see that $p$ matches $r$ or $r'$, and $k$ matches $r'$ or $r$. Both combinations correspond to a unique solution for $p, k$ in terms of $r, r'$, and also $uvu'v' = u_r v_r u_{r'} v_{r'}$. Therefore the double-sum collapses to 2 individual summands, and if we arbitrarily set the $uvu'v'$-coherence factor to be $\approx 1$ (it is less than 1) the expectation value for this state, is $-2\kappa^2 \approx -\frac{4\Lambda\omega}{\mathcal{N}}$. Moreover, to obtain the relative correction, we must divide by the energy of the unperturbed quasi-particle state. Since there are two quasiparticles, that energy $> 2\Delta$. Taking the ratio, we find that **the relative correction is** $\approx \frac{2\Lambda\omega}{\Delta\mathcal{N}} = \mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$**. This is negligibly small.** For opposite spin the only candidate is $T_{10}$ which, in the same way, leads to a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$.

Now consider the case of states $|n^{(0)}\rangle$ in which there are two quasiparticles of opposite spin, $b^\dagger_{r\uparrow}b^\dagger_{r'\downarrow}|BCS\rangle$. In the double-sum, we would require annihilation operators to pair with $b^\dagger_{r\uparrow}$ and $b^\dagger_{r'\downarrow}$ as well as creation operators to pair with $b_{r\uparrow}$ and $b_{r'\downarrow}$. This can happen in $V_{NO}$ with the terms $T_1, T_6, T_{11}, T_{16}$. Start by looking at $T_1 = u^2u'^2\ b^\dagger_{k\uparrow}\, b^\dagger_{-k\downarrow}\, b_{-p\downarrow}\, b_{p\uparrow}$. By setting the pair-matches we have $p = r$ and $k = r$ but this can only happen if $r' = -r$. If in fact, $r' = -r$, the double sum collapses to a single term and as before, we obtain a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$. Next look at $T_6 = -uvu'v'\ b^\dagger_{k\uparrow}\, b^\dagger_{-p\downarrow}\, b_{k\uparrow}\, b_{-p\downarrow}$. By

setting pair-matches we have $p = -r'$ and $k = r$. Again, a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$. Similar results hold for $T_{11}$ and $T_{16}$.

Notice that the above technique involves solving 2 equations in 2 unknowns. If the equations are independent, then there is exactly one 'match' and the relative correction is $\mathcal{O}(\frac{1}{\mathcal{N}})$. If the 2 equations are dependent (i.e., they are proportional to each other) , then they can only be solved for certain $r = -r'$. We are now ready to tackle states with more than two quasiparticles. There are several different kinds of situations. Since my objective is to estimate the size of perturbations, I'll only consider a representative example and will generalize from there.

Let $|n^{(0)}\rangle = b^\dagger_{r_1\uparrow} b^\dagger_{r_2\downarrow} b^\dagger_{r_3\uparrow} b^\dagger_{r_4\uparrow} |BCS\rangle$. All momenta are different from one another, otherwise by the Pauli exclusion principle the result would be 0. Take any $T_i$. If we pair off an unequal number of creation and annihilation operators, our expectation value will be left with an unequal number of annihilation and creation operators (since it began with an equal number prior to matching – creation operators for the 'ket' and their adjoints for the 'bra'). But in that case, the expectation value is 0. If we have an equal number of annihilation and creation operators in $T_i$, then we can choose to match the annihilation operators with any pair (provided spins match) out of $r_1, r_2, r_3, r_4$ and we can match the creation operators with any pair – possibly different. However, if the two matching pairs are different, than the remaining unmatched operators have different momenta, can anticommute and then annihilate the 'ket' (i.e. $\langle BCS|\alpha\beta^\dagger|BCS\rangle = -\langle BCS|\beta^\dagger\alpha|BCS\rangle = 0$ if $\alpha$ and $\beta$ are annihilation operators ). As a result, the $T_i$ operators pair off with at most 2 of the quasiparticles in $|n^{(0)}\rangle$, and the remaining operators pair off with each other. As an example, take $T_1$ whose annihilation operators have opposite spin. For $|n^{(0)}\rangle$ above, the possible pairs are $(r_1, r_2), (r_3, r_2)$ and $(r_4, r_2)$. If we took $T_7$ where both annihilation operators have spin up, the possible pairs are $(r_1, r_3), (r_1, r_4)$ and $(r_3, r_4)$. And if we took $T_{11}$ with all spins down, there are no candidate pairs (so a contribution of 0). Going back to $T_1$ and $T_7$ there are 3 possible contributions (a contribution involves the 2 possibilities per pairing) – one for each pairing – leading to a correction of magnitude $6\kappa^2$ (we set the $u$ and $v$ values to 1, which is an upper bound). The relative corrections for those terms are $\approx 6\frac{\Lambda}{g(\epsilon_F)}E_n = \frac{12\Lambda\omega}{\mathcal{N}E_n} = \mathcal{O}(\frac{1}{\mathcal{N}})$. We see that this approximate contribution comes from some, but not all, of the $T_i$.

Now we can generalize. Let $|n^{(0)}\rangle$ be a state with $M$ quasiparticles of various spins. The maximum amount of pairing if all spins were up, would occur

with $T_7$, in which case there would be a total of $2\binom{M}{2} = M(M-1)$ and a relative correction (magnitude) of $\approx \kappa^2 M^2$. With mixed spins, the maximum amount of pairing (per $T_i$) would be less. For example, assume half the spins are different. First look at the case when all momentum-sectors are different ($b_{p\uparrow}$ and $b_{-p\downarrow}$ are considered to be in the same momentum-sector). The contributing terms of the potential are $T_6, T_7, T_{10}, T_{11}$. In each case there are approximately $\frac{M^2}{4}$ pairings . So the total correction is $\kappa^2 M^2$. Second, look at the case when all particles appear in Cooper-pairs. That is, if the state $|n^{(0)}\rangle$ contains a particle labeled $(p, \uparrow)$, then it also contains a particle labeled $(-p, \downarrow)$. The term $T_1$ has $\frac{M}{2}$ matches, since the creation operators must also match the annihilation operators otherwise there would be a mismatch between the particles not annihilated in the 'bra' and 'ket'. Similarly with the term $T_{16}$. So both those terms are suppressed by $\mathcal{O}(\mathcal{N})$ relative to terms $T_6, T_7, T_{10}, T_{11}$. Those contribute $\approx \frac{M(M-1)}{4}$ pairings each so in total, the correction is $\approx \kappa^2 M^2$ for large $M$.

So far, we haven't considered the $u$ and $v$ contributions. When $\epsilon' \gg \Delta$, $u_{\epsilon'} \approx 1$ and $v_{\epsilon'} \approx 0$ and when $\epsilon' \ll -\Delta$, $v_{\epsilon'} \approx 1$ and $u_{\epsilon'} \approx 0$. For large values of $M$, the majority of quasiparticles have $|\epsilon'| \gg \Delta$ so either $u_{\epsilon'}$ or $v_{\epsilon'}$ are $\approx 0$. All $T_i$ of interest, except for $T_1$ and $T_{16}$ have at least one $u$ factor and one $v$ factor, so those contributions are suppressed for $|\epsilon'| \gg \Delta$. Both $T_1$ and $T_{16}$ are already suppressed by $\mathcal{O}(\mathcal{N})$ so those terms can be ignored. For the other $T_i$, the $u, v$ suppression (except for $|\epsilon'| < \Delta$) reduces the multiplicative factor by $\frac{\Delta}{\epsilon'}$.

We also need to estimate $E_n$. This requires a bit more analysis. What interests us here, are states where $M$ is a significant fraction of $\mathcal{N}$. So let's say for example, that $M = \frac{1}{2}\mathcal{N}$. Since two quasiparticles (they are fermions) can't occupy the same state, the most energetic[9] quasiparticle must have $|\epsilon'| \geq \frac{\omega}{2}$. This follows from the fact that near the Fermi energy $\epsilon_F$, the density (which takes into account momentum-degeneracy) is approximately constant. So if $M = \frac{\mathcal{N}}{2}$, that means half of the states in the Debye window are occupied, so the most energetic quasiparticle has to have $|\epsilon'|$ more than half the Debye energy (relative to $\epsilon_F$). For simplicity, assume that all quasiparticles have $|\epsilon'|$ less than $\omega/2$. Then we can use the density of states $g(\epsilon_F)$ to estimate the number of quasiparticles at any $|\epsilon'|$ below $\omega/2$. Each quasiparticle has energy $E_{\epsilon'} = \sqrt{\epsilon'^2 + \Delta^2}$. Recall that $\Delta \ll \omega$ so most of the $|\epsilon'|$ values are much greater than $\Delta$. For those, we can estimate $E_{\epsilon'} \approx |\epsilon'|$. For $|\epsilon'|$ values

[9]I'm misusing the word energetic or, for that matter the term $\epsilon'$ in this paragraph. Since $\epsilon' \approx \epsilon - \epsilon_F$, the value of $\epsilon'$ can be positive or negative. When I speak of $\epsilon'$ I really mean $|\epsilon'|$ and when I say energetic, I really mean the value of $|\epsilon'|$

$\ll \Delta$, we can estimate $E_{\epsilon'} \approx \Delta$. We can crudely estimate the total energy of $|n^{(0)}\rangle$ – with all these assumptions – as $E_n \approx g(\epsilon_F)\left(2\Delta^2 + 2\int_{\Delta}^{\frac{\omega}{2}} \epsilon d\epsilon\right) = \frac{\mathcal{N}}{2\omega}\left(\Delta^2 + \frac{\omega^2}{4}\right) \approx \frac{\mathcal{N}\omega}{8}$.

Now, in the mixed-spin Cooper-pair case, if we ignore the suppression factors coming from the $u, v$ terms, the relative correction – namely, the ratio of the perturbative correction to the energy of $|n^{(0)}\rangle$ – is $\frac{8\kappa^2 M^2}{\mathcal{N}\omega} = \frac{8\Lambda\omega\mathcal{M}^2}{\mathcal{N}^2\omega} = \frac{8\Lambda\mathcal{M}^2}{\mathcal{N}^2}$. When we include the $u, v$ suppression factors, we obtain $\frac{8\Lambda M^2}{\mathcal{N}^2}\frac{\Delta}{[|\epsilon'|]}$, where $[|\epsilon'|]$ denotes the average (over the $M$ quasiparticles) of $|\epsilon'|$. If we take as an example, $M = \frac{\mathcal{N}}{2}$ and $[|\epsilon'|] = \frac{\omega}{2}$, then the total relative correction would be $\approx 4\Lambda\frac{\Delta}{\omega}$. In physical systems we can pick (as an example) $\Lambda \approx 0.2$ and $\frac{\Delta}{\omega} \approx .02$. So we see for the mixed-spin Cooper state, the first order relative correction is of order of 2%. Other states that we have examined would have even smaller corrections. Could we look at situations with larger values of $M$? Maybe so, but we would be probing regions of energy-space where Cooper-pairs become unstable, so that the BCS approximation might be a poor starting point.

### 3.4.7 The residual potential – second order changes to eigenvalues

We now examine the second-order shifts:

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left|\langle m^{(0)}|V_{NO}|n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}} \tag{15}$$

We will avail ourselves of the following theorem.

> Define $b^0_{r\sigma} \equiv b_{r\sigma}$ and $b^1_{r\sigma} \equiv b^\dagger_{r\sigma}$. Then define $\mathcal{P} \equiv \langle BCS|b^{i_1}_{r_1\sigma_1} b^{i_2}_{r_2\sigma_2} ... b^{i_k}_{r_k\sigma_k} ... b^{i_n}_{r_n\sigma_n}|BCS\rangle$, where $i = 1$ or $i = 0$. If $n$ is odd, then $\mathcal{P} = 0$. If $n$ is even, then exactly $\frac{n}{2}$ of the indices $k$ have the value 0 and other half have the value 1. More precisely, half of the ladder operators are annihilation operators and the other half are their adjoints (creation operators). Otherwise, $\mathcal{P} = 0$.

To prove that theorem, it is easiest to use induction. Assume the theorem is true if there are $M$ annihilation operators and $N$ creation operators, where

$N > M$. Then anticommute each annihilation operator all the way to the right, at which point it annihilates $|BCS\rangle$. Either the annihilation operator, on its journey to the right, doesn't encounter it's adjoint (i.e., the corresponding creation operator) in which case $\mathcal{P} = 0$ or it encounters its adjoint at least once. Then anticommute so that $a_{s,\lambda}a^\dagger_{s,\lambda} = -a^\dagger_{s,\lambda}a_{s,\lambda} + 1$. The first term on the right has the annihilation operator to the right of the creation operator so as before, it continues its journey to the right. The second term, which is a $c$-number (namely, the value 1) has eliminated one annihilation operator and one creation operator which is its adjoint, so that the overall product now has $N-1$ creation operators and $M-1$ annihilation operators. By induction this gives $\mathcal{P} = 0$. So the theorem is proven by induction ***provided*** that it can be proven for the case $N = M + 1$. We can prove that in a similar way with induction on $M$. The case $M = 0$ has a single creation operator, and that annihilates $\langle BCS|$ so again $\mathcal{P} = 0$. Finally, we need to consider the possibility that $M > N$. The proof is identical except that creation operators are moved to the left. Finally it remains to show that when there an equal number of creation and annihilation operators, the expectation value is 0 unless every creation operator corresponds to an annihilation operator which is its adjoint. A similar induction argument can be used. QED.

Now examine some examples.

- We'll first look at the ground state, $|n^{(0)}\rangle = |BCS\rangle$. It turns out that the $T_4$ and $T_{13}$ terms make contributions to the second-order perturbation. In fact, $E^{(2)}_{gs} = -\sum_{mkpk'p'} \frac{\langle BCS|T_{13}(k',p')|m\rangle\langle m|T_4(k,p)|BCS\rangle}{E^{(0)}_m}$. For $\langle m|$ to match $T_4$, we must have $\langle m| = \langle BCS|b_{-k\downarrow}\, b_{k\uparrow}\, b_{-p\downarrow}\, b_{p\uparrow}$. Then $T_{13}$ matches $|m\rangle$ and that fixes $k'$ and $p'$. By including the coherence factors of $T_{13}$ and $T_4$, we end up with $E^{(2)}_{gs} = -\kappa^4 \sum_{kp} \frac{(u_p v_p)^2 (u_k v_k)^2}{2(E_k+E_p)} = -\kappa^4\Delta^4 \sum_{kp} \frac{1}{2(E_k+E_p)E_p^2E_k^2}$. We can estimate this with an integrated form of the sum, and rewriting $\kappa^2 = \frac{\Lambda}{g(\epsilon_F)}$, as $E^{(2)}_{gs} \approx -\Lambda^2\Delta^4 \int_0^\omega d\epsilon'_k d\epsilon'_p \frac{1}{2(E_k+E_p)E_p^2E_k^2}$. Each of the integrals (over $\epsilon'_k$ and $\epsilon'_p$) can be roughly split into 2 parts: $\int_0^\Delta$ and $\int_\Delta^\omega$ so altogether our double integral has 4 parts (both less than $\Delta$, both larger than $\Delta$ etc.). For $\epsilon'_k < \Delta$, we have $E_k \approx \Delta$. For $\epsilon'_k > \Delta$ we have $E_k \approx \epsilon'_k$. We could then come up with an approximate value for the double-integral. I've done that using MAXIMA and obtain a result $E^{(2)}_{gs} \approx -0.9\Lambda^2\Delta$.

  Is that a small perturbation? The question is more or less meaningless since we're comparing a nonzero value to zero. In point of fact, we need

to examine more carefully the role of the ground state. By definition, the ground state should have 0 energy. So if the residual potential shifts the ground state energy from zero, then we must redefine the energy (call it 'renormalization' if you want a fancy term – but in classical physics we would just refer to this step as "adding an arbitrary constant to the energy"). What matters are energy-differences. For example, we might care what the value of the mass gap is. As we go forward and explore second-order perturbations of the other quasiparticle energy eigenvalues, we will need to compare those to the ground state energy. In practice, that will mean we need to "renormalize" those second-order contributions by adding $\approx 0.9\Lambda^2\Delta$.

- Suppose $|n^{(0)}\rangle$ is a state with one quasiparticle, e.g. $|n^{(0)}\rangle = b^\dagger_{r\uparrow}|BCS\rangle$. Consider the term $T_1 = -\kappa^2 u_p^2 u_k^2 b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b_{-k\downarrow} b_{k\uparrow}$. This has two creation operators and two annihilation operators. The state $|n^{(0)}\rangle$ involves one creation operator. So any state $\langle m^{(0)}|$ which contributes to the sum in Eq. (15) must involve exactly one annihilation operator so that the inner product has 3 creation and 3 annihilation operators. The creation operator $b^\dagger_{r\uparrow}$ (applied to the BCS ground state) for $|n^{(0)}\rangle$ must be paired with an annihilation operator from $T_1$ (it can't be paired with the annihilation operator from $\langle m^{(0)}|$ because in Eq. (15), $\langle m^{(0)}| \neq \langle n^{(0)}|$. The only candidate with spin up, is $b_{k\uparrow}$. Therefore we must set $k = r$. So this transforms $|n^{(0)}\rangle$ to $|BCS\rangle$. Now we have a remaining annihilation operator on the right of $T_1$ and this annihilates $|BCS\rangle$, resulting in a 0 contribution.

  Continuing with this example where $|n^{(0)}\rangle = b^\dagger_{r\uparrow}|BCS\rangle$, consider the term $T_2 \propto b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b^\dagger_{-k\downarrow} b_{-k\downarrow}$. We see that this term has 3 creation operators and 1 annihilation operator. The annihilation operator has to pair with the creation operator $b^\dagger_{r\uparrow}$ (applied to the BCS ground state) for $|n^{(0)}\rangle$ but that's not possible because the $T_2$ annihilation operator is spin down. So $T_2$ has no contribution. On the other hand, the term $T_3$ doesn't have this problem because the one annihilation operator has spin up. So let's look at $T_3 \propto b^\dagger_{p\uparrow} b^\dagger_{-p\downarrow} b^\dagger_{k\uparrow} b_{k\uparrow}$. By pairing the annihilation operator, we have (as in the above paragraph) $k = r$. We now have 3 remaining unpaired creation operators in $T_3$ so they must pair with 3 annihilation operators for $\langle m^{(0)}|$. Since one of the creation operators is $b^\dagger_{k\uparrow}$, then one of the annihilation operators for $\langle m^{(0)}|$ must be $b_{k\uparrow}$ (acting on $\langle BCS|$). What about the remaining creation operators for $\langle m^{(0)}|$. These aren't yet determined. However, they have to match the remaining creation operators of $T_3$, so all in all, we have

$\langle m^{(0)}| = \langle BCS|b_{p\uparrow}b_{-p\downarrow}b_{r\uparrow}$. For every value – with the exception of $p = r$, which leads to 0 because $b_{p\uparrow}b_{p\uparrow} = 0$ – of the summand $p$ in Eq. (15), there is a state $\langle m^{(0)}|$ so that $\langle m^{(0)}|T_3|n^{(0)}\rangle \neq 0$. We're not quite done. Remember that in Eq. (15), the term $\langle m^{(0}|T_3|n^{(0)}\rangle \neq 0$ is multiplied by $\langle n^{(0)}|V_{NO}|m^{(0)}\rangle$. Expanding the states, we see we need a monomial $T$ for which $\langle BCS|b_{r\uparrow}Tb^\dagger_{p\uparrow}b^\dagger_{-p\downarrow}b^\dagger_{r\uparrow}|BCS\rangle \neq 0$. This requires that $T$ is some permutation of $b^\dagger_{r\uparrow}b_{p\uparrow}b_{-p\downarrow}b_{r\uparrow}$. If we write $V_{NO}$ which appears in $\langle n^{(0)}|V_{NO}|m^{(0)}\rangle$, as a sum over $p', k'$, we find that the values $p'$ and $k'$ are completely determined by $p$ and $r$, so the sum collapses to a finite number of terms (often just one term). After some amount of inspection, it can be seen that both $T_5$ and $T_{15}$ give non-zero contributions (note that the defining sums over $p', k'$ collapse to a single term). In summary, we've shown a non-zero contribution from the combination $\frac{\langle n^{(0)}|T_5+T_{15}|m^{(0)}\rangle\langle m^{(0)}|T_3|n^{(0)}\rangle}{E^{(0)}_n - E^{(0)}_m}$, which appears as a summand in Eq. (15).

Let us estimate the contribution of this term.

- The denominator $\mathcal{D} \equiv |E^{(0)}_n - E^{(0)}_m|$ is $\mathcal{D} = |\sqrt{\Delta^2+\epsilon_r^2} - \left(\sqrt{\Delta^2+\epsilon_r^2} + 2\sqrt{\Delta^2+\epsilon_p^2}\right)| = 2\sqrt{\Delta^2+\epsilon_p^2} > 2\Delta$.
- We have a sum over $p$. This sum has $\mathcal{O}(\mathcal{N})$ terms. In fact, since $T_5$ and $T_{15}$ both involve the coefficient $v_p$ and $T_3$ involves the coefficient $u_p$, the sum over $p$ effectively cuts off at $|\epsilon'_p| \approx \Delta$ , and the number of terms is $\approx \mathcal{N}\frac{\Delta}{\omega}$.
- Furthermore, since each $T_i$ has a coefficient $\kappa^2$, then each numerator (which involves products of two $T_i$'s) has a coefficient $\kappa^4 \approx \frac{4\Lambda^2\omega^2}{\mathcal{N}^2}$.
- Estimate the remaining $u$ and $v$ coefficients as approximately 1.
- Then multiplying everything together we obtain $\frac{\langle n^{(0)}|T_5+T_{15}|m^{(0)}\rangle\langle m^{(0)}|T_3|n^{(0)}\rangle}{E^{(0)}_n - E^{(0)}_m} = \mathcal{O}\left(\frac{\Lambda^2\omega}{\mathcal{N}}\right) \approx \mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$.

**If we then take the ratio of this correction to the original energy of $|n^{(0)}\rangle$, we get $\mathcal{O}\left(\Lambda^2\frac{\omega}{\Delta}\frac{1}{\mathcal{N}}\right) \approx \mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$.** As before, the denominator $\mathcal{N}$ is huge so the relative correction is approximately 0.

So far, we've only looked at $\langle m^{(0)}|T_i|n^{(0)}\rangle$ where $i = 1, 2, 3$. However, our methods apply to the remaining $T_i$'s **except** for $T_4$ which we'll deal with separately in a moment. For the other $T_i$'s, the following holds:

- One of the annihilation operators in $T_i$ must pair with $b^\dagger_{r\uparrow}$ (remember that $|n^{(0)})\rangle = b^\dagger_{r\uparrow}|BCS\rangle$. Therefore, either $k$ or $p$ must be set

to $r$. That leaves only a single sum rather than a double sum. Without loss of generality, assume we're left with a sum over $k$.

- The state $\langle m^{(0)}|$ must then contain a quasiparticle in the $k$-sector. Therefore in $\langle n^{(0)}| \sum_j \sum_{p'k'} T_j(p',k')|m^{(0)}\rangle$, we either have $(p'=r,k'=k)$ or $(p'=k,k'=r)$. There is no remaining double sum over $p',k'$.
- Altogether, the last two items show that we have (at most) a sum over $\mathcal{N}$ terms.
- When we compute $E_n^{(0)} - E_m^{(0)}$, we'll obtain a value $> 2\Delta$ for precisely the reason that we did above (recall that one of the quasiparticles in $|m^0)\rangle$ must be in the $p$ sector so its energy precisely cancels the energy of $|n^0\rangle$.
- There is an overall factor of $\kappa^4$.
- Although there might also be a suppression from $u$ and $v$ coefficients, I haven't bothered analyzing that because it makes no difference to the order of magnitude of the relative correction.
- Similar to above (but without the $uv$ suppression), **the relative correction is $\mathcal{O}\left(\Lambda^2 \frac{\omega^2}{\Delta^2}\frac{1}{\mathcal{N}}\right) \approx \mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$.** Again negligible.

What about $T_4$? This we deal with more or less the way we dealt with the ground-state. None of the operators are annihilation operators, so none of them match up to $b_r^\dagger$. Instead, we end up with 5 creation operators including $b_r^\dagger$. Then $\langle m^0| = \langle BCS|b_{-k\downarrow}^\dagger\, b_{-p\downarrow}^\dagger\, b_{-k\downarrow}\, b_{-p\downarrow} b_{r\uparrow}$. Just as for the ground state, we then need to match $\langle n^{(0)}|T_{10}(p'k')|m^{(0)}\rangle$ with $\langle m^{(0)}|$ and the result is just as we had for the ground state **except** that the double sum needs to exclude $k=r$ and $p=r$, since in both cases we would have the square of a creation operator which would give 0. Once we calculate $E_R^{(2)}$ by adding $\kappa^4\Delta^4 \sum_{kp} \frac{1}{2(E_k+E_p)E_p^2E_k^2} \approx 0.9\Lambda^2\Delta$, we end up with a contribution from the 5-quasiparticle intermediate states of $2\kappa^4\Delta^4 \sum_p \frac{1}{2(E_r+E_p)E_p^2E_r^2}$. We can evaluate this with the usual integration approximation, but now we've had some experience with this kind of thing, we need only note the following. The conversion from sum to (single) integral leads to a density factor of $g(\epsilon_F)$. This is multiplied by one of the factors of $\kappa^2$ to give $\Lambda$. The integral is bounded by $\omega$ so the result of integration is ultimately bounded by $\omega$ (in point of fact it will turn out to be $\mathcal{O}(\Delta)$)). But more importantly, there is a factor of $\kappa^2$ left over which is $\mathcal{O}(\frac{1}{\mathcal{N}})$. Thus the contribution ends up being $\mathcal{O}(\frac{1}{\mathcal{N}})$ which as usual is negligible.

- In the above example where we were perturbing a single quasi-particle state, we saw that the double sum was reduced to a single sum, leading to a factor of $\mathcal{N}$. This canceled $\frac{1}{\mathcal{N}}$ coming from one of the $\kappa^2$ coefficients. However, we are left with a remaining $\kappa^2$ coefficient which is proportional to $\frac{1}{\mathcal{N}}$. That ultimately led to an overall relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$. Is this suppression by $\mathcal{O}(\mathcal{N})$ a feature of the energy-correction to all states $|n^{(0)}\rangle$ and not just the single-quasiparticle states? In the next example I'll look at a state with a correction of $\mathcal{O}\left(\frac{\log \mathcal{N}}{\mathcal{N}}\right)$. This is still negligible, but not as negligible as $\mathcal{O}(\frac{1}{\mathcal{N}})$.

  We'll take $|n^{(0)}\rangle = b^\dagger_{r\uparrow} b^\dagger_{-r\downarrow}|BCS\rangle$. This is a state with two opposite spin quasiparticles and zero total momentum. Following the same kind of logic used in the first example, we first look for $T_i$ with two annihilation operators to pair off with the creation operators $b^\dagger_{r\uparrow}, b^\dagger_{-r\downarrow}$. Moreover, since we'll also be evaluating $\langle n^{(0)}|T_i = \langle BCS|b_{-r\downarrow} b_{r\uparrow} T_i$, then $T_i$ also needs to have two creation operators to pair off with the annihilation operators $b_{-r\downarrow}, b_{r\uparrow}$. These conditions strongly constrain the options for $T_i$.

  In fact, the only two viable options are $T_1$ and $T_{16}$. It might seem as though $T_6$ and $T_{11}$ would also be options but to see that this isn't the case, consider $T_6 \propto b^\dagger_{p\uparrow} b^\dagger_{-k\downarrow} b_{p\uparrow} b_{-k\downarrow}$. In order for the two annihilation operators to pair with the creation operators associated with $|n^{(0)}\rangle$, both $p$ and $k$ would have to be equal $r$ (or $-r$). But then, for $\langle m^{(0)}|$ to 'match' $T_6$, it would also have to be a state with the same quasiparticles as $|n^{(0)}\rangle$. However, such a state is excluded from the sum in Eq. (15).

  So let's focus on $T_1$ and $T_{16}$. It's straightforward to see that the corresponding terms for $\langle n^{(0)}|T_i|m^{(0)}\rangle$ are also $T_1$ and $T_{16}$. Moreover, for all 4 combinations, the denominator is $E^{(0)}_n - E^{(0)}_m$. All the inner products are equal, so we end up with

$$E^{(2)}_n = \sum_{k\neq r} \frac{4\kappa^4}{\left(E^{(0)}_r - E^{(0)}_k\right)} \left(u_r^2 u_k^2 + v_r^2 v_k^2\right)^2 \tag{16}$$

  We've simplified notation so that energies $E^{(0)}_k$ are the energies for two quasiparticles, each with momentum $k$ or $-k$ and with opposite spin. Only the magnitude of $k$ matters, so we have $E^{(0)}_k = 2\sqrt{\epsilon'^2_k + \Delta^2}$.

  To calculate this, we have from Appendix A, that on average, adjacent 'bare' energy levels $\epsilon'_a$ and $\epsilon'_b$ are separated by an energy $\delta' = \delta \frac{\sqrt{2\epsilon_F}}{\sqrt{m^*}} =$

$\frac{\left(128\pi^2\epsilon_F m^{*3}\omega^2\right)^{\frac{1}{6}}}{\mathcal{N}^{\frac{1}{3}}}\frac{\sqrt{2\epsilon_F}}{\sqrt{m^*}}$ .[10] Next we approximate our sum over intermediate states $\epsilon'_k$ by a pair of integrals[11] (we need two integrals so that we can cut out the case where $\epsilon'_r = \epsilon'_k$)

$$\sum_{\epsilon'_k \neq \epsilon'_r; |\epsilon'_k|, |\epsilon'_r| < \omega} \rightarrow g(\epsilon_F) \int_{-\omega}^{\epsilon'_r - \delta'} d\epsilon' + g(\epsilon_F) \int_{\epsilon'_r + \delta'}^{\omega} d\epsilon'.$$

We can proceed to estimate the sum in Eq. (16). As a simplification for estimating orders of magnitude, set $u_r^2 u_k^2 + v_r^2 v_k^2 \approx \frac{1}{2}$. Recall $E_i^{(0)} =$

[10] We have to be careful, because many states of differing momenta can share each energy level. There are several ways to deal with this degeneracy. I haven't delved into how to reconcile my approach below with the formal treatment of second order perturbation theory as done when there are degenerate energy levels. I believe my approach is correct but that should be verified. Formally one needs to find certain subspaces spanned by $H_0$ eigenvectors for both $\epsilon'_a$ and $\epsilon'_b$ and then do perturbation theory within each subspace. I think another approach has to do with the volume of degeneracy compared to the volume of the full space. Think of concentric 2D spherical shells in momentum space where each shell has the same energy (sum of squares of the momenta). Then if we approximate our sums by integrals, we could go to spherical coordinates and the energy $\epsilon'(\mathbf{p})$ is $\epsilon'(|\mathbf{p}|)$. . We avoid degeneracy by restricting $\epsilon'_a$ and $\epsilon'_b$ to be points at the same angle within the sphere – then integrating over the angles. This leads to an integration parameter of $4\pi p^2 dp$. Provided that the momenta all lie within the Debye band, the value of $4\pi p^2 dp$ is approximately constant within the Debye band, and can be absorbed into $g(\epsilon)$. In summary, when we say $\epsilon'_a \neq \epsilon'_b$, we assume that all degeneracies have been removed so that there is only one example of each $\epsilon'$.

[11] This technique – when $\delta' \rightarrow 0$ – is known as 'the Cauchy principal value' and denoted $\mathcal{P}$.

$2\sqrt{\epsilon_i'^2+\Delta^2}$ so our sum becomes

$$\begin{aligned}
E_n^{(2)} &\approx \sum_{k\neq r} \frac{\kappa^4}{2\left(\sqrt{\epsilon_r'^2+\Delta^2}-\sqrt{\epsilon_k'^2+\Delta^2}\right)} \\
&\approx g(\epsilon_F)\int_{-\omega}^{\epsilon_r'-\delta'} d\epsilon' \frac{\kappa^4}{2\left(\sqrt{\epsilon_r'^2+\Delta^2}-\sqrt{\epsilon'^2+\Delta^2}\right)} + \\
&\qquad g(\epsilon_F)\int_{\epsilon_r'+\delta'}^{\omega} d\epsilon' \frac{\kappa^4}{2\left(\sqrt{\epsilon_r'^2+\Delta^2}-\sqrt{\epsilon'^2+\Delta^2}\right)} \\
&\approx g(\epsilon_F)\frac{\kappa^4}{2}\Bigg(-\frac{\sqrt{\epsilon_r'^2+\Delta^2}}{\epsilon_r'}\log\frac{(\omega-\epsilon_r')\left(\sqrt{(\epsilon_r'-\delta')^2+\Delta^2}-\epsilon_r'\right)}{(\omega+\epsilon_r')\left(\sqrt{(\epsilon_r'+\delta')^2+\Delta^2}-\epsilon_r'\right)} + \\
&\qquad \log\frac{\left(\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'-\delta)^2+\Delta^2}+\epsilon_r'-\delta'\right)}{\left(-\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'+\delta')^2+\Delta^2}+\epsilon_r'+\delta'\right)}\Bigg)
\end{aligned} \tag{17}$$

Let's explore the shift for four cases:

1. Start with a state of energy $\epsilon_r' = \frac{\Delta}{2}$. Note that $\Delta \ll \omega$ and $\delta' \ll \Delta$ so Taylor expand accordingly. Then

$$\begin{aligned}
E_n^{(2)} &\approx g(\epsilon_F)\frac{\kappa^4}{2}\left(\sqrt{5}\frac{\Delta}{\omega}+2\log(2\frac{\omega}{\Delta})\right) \approx \frac{\Lambda^2\omega}{\mathcal{N}}\left(\sqrt{5}\frac{\Delta}{\omega}+2\log(2\frac{\omega}{\Delta})\right) \\
&= \mathcal{O}(\frac{1}{\mathcal{N}})
\end{aligned}$$

2. Next consider a state with $\epsilon_r' = \frac{\omega}{4}$.

$$\begin{aligned}
E_n^{(2)} &\approx g(\epsilon_F)\frac{\kappa^4}{2}\left(\log\frac{5}{3}\left(\frac{4+\sqrt{17}}{-4+\sqrt{17}}\right)\right) \\
&\approx 4.7\frac{\Lambda^2\omega}{\mathcal{N}} \\
&= \mathcal{O}(\frac{1}{\mathcal{N}}).
\end{aligned}$$

3. Now consider the lowest-energy excitation with $\epsilon_r' = \delta'$. Then, noting that $\Delta \gg \delta'$ and doing Taylor expansions where appropriate,

$$E_n^{(2)} \approx g(\epsilon_F)\kappa^4\left(\frac{\Delta}{\omega}+\log\frac{2\omega}{\Delta}\right).$$

Again a correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$.

4. Finally, notice that the largest (smallest) possible value of $\epsilon'_r$ is $\epsilon'_r = \pm(\omega - \delta')$, since the denominator of the integrand is forbidden (by definition) from being 0. Let us consider the state with $\epsilon'_r = (\omega - \delta')$. $E_n^{(2)}$ has a term proportional to $\log(\delta') \propto \log \mathcal{N}$ so multiplying by the factor of $\frac{1}{\mathcal{N}}$ coming from the factor $g(\epsilon_F)\kappa^4$ we obtain $E_n^{(2)} = \mathcal{O}\left(\frac{\log \mathcal{N}}{\mathcal{N}}\right)$. This is still an extremely small correction. We won't examine this situation in any further detail because the Debye limits are only approximate and must be treated carefully using physics beyond what we have been considering here.

It would have been nice if all of this had been the full story on the perturbation of two-quasiparticle states. However, as we saw for both the ground state and one-particle states, we still need to consider intermediate states that pair off with $T_i$ that don't fully pair off with the creation operators in the definition of $|n^{(0)}\rangle$. Specifically, we want to consider first the $T_i$ with only one annihilation operator that pairs off with one of the two-quasiparticle creation operators for $|n^{(0)}\rangle$, and next the $T_i$ with no annihilation operators (similar to what we did with the ground state, for example).

The $T_i$'s with only one annihilation operator are $T_2$, $T_3$, $T_8$ and $T_{12}$. Consider $T_2 = -u^2u'v'\, b^\dagger_{k\uparrow}\, b^\dagger_{-k\downarrow}\, b^\dagger_{-p\downarrow}\, b_{-p\downarrow}$. The annihilation operator will pair off with $b^\dagger_{-r\downarrow}$, leaving 4 creation operators, $b^\dagger_{k\uparrow}\, b^\dagger_{-k\downarrow}\, b^\dagger_{-r\downarrow}\, b^\dagger_{r\uparrow}$. Those match against the intermediate state $\langle m^{(0)}|b_{r\uparrow}\, b_{-r\downarrow}\, b_{-k\downarrow}\, b_{k\uparrow}$. Then, in order to match $|m^{(0)}\rangle$, we'll take $T_9$ in $\langle n^{(0)}|T_9|m^{(0)}|$. The result will be a single sum over $k$ and an overall factor of $\kappa^4$. As discussed previously, that will lead to an overall suppression of $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$. A similar result is obtained with the combination of $T_3$ and $T_5$ as well as the other cases involving $T_8$ and $T_{12}$.

With $T_i$'s with no annihilation operators, the situation is almost identical to what we encountered when looking at those terms for the perturbations of single-quasiparticle states. Namely, we end up with the same sum as we did for the ground state except that we need to exclude the terms in the sum that match the momenta $r$. As before, we end up with a contribution suppressed by $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$.

- Our last example, will be one where the unperturbed state $|n^{(0)}\rangle$ consists of $M$ quasiparticles with spin up, and $M$ with opposite spin and opposite (i.e., the negative) momenta. Start the analysis by examining $\langle m^{(0)}|T_1|n^{(0)}\rangle$. Let $p$ match one of the spin-up quasiparticles, say

$(r\uparrow)$. That is, the $T_1$-factor $b_{-p\downarrow}b_{p\uparrow}$, pairs off with the factor $b^\dagger_{-p\downarrow}b^\dagger_{p\uparrow}$ in the specification of $|n^{(0)}\rangle$. The remaining creation operators in the specification of $|n^{(0)}\rangle$ must match annihilation operators in the specification of $\langle m^{(0)}|$. There is one such pairing for each $k$ except when $k=p$. That pair is excluded because then $|n^{(0)}\rangle = |m^{(0)}\rangle$. The $p$-pairing can be with any one of the $M$ quasiparticles in $|n^{(0)}\rangle$, so in total there are $M\mathcal{N}$ terms (as usual, the factor $\langle n^{(0)})|T_i|m^{(0)}|\rangle$ is fully determined by matching). Proceeding, we see that the analysis will be exactly the same as what we did in the previous example, except that the $p$ terms can match any one of the $M$ momenta in $|n^{(0)}\rangle$. The correction $E_n^{(2)}$ becomes the sum over $\epsilon'_r$ of Eq. (17), where $\epsilon'_r$ are the momenta of the quasiparticles in $|n^{(0)}\rangle$. In the previous example, we considered four cases for $\epsilon'_r$. In the present example, each of the quasiparticle momenta could be in any one of those four cases (of course, they could be anything else but I just want to look at the analogue of the previous example). In the previous example, for the first two cases, $\epsilon'_r \geq \frac{\Delta}{2}$ and $E_n^{(2)} \approx \alpha\frac{\Lambda^2\omega}{\mathcal{N}}$ for some constant $\alpha = \mathcal{O}(1)$. In the present example, this number would be multiplied by $M$. Let's take $M = \frac{\mathcal{N}}{4}$ (so, including both spins, we have a total of $\frac{\mathcal{N}}{2}$ quasiparticles). Then $E_n^{(2)}$ for case 1 (if all quasiparticle energies were similar), would be $\frac{\Lambda^2\omega}{4}\left(\sqrt{5}\frac{\Delta}{\omega} + 2\log(2\frac{\omega}{\Delta})\right)$. For case 2 (if all quasiparticle energies were similar), $E_n^{(2)} \approx 2.3\Lambda^2\omega$. Before continuing to cases 3 and 4, notice that the relative correction requires that we divide by $E_N^{(0)}$. Each quasi-particle has an energy $> \Delta$ and there are $M = \frac{\mathcal{N}}{2}$ quasiparticles, so for both cases 1 and 2, the relative energies are $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$. What about case 3? In the previous example, we looked at $\epsilon'_r = \delta'$. However, since all quasiparticles must have different energies, and since there are $\frac{\mathcal{N}}{2}$ quasiparticles, most energies will be far greater than $\delta'$. Even if all energies were $\delta'$, the relative energy-correction would be $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$ as before. Similarly for case 4, most energies would be far less than the Debye energy. However, if all energies were approximately the Debye energy, then the relative correction (taking into account an overall ground-state correction of of $\kappa^4\Delta^4\sum_{kp}\frac{1}{2(E_k+E_p)E_p^2E_k^2}$) would be $\mathcal{O}\left(\frac{\log\mathcal{N}}{\mathcal{N}}\right)$.

What about contributions to $\langle m^{(0)}|T_i|n^{(0)}\rangle$ from $T_i$ with only one or no annihilation operators? The analysis is similar to what we did for the two-particle case and the perturbation is suppressed by $\mathcal{O}\left(\frac{\log\mathcal{N}}{\mathcal{N}}\right)$.

In all of these examples, the second-order correction is negligible.

### 3.4.8 The $q$-dependent potential – notation change

We turn to the term in the Hamiltonian $V_Q = -\frac{1}{2}\sum'_{\substack{pqk \\ \sigma_1=\sigma_3\sigma_2=\sigma_4}} c^\dagger_{(k+q)\sigma_2} c_{k\sigma_3} c_{p\sigma_4}$ where $\sum'$ means "exclude momenta combinations already included in $H_{CP}$". An alternate summation convention is obtained by the following substitutions of the summation indices: $p \to (q+k)$; $k \to (-q-k')$; $q \to q$, which leads to $V_Q = -\frac{1}{2}\kappa^2 \sum'_{\substack{kk'q \\ \sigma_1=\sigma_3\sigma_2=\sigma_4}} c^\dagger_{k\sigma_1} c^\dagger_{-k'\sigma_2} c_{(-k'-q)\sigma_3} c_{(k+q)\sigma_4}$ (where we have made the substitution discussed earlier of $V_q = \kappa^2$ and also reversed the order – with a sign change – of the the last two terms in the monomial). With this new convention, our notation follows that of Fetter and Walecka so that we can more easily refer to their derivations.

### 3.4.9 The $q$-dependent potential – Wick expansion

In the following sections, we will treat the quartic and quadratic normal-ordered contributions separately. To make this clear, I'll define operators $\tilde{O}$ to be the quadratic part of $O$, i.e. $O$ minus the quartic-normal-ordered part.

Recall the Wick-expanded $q$-independent Hamiltonian $H_{CP}$

$$\begin{aligned} H_{CP} = \sum_{p\sigma} \left(\epsilon_p - \kappa^2 \langle c^\dagger_{-p\sigma} c_{-p\sigma}\rangle_{CP}\right) c^\dagger_{p\sigma} c_{p\sigma} - \kappa^2 \sum_{pk} [&: c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} :_{CP} \\ + \ &\langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}\rangle_{CP} : c_{-p\downarrow} c_{p\uparrow} :_{CP} + : c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} :_{CP} \langle c_{-p\downarrow} c_{p\uparrow}\rangle_{CP} \\ + \ &\langle c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}\rangle_{CP} \langle c_{-p\downarrow} c_{p\uparrow}\rangle_{CP} + \langle c^\dagger_{k\uparrow} c_{p\uparrow}\rangle_{CP} \langle c^\dagger_{-k\downarrow} c_{-p\downarrow}\rangle_{CP}] \end{aligned}$$

and also recall that $H_0 = \sum_{p\sigma} \epsilon_p c^\dagger_{p\sigma} c_{p\sigma}$. Define $V \equiv \tilde{H}_{CP} - H_0 - \sum_{p\sigma} \kappa^2 \langle c^\dagger_{-p\sigma} c_{-p\sigma}\rangle_{CP} c^\dagger_{p\sigma} c_{p\sigma}$ where $\tilde{H}_{CP} = H + \kappa^2 \sum_{pk} : c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} c_{-p\downarrow} c_{p\uparrow} :_{CP}$. For the $q$-dependent Hamiltonian, define $\tilde{V}_Q = -\frac{1}{2}\kappa^2 \sum_{kk'q\sigma_1\sigma_2\sigma_3\sigma_4} c^\dagger_{k\sigma_1} c^\dagger_{-k'\sigma_2} c_{(-k'-q)\sigma_3} c_{(k+q)\sigma_4} + V_{QNO}$ (i.e. the full $q$-dependent potential minus the normal-ordered quartic part). The term $\tilde{V}_Q - \tilde{V}$ is the perturbation, due to the quadratic-part of the potential, to $H_{CP}$.

Expand $\tilde{V}$ in terms of $b$-operators.

$$\tilde{V} = \sum_p \Big[ \left( -\left(u_p^2 - v_p^2\right)\left(\kappa^2 v_p^2\right) + 2u_p v_p \Delta \right) \left( b_{p\uparrow}^\dagger b_{p\uparrow} + b_{-p\downarrow}^\dagger b_{-p\uparrow} \right) \\ - \left( u_p v_p (\kappa^2 v_p^2) \left(u_p^2 - v_p^2\right) \right) \left( b_{-p\downarrow} b_{p\uparrow} + b_{p\uparrow}^\dagger b_{-p\downarrow}^\dagger \right) \Big] + \text{ c-numbers}$$

When looking at quadratic terms, it will be more convenient for us to work with $\tilde{V}_Q$ and then subtract $V$, rather than worrying about which summation indices to exclude if we directly worked with $V_Q$. Another advantage of working with $\tilde{V}_Q$ is that Fetter and Walecka have done the work for us. They obtain

$$\tilde{V}_Q = -\kappa^2 \sum_{kk'q} \big[ \delta_{q,0} v_k^2 : c_{-k'\downarrow}^\dagger c_{-k'\downarrow} :_{BCS} + \delta_{q,0} v_{k'}^2 : c_{k\uparrow}^\dagger c_{k\uparrow} :_{BCS} + \\ \delta_{k,k'} u_k v_k : c_{(-k-q)\downarrow} c_{(k+q)\uparrow} :_{BCS} + \delta_{k,k'} u_{k+q} v_{k+q} : c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger :_{BCS} + \\ \delta_{q,0} v_k^2 v_{k'}^2 + \delta_{k,k'} u_k v_k u_{k+q} v_{k+q} \big].$$

Next, expand the $c$ operators in terms of the $b$ operators. Again, this is done by Fetter and Walecka.[12]

$$\tilde{V}_Q = -\kappa^2 \sum_{kk'} \big[ u_k v_k u_{k'} v_{k'} + v_k^2 v_{k'}^2 + \left( b_{k\uparrow}^\dagger b_{k\uparrow} + b_{-k\downarrow}^\dagger b_{-k\downarrow} \right) \left( \left(u_k^2 - v_k^2\right) v_{k'}^2 - 2u_k v_k u_{k'} v_{k'} \right) \\ + \left( b_{k\uparrow}^\dagger b_{-k\downarrow}^\dagger + b_{-k\downarrow} b_{k\uparrow} \right) \left( \left(u_k^2 - v_k^2\right) u_{k'} v_{k'} + 2u_k v_k v_{k'}^2 \right) \big] \\ = \sum_p \Big[ \left( -\left(u_p^2 - v_p^2\right)\alpha + 2u_p v_p \Delta \right) \left( b_{p\uparrow}^\dagger b_{p\uparrow} + b_{-p\downarrow}^\dagger b_{-p\downarrow} \right) \\ - \left( \left(u_p^2 - v_p^2\right)\Delta + 2u_p v_p \alpha \right) \left( b_{-p\downarrow} b_{p\uparrow} + b_{p\uparrow}^\dagger b_{-p\downarrow}^\dagger \right) \Big] + \text{ c-numbers}$$

where $\alpha = \kappa^2 \sum_{\boldsymbol{k}} v_k^2$.

Finally, the perturbative correction $\delta\tilde{V}_Q$ is

$$\delta\tilde{V}_Q \equiv \tilde{V}_Q - \tilde{V} = \sum_p \alpha^{(p)} \left[ -\left(u_p^2 - v_p^2\right) \left( b_{p\uparrow}^\dagger b_{p\uparrow} + b_{-p\downarrow}^\dagger b_{-p\downarrow} \right) + 2u_p v_p \left( b_{-p\downarrow} b_{p\uparrow} + b_{p\uparrow}^\dagger b_{-p\downarrow}^\dagger \right) \right] \tag{18}$$

where $\alpha^{(p)} = \kappa^2 \sum_{k \neq p} v_k^2$

[12]Fetter and Walecka are much more careful about the Debye limits than I have been. In particular, they impose the limits implicitly by including in the Hamiltonian, factors that we've ignored – for example $\langle (k)(-k') | \mathcal{V} | (k+q)(-k'-q) \rangle$. The suppression of Cooper pairs beyond the Debye limit, then arises from the inner products of $\mathcal{V}$, which allows us to change $\sum_{k,k',q} \delta_{k,k'} \langle k(-k') | \mathcal{V} | (k+q)(-k'-q) \rangle \mathcal{I}_{k'+q}$ to $\sum_{k,k'} \langle (k)(-k) | \mathcal{V} | (k')(-k') \rangle \mathcal{I}_{k'}$, where $\mathcal{I}_{k'+q}$ represents a portion of the integrand.

### 3.4.10 The $q$-dependent quadratic term – exact solution

Lancaster showed us how to write the $b$ ladder operators in terms of other ladder operators $\bar{b}$ so that the resulting Hamiltonian $\bar{H}_0$ is in canonical form $\bar{H}_0 = \sum_{p\sigma} \bar{E}_p \bar{b}^\dagger_{p\sigma} \bar{b}_{p\sigma}$+ c-numbers. Fetter and Walecka do the calculation and obtain

$$\bar{E}_p = \sqrt{(\epsilon_p - \alpha)^2 + \hat{\Delta}^2} \tag{19}$$

where

$$\bar{u}_p = \frac{1}{\sqrt{2}} \sqrt{1 + \frac{\epsilon_p - \alpha}{\bar{E}_p}}$$

$$\bar{v}_p = \frac{1}{\sqrt{2}} \sqrt{1 - \frac{\epsilon_p - \alpha}{\bar{E}_p}}$$

$$\bar{\Delta} = \kappa^2 \sum_p \bar{u}_p \bar{v}_p$$

The definition of $\bar{\Delta}$ is a bit deceiving because as explained earlielr, the self-consistency equation leads to the equation $\frac{1}{\Lambda} = \sinh^{-1} \omega/\bar{\Delta}$, which is precisely the same equation as for $\Delta$. So $\bar{\Delta} = \Delta$.

Notice that for this exact solution, the ground state differs from the ground state, $|BCS\rangle$ of $H_{CP}$. As usual, the energy of the ground state is 0 and all $c$-number contributions are ignored since they change every state by the same amount.

### 3.4.11 The $q$-dependent quadratic terms – first order corrections

We'll start by applying first-order perturbation theory to $\delta\tilde{V}_Q$ defined in Eq. (18).

- Take $|n^{(0)}\rangle = |BCS\rangle$ Then $E^{(1)}_n = \langle BCS|\delta\tilde{V}_Q|BCS\rangle = 0$.
- Next take the one-quasiparticle state, $|n^{(0)}\rangle = b^\dagger_{r\sigma}|BCS\rangle$. Then $E^{(1)}_n = \langle BCS|\delta\tilde{V}_Q|BCS\rangle = -\alpha^{(r)}(u_r^2 - v_r^2) = -\alpha^{(r)}\frac{\epsilon_r}{E_r}$ where to first order in $\kappa^2$, $\epsilon'_r \approx \epsilon$.
- Finally consider an M-quasiparticle state, $|n^{(0)}\rangle = \prod_{n=1}^M b^\dagger_{r_n\sigma_n}|BCS\rangle$. In that case, $E^{(1)}_n = \langle BCS|\delta\tilde{V}_Q|BCS\rangle = -\sum_{r=1}^M \alpha^{(r)}(u_r^2 - v_r^2) =$

$-\sum_{r=1}^{M} \alpha^{(r)} \frac{\epsilon_r}{E_r}$.

Compare the perturbative results to the results we obtain by expanding the exact quadratic solution.

- The vacuum state has 0 energy to first order.
- The one-particle state has energy $\bar{E}_r = \sqrt{(\epsilon_r - \alpha)^2 + \Delta^2}$ (we've used the fact that $\bar{\Delta} = \Delta$). We expand in $\kappa^2$ to obtain $\bar{E}_r = \sqrt{(\epsilon_r - \kappa^2 v_r^2)^2 + \Delta^2}\sqrt{1 + \frac{-2\alpha^{(r)}\epsilon_r + \left(\alpha^{(r)}\right)^2}{(\epsilon_r - \kappa^2 v_r^2)^2 + \Delta^2}} \approx E_r \left(1 - \alpha^{(r)} \frac{\epsilon_r}{E_r^2}\right)$. This agrees with what we obtained in first-order perturbation theory.
- Since the $M$-quasiparticle state has an energy which is the sum of the energies of the individual quasi-particle states, the expansion in $\kappa^s$ is $\sum_r E_r \left(1 - \alpha^{(r)} \frac{\epsilon_r}{E_r^2}\right)$, again what we obtained in first-order perturbation theory.

### 3.4.12 The $q$-dependent quadratic terms – second order corrections

Now we'll apply second-order perturbation theory to $\delta\tilde{V}_Q$ defined in Eq. (18).

- Take $|n^{(0)}\rangle = |BCS\rangle$. Then $E_n^{(2)} = -\sum_m \frac{\langle BCS|\delta\tilde{V}_Q|m^{(0)}\rangle\langle m^{(0)}|\delta\tilde{V}_Q|BCS\rangle}{E_m^{(0)}}$. The only contributions to $\langle m^{(0)}|\delta\tilde{V}_Q|BCS\rangle$ come from the terms in $\delta\tilde{V}_Q$ proportional to $b_{p\uparrow}^\dagger b_{-p\downarrow}^\dagger$. They match up with the two-quasiparticle states (which have energy $2E_p$) $\langle m^{(0)}| = \langle BCS|b_{p\uparrow}b_{-p\downarrow}$ and therefore with $\langle BCS|b_{-p\downarrow}b_{p\uparrow}|m^{(0)}\rangle$. Altogether this gives $E_{BCS}^{(2)} = -4\sum_p \frac{u_p^2 v_p^2 \left(\alpha^{(p)}\right)^2}{2E_p}$.

  As discussed in Section 3.4.7, the ground state has been shifted, so in order to reset energies relative to a ground state with energy 0, we will need to add $\delta E_g = 4\sum_p \frac{u_p^2 v_p^2 \left(\alpha^{(p)}\right)^2}{2E_p}$ to the second-order perturbed energy of every state.

  There is one apparent glitch in this analysis. Remember that all of our energies $\epsilon_p$, prior to adjustments (such as what we have with $\epsilon'$), include a subtraction of the the Fermi energy $\epsilon_F$ (or equivalently, the chemical potential $\mu$). To make this clear, write $\epsilon = \epsilon^{(0)} - \epsilon_F$. Then

with this definition, we see that $\epsilon' \neq \epsilon$ and in particular, $\epsilon'$ has different integration boundaries than $\epsilon$. Strictly speaking, it might seem that we should be integrating the unaltered energy from $-\omega$ to $\omega$. But we will, in what follows integrate the adjusted energy from $-\omega$ to $+\omega$. The reason this turns out to be correct, is that the ground state adjustments are made to the entire Hamiltonian, and not just to the part of the Hamiltonian involving the Cooper interaction term. Because of that, the Fermi energy is also adjusted by the same amount that the BCS ground state is adjusted.

- Next take the one-quasiparticle state, $|n^{(0)}\rangle = b^\dagger_{r\sigma}|BCS\rangle$. $E^{(2)}_n = \sum_m \frac{\langle n^{(0)}|\delta\tilde{V}_Q|m^{(0)}\rangle\langle m^{(0)}|\delta\tilde{V}_Q|n^{(0)}\rangle}{E^{(0)}_n - E^{(0)}_m}$. Again, the only contributions[13]. $\langle m^{(0)}|\delta\tilde{V}_Q|BCS\rangle$ come from the terms in $\delta\tilde{V}_Q$ proportional to $b^\dagger_{p\uparrow}b^\dagger_{-p\downarrow}$. Together with the creation operator for $|n^{(0)}\rangle$, we have three creation operators $b^\dagger_{p\uparrow}b^\dagger_{-p\downarrow}b^\dagger_{r\sigma}$ so those match the state $\langle m^{(0)}| = \langle BCS|b_{r\sigma}b_{p\uparrow}b_{-p\downarrow}$. Notice – and this is critically important – that one term is missing from the sum over $p$. For example, if $\sigma = \uparrow$, then the term with $p = r$ is missing since that term is $b^\dagger_{r\uparrow}b^\dagger_{-r\downarrow}b^\dagger_{r\sigma} = 0$ (the square of a ladder operator is 0 because of the Pauli exclusion principle). Also, the energy denominator for $p$ is $E^{(0)}_r - 2E^{(0)}_p - E^{(0)}_r = -2E^{(0)}_p$. When we add everything up, we get the same result as for the previous example with $|n^{(0)}\rangle = |BCS\rangle$ except that the sum is missing the term $p = r$. That is, $E^{(2)}_n = -4\sum_p \frac{u^2_p v^2_p \left(\alpha^{(p)}\right)^2}{2E^{(0)}_p} + 4\frac{u^2_r v^2_r \left(\alpha^{(r)}\right)^2}{2E^{(0)}_r}$. The adjusted energy-perturbation (after adding $\delta E_g$) is $E^{(2)}_R = E^{(2)} + \delta E_g = 4\frac{u^2_r v^2_r \left(\alpha^{(r)}\right)^2}{2E^{(0)}_r}$. To estimate this value, notice that (once we substitute an integral for the sum) $\alpha^{(r)} \propto \kappa^2 g(\epsilon_F)\int_{-\omega}^{\omega} dk v^2_k \approx \Lambda\omega$. So $\left(\alpha^{(r)}\right)^2 \approx \Lambda^2\omega^2$. Since $u^2_r v^2_r = \frac{\Delta^2}{4\left(E^{(0)}_r\right)^2}$, we end up with $E^{(2)}_R \approx \frac{\Lambda^2\omega^2\Delta^2}{2(E^0_r)^3}$. **This is a problematic result!! For low-energy quasiparticles (e.g. $\epsilon'_r < \Delta$), with typical experimental values for $\Lambda$, $\Delta$ and $\omega$, the relative perturbation $\frac{E^{(2)}_R}{E^{(0)}_r} \approx \frac{\Lambda^2\omega^2}{2\Delta^2}$ has a value of $\approx 50$.**

- Finally consider an M-quasiparticle state, $|n^{(0)}\rangle = \prod_{n=1}^{M} b^\dagger_{r_n\sigma_n}|BCS\rangle$. Following virtually the identical steps to the previous example, we ar-

[13] It would appear that there should be contributions from terms in $\delta\tilde{V}_Q$ proportional to $b^\dagger_{p\uparrow}b_{p\downarrow}$. However, the only states $|m^{(0)}\rangle$ that could contribute in that case, would be the states $|m^{(0)}\rangle = |n^{(0)}\rangle$

rive at $E_R^{(2)} = E^{(2)} + \delta E_g = 4\sum_{r=1}^{M} \frac{u_r^2 v_r^2 \left(\alpha^{(r)}\right)^2}{2E_r}$. **Just as in the 1-quasiparticle case, the relative perturbation – for small energies – is very large.**

Now we can expand the exact quadratic solution through order $\kappa^2$. Let's do it for each quasiparticle. Recall that $\bar{E}_r = E_r\sqrt{1 + \frac{-2\alpha^{(r)}\epsilon_r + \left(\alpha^{(r)}\right)^2}{(\epsilon_r - \kappa^2 v_r^2)^2 + \Delta^2}}$. A formal Taylor expansion in $\kappa^2$ of the squareroot leads to $E_r \approx E_r + \frac{\left(\alpha^{(r)}\right)^2}{2E_r^2} - \frac{\left(\alpha^{(r)}\right)^2 \epsilon_r^2}{2E_r^4} \approx E_r + \left(\alpha^{(r)}\right)^2 \frac{E_r - \epsilon_r^2}{2E_r^3} \approx E_r + \left(\alpha^{(r)}\right)^2 \frac{\Delta^2}{2E_r^3}$ where the last equality ignores a term of higher order in $\kappa^2$. The total energy for the $M$ particles is obtained by summing over the values of $r$ for those $M$ particles. The result agrees with our perturbative result. **HOWEVER, what we just did was a *formal* Taylor expansion. For some energy values, the perturbative" term is, as we've seen above, is large, and the series will not converge.**

What our analysis appears to show, is that in a formal perturbative expansion, the exact result agrees – at least through second order – with standard perturbation theory. But what the analysis also shows, is that standard perturbation theory would not be applicable and without an exact solution, we wouldn't arrive at a physically applicable result. Fortunately, we can use the exact solution, at least for the quadratic terms. What remains to be seen is the size of perturbative contributions for the residual potential.

### 3.4.13 The $q$-dependent residual potential – expansion in terms of $b$ operators

The normal-ordered quartic potential $V_{QNO} = \kappa^2 \sum_{pqk} : c^\dagger_{(p-q)\uparrow} c^\dagger_{(k+q)\downarrow} c_{p\uparrow} c_{k\downarrow} :_{BCS}$ is expressed using the $c$ operators and can therefore be expanded as a sum of quartic terms using the $b$ operators.

An alternate summation convention is obtained by the following substitutions of the summation indices: $p \to (q+k)$; $k \to (-q-p)$; $q \to q$, which leads to

$$V_{QNO} = -\kappa^2 \sum_{pqk} c^\dagger_{k\uparrow} c^\dagger_{-p\downarrow} c_{(-p-q)\downarrow} c_{(k+q)\uparrow}$$

(where we have reversed the order – with a sign change – of the the last two terms in the monomial). With this new convention, our notation follows that of Fetter and Walecka (allowing for easier comparisons later). Simplify

notation by the following assignments.

$$a = b_{k\uparrow}, \qquad b = b_{-k\downarrow}, \qquad c = b_{p\uparrow}, \qquad d = b_{-p\downarrow},$$
$$e = b_{p+q\uparrow}, \qquad f = b_{-p-q\downarrow}, \qquad g = b_{k+q\uparrow}, \qquad h = b_{-k-q\downarrow},$$

Also introduce $u_i, v_i$ in the expansions

$$c^\dagger_{k\uparrow} = u_1 a^\dagger + v_1 b, \qquad c^\dagger_{-p\downarrow} = u_2 d^\dagger - v_2 c,$$
$$c_{(-p-q)\downarrow} = u_3 f - v_3 e^\dagger, \quad c_{(k+q)\uparrow} = u_4 g + v_4 h^\dagger,$$

with

$$u_1 \equiv u_k, \qquad v_1 \equiv v_k,$$
$$u_2 \equiv u_p, \qquad v_2 \equiv v_p,$$
$$u_3 \equiv u_{p+q}, \quad v_3 \equiv v_{p+q},$$
$$u_4 \equiv u_{k+q}, \quad v_4 \equiv v_{k+q},$$

The expanded and normal-ordered result is

$$\boxed{V_{QNO} = -\kappa^2 \sum_{i=1}^{16} T_i}$$

with the $T_i$ given below.

$$
\begin{aligned}
T_1 &= \; u_1u_2u_3u_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{-p\downarrow}\, b_{-p-q\downarrow}\, b_{k+q\uparrow},\\
T_2 &= -\, u_1u_2u_3v_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{-p\downarrow}\, b^\dagger_{-k-q\downarrow}\, b_{-p-q\downarrow},\\
T_3 &= -\, u_1u_2v_3u_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{-p\downarrow}\, b^\dagger_{p+q\uparrow}\, b_{k+q\uparrow},\\
T_4 &= -\, u_1u_2v_3v_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{-p\downarrow}\, b^\dagger_{p+q\uparrow}\, b^\dagger_{-k-q\downarrow},\\
T_5 &= -\, u_1v_2u_3u_4 \; b^\dagger_{k\uparrow}\, b_{p\uparrow}\, b_{-p-q\downarrow}\, b_{k+q\uparrow},\\
T_6 &= -\, u_1v_2u_3v_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{-k-q\downarrow}\, b_{p\uparrow}\, b_{-p-q\downarrow},\\
T_7 &= -\, u_1v_2v_3u_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{p+q\uparrow}\, b_{p\uparrow}\, b_{k+q\uparrow},\\
T_8 &= +\, u_1v_2v_3v_4 \; b^\dagger_{k\uparrow}\, b^\dagger_{p+q\uparrow}\, b^\dagger_{-k-q\downarrow}\, b_{p\uparrow},\\
T_9 &= -\, v_1u_2u_3u_4 \; b^\dagger_{-p\downarrow}\, b_{-k\downarrow}\, b_{-p-q\downarrow}\, b_{k+q\uparrow},\\
T_{10} &= -\, v_1u_2u_3v_4 \; b^\dagger_{-p\downarrow}\, b^\dagger_{-k-q\downarrow}\, b_{-k\downarrow}\, b_{-p-q\downarrow},\\
T_{11} &= +\, v_1u_2v_3u_4 \; b^\dagger_{-p\downarrow}\, b^\dagger_{p+q\uparrow}\, b_{-k\downarrow}\, b_{k+q\uparrow},\\
T_{12} &= -\, v_1u_2v_3v_4 \; b^\dagger_{-p\downarrow}\, b^\dagger_{p+q\uparrow}\, b^\dagger_{-k-q\downarrow}\, b_{-k\downarrow},\\
T_{13} &= -\, v_1v_2u_3u_4 \; b_{-k\downarrow}\, b_{p\uparrow}\, b_{-p-q\downarrow}\, b_{k+q\uparrow},\\
T_{14} &= +\, v_1v_2u_3v_4 \; b^\dagger_{-k-q\downarrow}\, b_{-k\downarrow}\, b_{p\uparrow}\, b_{-p-q\downarrow},\\
T_{15} &= +\, v_1v_2v_3u_4 \; b^\dagger_{p+q\uparrow}\, b_{-k\downarrow}\, b_{p\uparrow}\, b_{k+q\uparrow},\\
T_{16} &= +\, v_1v_2v_3v_4 \; b^\dagger_{p+q\uparrow}\, b^\dagger_{-k-q\downarrow}\, b_{-k\downarrow}\, b_{p\uparrow}.
\end{aligned}
\tag{20}
$$

As in Section 3.4.7 we could show that dependence as $T_i(p,k,q)$ so that we would write $V_{NO} = -\kappa^2 \sum_{pk} \sum_{i=1}^{16} T_i(p,k,q)$. In what follows, we usually suppress the $p,k,q$ arguments of $T_i$.

### 3.4.14 The $q$-dependent residual potential – first order corrections

This section is partly a cut-and-paste from section 3.4.6. We'll start with the first-order correction to energy eigenvalues. Each state $|n^{(0)}\rangle$ is of the form $\prod_{(k_i)} b^\dagger_{k_i}|0\rangle_{BCS}$ or ${}_{BCS}\langle 0| \prod_{(k_i)} b_{k_i}$ for some collection $(k_i)$. The annihilation

operators (in terms of the $b$ ladder operators) of $V_{QNO}$ operate on the right and the creation operators act on the left.

First consider an excited state consisting of exactly one quasiparticle, e.g. $b^\dagger_{r\uparrow}|BCS\rangle$. The first-order energy correction is $\langle BCS|b_{r\uparrow}V_{QNO}b^\dagger_{r\uparrow}|BCS\rangle$. The only possible non-zero contributions would require a monomial in $V_{QNO}$ which includes $b^\dagger_{r\uparrow}$ and $b_{r\uparrow}$. These would be required in order to pair off with the excited quasiparticle for both the 'bra' and 'ket'. There remain two unpaired operators in that monomial (since the monomial is a quartic) and because they are normal-ordered, their $|BCS\rangle$ expectation value is 0.

What about states with two quasiparticles? First consider a state of the form $b^\dagger_{r\uparrow}b^\dagger_{r'\uparrow}|BCS\rangle$, where both quasiparticles have the same spin. The only contributing monomials of $V_{QNO}$ would be ones where all 4 operators have that spin. Moreover, there would need to be two annihilation operators and two creation operators. The only candidate is term $T_7 = -\,u_1v_2v_3u_4\; b^\dagger_{k\uparrow}b^\dagger_{(p+q)\uparrow}\,b_{p\uparrow}\,b_{(k+q)\uparrow}$. By matching annihilation and creation operators, we see that $p$ matches $r$ or $r'$, and $k+q$ matches $r'$ or $r$. Furthermore $k$ matches $r$ or $r'$ and $p+q$ matches $r'$ or $r$. All four combinations correspond to a unique solution for $p, q, k$ in terms of $r, r'$, and also $u_1, v_2, v_3, u_4$ in terms of $u_r, v_r, u_{r'}, v_{r'}$. Therefore the triple-sum collapses to 4 individual summands, and if we arbitrarily set the $uvuv$-coherence factors to be $\approx 1$ (they are less than 1) the expectation value for this state, is $-4\kappa^2 \approx -\frac{8\Lambda\omega}{\mathcal{N}}$. Moreover, to obtain the relative correction, we must divide by the energy of the unperturbed quasi-particle state. Since there are two quasiparticles, that energy $> 2\Delta$. Taking the ratio, we find that **the relative correction is** $\approx \frac{4\Lambda\omega}{\Delta\mathcal{N}} = \mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$**. This is negligibly small.** For opposite spin the only candidate is $T_{10}$ which, in the same way, leads to a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$.

Now consider the case of states $|n^{(0)}\rangle$ in which there are two quasiparticles of opposite spin, $b^\dagger_{r\uparrow}b^\dagger_{r'\downarrow}|BCS\rangle$. (We will permit $r' = -r$ since the analysis is independent of whether the total momentum is 0 or not). In the triple-sum, we would require annihilation operators to pair with $b^\dagger_{r\uparrow}$ and $b^\dagger_{r'\downarrow}$ as well as creation operators to pair with $b_{r\uparrow}$ and $b_{r'\downarrow}$. This can happen in $V_{QNO}$ with the terms $T_1, T_6, T_{11}, T_{16}$. Start by looking at $T_1 = u_1u_2u_3u_4\; b^\dagger_{k\uparrow}b^\dagger_{-p\downarrow}\,b_{(-p-q)\downarrow}\,b_{(k+q)\uparrow}$. By setting the pair-matches we have $k+q = r$, $p+q = -r'$, $k = r$, $p = -r'$. These equations can be solved uniquely by $q = 0$, $k = r$, $p = -r'$. The triple sum collapses to a single term and as usual, we obtain a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$. Next look at $T_6 = -\,u_1v_2u_3v_4\; b^\dagger_{k\uparrow}b^\dagger_{(-k-q)\downarrow}\,b_{p\uparrow}\,b_{(-p-q)\downarrow}$. By setting pair-matches we have

$p = r$, $p + q = -r'$, $k = r$, $k + q = -r'$. These equations can be solved uniquely by $p = k = r$, $q = -(r + r')$. Again, a relative correction of $\mathcal{O}(\frac{1}{\mathcal{N}})$. Similar results hold for $T_{11}$ and $T_{16}$.

Notice that the above technique involves solving 4 equations in 3 unknowns. In each case, any subset of 3 equations turned out to be independent and could therefore be solved with exactly one value of $p, k, q$. Moreover, all subsets of 3 equations had the same solutions. I believe it's possible that all subsets of equations could be solvable but that some solutions would differ from the others. If that were to happen, then this would mean there was no solution and therefore that the $T_i$ contribution is 0. Since I'm only interested here in whether perturbative contributions would be large, I am willing to ignore any contributions that are 0. There therefore isn't any reason (here) for checking whether the equations are solvable. On the other hand, if the determinant were 0 for each set of 3 equations, that would mean we could have an undetermined variable. For example, it might turn out that a solution can be found for any value of $p$. If so, there would be $\mathcal{N}$ (or more) summands contributing to the first-order correction so this could lead to a cancellation of the $\frac{1}{\mathcal{N}}$ which has appeared in all of our results. **What can be shown is that for each $T_i$ there isn't any situation with an undetermined variable. That is, all sets of 3 equations have nonzero determinants.** (That result was obtained by explaining all this to ChatGPT and having it do the calculations.)

We are now ready to tackle states with more than two quasiparticles. There are several different kinds of situations. Since my objective is to estimate the size of perturbations, I'll only consider a representative example and will generalize from there.

Let $|n^{(0)}\rangle = b^\dagger_{r_1\uparrow} b^\dagger_{r_2\downarrow} b^\dagger_{r_3\uparrow} b^\dagger_{r_4\uparrow} |BCS\rangle$. All momenta are different from one another, otherwise by the Pauli exclusion principle the result would be 0. Take any $T_i$. If we pair off an unequal number of creation and annihilation operators, our expectation value will be left with an unequal number of annihilation and creation operators (since it began with an equal number prior to matching – creation operators for the 'ket' and their adjoints for the 'bra'). But in that case, the expectation value is 0. If we have an equal number of annihilation and creation operators in $T_i$, then we can choose to match the annihilation operators with any pair (provided spins match) out of $r_1, r_2, r_3, r_4$ and we can match the creation operators with any pair – possibly different. However, if the two matching pairs are different, than the remaining unmatched operators have different momenta, can anticommute and then annihilate the

'ket' (i.e. $\langle BCS|\alpha\beta^\dagger|BCS\rangle = -\langle BCS|\beta^\dagger\alpha|BCS\rangle = 0$ if $\alpha$ and $\beta$ are annihilation operators ). As a result, the $T_i$ operators pair off with at most 2 of the quasiparticles in $|n^{(0)}\rangle$, and the remaining operators pair off with each other. As an example, take $T_1$ whose annihilation operators have opposite spin. For $|n^{(0)}\rangle$ above, the possible pairs are $(r_1, r_2), (r_3, r_2)$ and $(r_4, r_2)$. If we took $T_7$ where both annihilation operators have spin up, the possible pairs are $(r_1, r_3), (r_1, r_4)$ and $(r_3, r_4)$. Consider $(r_1, r_3)$. one set of equations to solve is $p = r_1, k + q = r_3$. Those match the annihilation operators in $T_7$ with the $r_1$ and $r_3$ quasiparticles in $|n^{(0)}\rangle$. Then the creation operators in $T_7$ must match the $r_1$ and $r_3$ quasiparticles in $\langle n^{(0)}|$. So, for example, $k = r_1, p + q = r_3$. These equations have a single solution, but there are 4 different ways for matching to occur. A similar situation occurs for $T_1, T_6, T_{11}, T_{16}$ although only one matching can occur (because the spins distinguish the particles). On the other hand, if we took $T_{11}$ with all spins down, there are no candidate pairs (so a contribution of 0). Putting all this together, and setting the $u$ and $v$ values to 1, which is an upper bound, we find that the relative corrections for those terms are $\approx 24\frac{\Lambda}{g(\epsilon_F)E_n} = \frac{48\Lambda\omega}{\mathcal{N}E_n} = \mathcal{O}(\frac{1}{\mathcal{N}})$.

Now we can generalize. Let $|n^{(0)}\rangle$ be a state with $M$ quasiparticles of various spins. The maximum amount of pairing if all spins were up, would occur with $T_7$, in which case there would be a total of $4\binom{M}{2} = 2M(M-1)$ and a relative correction of $\approx 2\kappa^2 M^2$. With mixed spins, the maximum amount of pairing (per $T_i$) would be less. For example, assume half the spins are different. The contributing terms of the potential are $T_1, T_6, T_7, T_{10}, T_{11}, T_{16}$. For $T_7$ and $T_{11}$ there are each $\approx \frac{M^2}{2}$ pairings. For the others, there are each $\frac{M^2}{4}$ pairings. So the total correction is $\approx 2\kappa^2 M^2$.

So far, we haven't considered the $u$ and $v$ contributions. When $\epsilon' \gg \Delta$, $u_{\epsilon'} \approx 1$ and $v_{\epsilon'} \approx 0$ and when $\epsilon' \ll -\Delta$, $v_{\epsilon'} \approx 1$ and $u_{\epsilon'} \approx 0$. For large values of $M$, the majority of quasiparticles have $|\epsilon'| \gg \Delta$ so either $u_{\epsilon'}$ or $v_{\epsilon'}$ are $\approx 0$. All $T_i$ of interest, except for $T_1$ and $T_{16}$ have at least one $u$ factor and one $v$ factor, so those contributions are suppressed by $\approx \frac{\Delta}{\epsilon'}$ for $|\epsilon'| \gg \Delta$. One of either $T_1$ (all $u$'s) or $T_{16}$ (all $v$'s) is also suppressed (in fact, by $\mathcal{O}\left(\frac{\Delta^2}{\epsilon'^2}\right)$), so instead of 6 unsuppressed contributions from $T_i$, only 1 contribution remains. Specifically, including the suppression, the total correction is $\approx \kappa^2 M^2 \left(\frac{1}{2} + \frac{\Delta}{\epsilon'}\right)$.

We also need to estimate $E_n$. This requires a bit more analysis. What interests us here, are states where $M$ is a significant fraction of $\mathcal{N}$. So let's say for example, that $M = \frac{1}{2}\mathcal{N}$. Since two quasiparticles (they are fermions)

can't occupy the same state, the most energetic[14] quasiparticle must have $|\epsilon'| \geq \frac{\omega}{2}$. This follows from the fact that near the Fermi energy $\epsilon_F$, the density (which takes into account momentum-degeneracy) is approximately constant. So if $M = \frac{\mathcal{N}}{2}$, that means half of the states in the Debye window are occupied, so the most energetic quasiparticle has to have $|\epsilon'|$ more than half the Debye energy (relative to $\epsilon_F$). For simplicity, assume that all quasiparticles have $|\epsilon'|$ less than $\omega/2$. Then we can use the density of states $g(\epsilon_F)$ to estimate the number of quasiparticles at any $|\epsilon'|$ below $\omega/2$. Each quasiparticle has energy $E_{\epsilon'} = \sqrt{\epsilon'^2 + \Delta^2}$. Recall that $\Delta \ll \omega$ so most of the $|\epsilon'|$ values are much greater than $\Delta$. For those, we can estimate $E_{\epsilon'} \approx |\epsilon'|$. For $|\epsilon'|$ values $\ll \Delta$, we can estimate $E_{\epsilon'} \approx \Delta$. We can crudely estimate the total energy of $|n^{(0)}\rangle$ – with all these assumptions – as $E_n \approx g(\epsilon_F)\left(2\Delta^2 + 2\int_{\Delta}^{\frac{\omega}{2}} \epsilon d\epsilon\right) = \frac{\mathcal{N}}{2\omega}\left(\Delta^2 + \frac{\omega^2}{4}\right) \approx \frac{\mathcal{N}\omega}{8}$.

If, in our example where $M = \frac{\mathcal{M}}{2}$, we take the ratio of the suppressed corrections to the unperturbed energy $E_n$, we obtain the relative correction $\approx \frac{8M^2\kappa^2\left(\frac{1}{2}+\frac{\Delta}{\epsilon'}\right)}{\mathcal{N}\omega} \approx 4\Lambda\left(\frac{1}{2} + \frac{\Delta}{\epsilon'}\right)$. As mentioned earlier, in physical systems we can pick (as an example) $\Lambda \approx 0.2$ and $\frac{\Delta}{\omega} \approx .02$. With those values we see that the relative correction for this example, is $\approx 0.4$. Although this isn't as small as what we obtained for a first-order correction in the $q$-independent approximation, it can still be regarded as a reasonable first order perturbation.

### 3.4.15 The $q$-dependent residual potential – second order corrections

We now examine the second-order shifts:

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left|\langle m^{(0)}|V_{QNO}|n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}} \tag{21}$$

We repeat here some results from Section 3.4.7 and Appendix A that we will use in what follows: First a theorem.

[14] I'm misusing the word energetic or, for that matter the term $\epsilon'$ in this paragraph. Since $\epsilon' \approx \epsilon - \epsilon_F$, the value of $\epsilon'$ can be positive or negative. When I speak of $\epsilon'$ I really mean $|\epsilon'|$ and when I say energetic, I really mean the value of $|\epsilon'|$.

Define $b^0_{r\sigma} \equiv b_{r\sigma}$ and $b^1_{r\sigma} \equiv b^\dagger_{r\sigma}$. Then define $\mathcal{P} \equiv \langle BCS|b^{i_1}_{r_1\sigma_1} b^{i_2}_{r_2\sigma_2} ... b^{i_k}_{r_k\sigma_k} ... b^{i_n}_{r_n\sigma_n}|BCS\rangle$, where $i_m = 1$ or $i_m = 0$. If $n$ is odd, then $\mathcal{P} = 0$. If $n$ is even, then exactly $\frac{n}{2}$ of the indices $k$ have the value 0 and other half have the value 1. More precisely, half of the ladder operators are annihilation operators and the other half are their adjoints (creation operators). Otherwise, $\mathcal{P} = 0$.

Some other results: Near the Fermi surface, we have a dispersion relation $\epsilon' \approx \frac{p_F}{m^*}(p - p_F)$ where $p_F$ is the Fermi momentum $\sqrt{2m^*\epsilon_F}$, and $m^*$ is the effective mass. From the dispersion relation, we see that bare energy levels ($\epsilon'$) have a spacing of approximately $\delta\frac{p_F}{m^*} = \delta\frac{\sqrt{2\epsilon_F}}{\sqrt{m^*}}$. The average number of states between energy-levels (i.e. between constant-energy hypersurfaces) is $4\pi\frac{p^2}{\delta^2} \approx 4\pi\frac{p_F^2}{\delta^2} = \frac{8\pi m^*\epsilon_F}{\delta^2}$. The total number of states in the Debye window is then $\mathcal{N} \approx \frac{8\pi\omega}{\delta^3}\sqrt{2\epsilon_F m^{*3}}$. Recall that the density $g(\epsilon_F)$ is about $\frac{\mathcal{N}}{2\omega}$. From the expression we just derived for $\mathcal{N}$, we see that $g(\epsilon_F) \approx \frac{4\pi\sqrt{2\epsilon_F m^{*3}}}{\delta^3}$. Also, $\delta = \frac{\left(128\pi^2\epsilon_F m^{*3}\omega^2\right)^{\frac{1}{6}}}{\mathcal{N}^{\frac{1}{3}}}$.

Now examine some examples.

- We'll first look at the ground state, $|n^{(0)}\rangle = |BCS\rangle$ and will follow the approach described in Section 3.4.7/ As in that section, $T_4$ and $T_{13}$ terms (but from Eq. (20)) make contributions to the second-order perturbation. In fact, $E^{(2)}_{gs} = -\sum_{mkpqk'p'q'} \frac{\langle BCS|T_{13}(k',p',q')|m\rangle\langle m|T_4(k,p,q)|BCS\rangle}{E^{(0)}_m}$. For $\langle m|$ to match $T_4$, we must have $\langle m^{(0)}| = \langle BCS|b_{-k-q\downarrow}\, b_{p+q\uparrow}\, b_{-p\downarrow}\, b_{k\uparrow}$. Then $T_{13}$ matches $|m\rangle$ and that fixes $k'$,$p'$ and $q'$. By including the coherence factors of $T_{13}$ and $T_4$, we end up with $E^{(2)}_{gs} = -\kappa^4 \sum_{kpq} \frac{(u_p v_p)(u_k v_k)(u_{p+q} v_{p+q})\left(u_{k+q} v_{k+q}\right)}{\left(E_k+E_p+E_{p+q}+E_{k+q}\right)} = -\kappa^4\Delta^4 \sum_{kpq} \frac{1}{\left(E_k+E_p+E_{p+q}+E_{k+q}\right)E_p E_k E_{p+q} E_{k+q}}$. Unlike what we did in Section 3.4.7, I won't make any attempt to evaluate this sum. As before, we need to subtract this sum from all other eigen-energies so that they are all "renormalized" so the ground state has an adjusted energy of 0. However, it is worth noting that if we were to evaluate the triple sum, we would find that it is proportional to $\mathcal{N}$ – a huge number. Since that number is simply an overall constant in the theory, it doesn't matter that it is huge. What will matter is the difference between it, and related terms in the eigen-energy perturbations.
- Suppose $|n^{(0)}\rangle$ is a state with one quasiparticle, e.g. $|n^{(0)}\rangle = b^\dagger_{r\uparrow}|BCS\rangle$. Consider the term $T_1 = u_1u_2u_3u_4\ b^\dagger_{k\uparrow} b^\dagger_{-p\downarrow}\, b_{-p-q\downarrow}\, b_{k+q\uparrow}$. This has two

creation operators and two annihilation operators. The state $|n^{(0)}\rangle$ involves one creation operator. So any state $\langle m^{(0)}|$ which contributes to the sum in Eq. (21) must involve exactly one annihilation operator so that the inner product has 3 creation and 3 annihilation operators. The creation operator $b^\dagger_{r\uparrow}$ (applied to the BCS ground state) for $|n^{(0)}\rangle$ must be paired with an annihilation operator from $T_1$ (it can't be paired with the annihilation operator from $\langle m^{(0)}|$ because in Eq. (15), $\langle m^{(0)}| \neq \langle n^{(0)}|$. The only candidate with spin up, is $b_{k+q\uparrow}$. Therefore we must set $k+q=r$. So this transforms $|n^{(0)}\rangle$ to $|BCS\rangle$. Now we have a remaining annihilation operator on the right of $T_1$ and this annihilates $|BCS\rangle$, resulting in a 0 contribution.

Continuing with this example where $|n^{(0)}\rangle = b^\dagger_{r\uparrow}|BCS\rangle$, consider the term $T_2 \propto b^\dagger_{k\uparrow} b^\dagger_{-p\downarrow} b^\dagger_{-k-q\downarrow} b_{-p-q\downarrow}$. We see that this term has 3 creation operators and 1 annihilation operator. The annihilation operator has to pair with the creation operator $b^\dagger_{r\uparrow}$ (applied to the BCS ground state) for $|n^{(0)}\rangle$ but that's not possible because the $T_2$ annihilation operator is spin down. So $T_2$ has no contribution. On the other hand, the term $T_3$ doesn't have this problem because the one annihilation operator has spin up. So let's look at $T_3 \propto b^\dagger_{k\uparrow} b^\dagger_{-p\downarrow} b^\dagger_{p+q\uparrow} b_{k+q\uparrow}$. By pairing the annihilation operator, we have $k+q=r$. We now have 3 remaining unpaired creation operators in $T_3$ which must pair with 3 annihilation operators for $\langle m^{(0)}|$, so $\langle m^{(0)}| = \langle BCS| b_{r-q\uparrow} b_{-p\downarrow} b_{p+q\uparrow}$. From this we see that for every value – with the exception of $p+q=k$ which leads to 0 because $b_{k\uparrow}b_{k\uparrow}=0$ – of the summands $p$ and $q$ in Eq. (15), there is a state $\langle m^{(0)}|$ so that $\langle m^{(0)}|T_3|n^{(0)}\rangle \neq 0$. We're not quite done. Remember that in Eq. (21), the term $\langle m^{(0)}|T_3|n^{(0)}\rangle$ is multiplied by $\langle n^{(0)}|V_{QNO}|m^{(0)}\rangle$. Expanding the states, we see we need a monomial $T_i(p',k',q')$ for which $\langle BCS| b_{r\uparrow} \left(\sum_{p'k'q'} T_i(p',k',q')\right) b^\dagger_{p+q\uparrow} b^\dagger_{-p\downarrow} b^\dagger_{r-q\uparrow}|BCS\rangle \neq 0$. This requires that for some values of $(p',k',q')$, $T_i(p',k',q')$ is (up to a sign) $b^\dagger_{r\uparrow} b_{p+q\uparrow} b_{-p\downarrow} b_{r-q\uparrow}$. Both $T_5$ and $T_{15}$ are candidates for such a $T_i$. Take $T_5$ as an example. $T_5(p',k',q') = b^\dagger_{k'\uparrow} b_{p'\uparrow} b_{-p'-q'\downarrow} b_{k'+q'\uparrow}$. Since only one of the $T_5$ operators is a creation operator, it must be $b^\dagger_{r\uparrow}$, so $k'=r$. Also, since only one of the $T_5$ operators has spin down, it must be $b_{-p'\downarrow}$, so $p'+q'=p$. There are two remaining spin-up creation operators in $T_5$ and those must pair with the remaining creation operators in the definition of $|m^{(0)}\rangle$. There are exactly two ways of doing that (each of which gives a unique matching solution). In summary, we have seen that for a fixed value of $q,p,k$ there are two possible pairs of $(q',p')$ and one value of $k'$ so that $\langle n^{(0)}|T_5(p',q',k')|m^{(0)}\rangle\langle m^{(0)}|T_3(p,q,k)|n^{(0)}\rangle \neq 0$. Likewise

if we replace $T_5$ by $T_{15}$. If we ignore the $u$ and $v$ coherence factors, the net contribution of $T_5(p', q', k')$ is then $2\kappa^2$. Moreover, if we replace $T_3$ by $T_8$ (paired again with $T_5$ or $T_{15}$), there is another factor of $2\kappa^2$. Those exhaust all possibilities for $|n^{(0)}\rangle$. The next step in calculating the perturbative correction to $E_n$ is to sum over $p$ and $q$. What follows is an analysis similar, but more qualitative, to what was done in Section 3.4.7. Begin by examining $\sum_{p,q} \frac{\langle n^{(0)}|T_5(p',k',q')|m^{(0)}\rangle\langle m^{(0)}|T_3(p,k,q)|n^{(0)}\rangle}{E_n^{(0)}-E_m^{(0)}}$ where $T_i$ are defined in Eq. (20) and the variables $q, p', k', q', m^{(0)}$ are determined (as discussed above) – up to a finite number of configurations – by $p$ and $q$. One of the configurations for $|m^{(0)}\rangle$ has $\mathcal{D} = E_n^{(0)} - E_m^{(0)} = E_n - E_1 + E_2 + E_3$ where

$$E_1 = \sqrt{\Delta^2 + \epsilon_{p+q}'^2}$$
$$E_2 = \sqrt{\Delta^2 + \epsilon_p'^2}$$
$$E_3 = \sqrt{\Delta^2 + \epsilon_{r-q}'^2}$$
$$E_n = \sqrt{\Delta^2 + \epsilon_r'^2}$$

and $p, q, r$, are all 3-vectors. Let $(p_0, q_0)$ be a pair of vectors such that $\mathcal{D} = 0$. Imagine – as a starting point – that we can ignore $E_2$ and $E_3$. $E_n$ is determined by our choice of the state $|n^{(0)}\rangle$. We will begin by varying $p$ and $q$ away from $(p_0, q_0)$ so that $\epsilon_{p+q}'^2 = \epsilon_{p_0+q_0}'^2$. For all those values of $(p, q)$, we have $\mathcal{D} = 0$. Just as in Appendix A, we've carved out a hypersurface of constant energy and so just as in Appendix A, the number of states between adjacent hypersurfaces is $\frac{8\pi m^* \epsilon_F}{\delta^2}$. Note that there is a double sum which becomes a double-integral. Then change variables, so that $\sum_{p,q} \to g^2(\epsilon_F) \int dpdq \to g^2(\epsilon_F) \int d(p+q)dq = g^2(\epsilon_F) \int d\tilde{p}dq$. The denominator then becomes a function of $\tilde{p}$. For now, ignore the coherence factors in the numerator. The factor $g(\epsilon_F) \int dq$ becomes $\mathcal{N}$. Then, continuing to ignore the coherence factors, we could

approximate the sum in Eq. (21) as we did in Section 3.4.7,[15]

$$
\begin{aligned}
E_n^{(2)} \approx \mathcal{N} g(\epsilon_F) \int_{-\omega}^{\epsilon_r'-\delta'} d\epsilon' \frac{\kappa^4}{2\left(\sqrt{\epsilon_r'^2+\Delta^2}-\sqrt{\epsilon'^2+\Delta^2}\right)} + \\
\mathcal{N} g(\epsilon_F) \int_{\epsilon_r'+\delta'}^{\omega} d\epsilon' \frac{\kappa^4}{2\left(\sqrt{\epsilon_r'^2+\Delta^2}-\sqrt{\epsilon'^2+\Delta^2}\right)} \\
\approx \mathcal{N} g(\epsilon_F) \frac{\kappa^4}{2} \Bigg( -\frac{\sqrt{\epsilon_r'^2+\Delta^2}}{\epsilon_r'} \log \frac{(\omega-\epsilon_r')\left(\sqrt{(\epsilon_r'-\delta')^2+\Delta^2}-\epsilon_r'\right)}{(\omega+\epsilon_r')\left(\sqrt{(\epsilon_r'+\delta')^2+\Delta^2}-\epsilon_r'\right)} + \\
\log \frac{\left(\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'-\delta)^2+\Delta^2}+\epsilon_r'-\delta'\right)}{\left(-\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'+\delta')^2+\Delta^2}+\epsilon_r'+\delta'\right)} \Bigg)
\end{aligned}
\tag{22}
$$

In Section 3.4.7 we explored this result (but without the extra coefficient of $\mathcal{N}$) for four different values of $\epsilon_r'$. The first three were $\frac{\Delta}{2}, \frac{\omega}{4}, \delta'$. In each of those cases, the result was independent (to leading order) of $\delta'$. Moreover, using the results computed in Section 3.4.7, but with the extra factor of $\mathcal{N}$, we can see that the results are now all $\mathcal{O}(1)$ (instead of $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$. One might have expected a log divergence as $\mathcal{N} \to \infty$ since the integrand's denominator goes like $\frac{1}{\epsilon'-\epsilon_r'}$ as $\epsilon' \to \epsilon_r$. However, because the integrand changes sign around that potential singularity, the divergence doesn't appear. Mathematically, this is captured in the observation that in the limit of $\delta' \to 0$, the integral can be shown to be the Cauchy principal value. This observation carries forward to the case where we include the coherence factors since they are smooth around $\epsilon_r'$.

We also explored, in Section 3.4.7 the cases where $\epsilon_r' = \pm(\omega - \delta')$. There we found a $\delta'$-dependence and in the present case (with the extra factor of $\mathcal{N}$) our result is actually $\mathcal{O}(\mathcal{N})$ – which is unmanageably large! However, as pointed out in Section 3.4.7, when $\epsilon_r'$ is near the Debye boundary, Cooper pairs become unstable, so we shouldn't be surprised that perturbation theory would break down there. In fact, the $\log \mathcal{N}$ behavior is almost welcome, since we don't want the lowest-order BCS

[15] You might notice that if the upper bounds of $p$ and $q$ were both $\omega$ then the upper bounds of $|p+q|$ would not be $\omega$ (remember that $p$ and $q$ are both vectors). However, the Debye theory presumably is much more subtle about which momenta of the theory are limited by $\omega$ so it's really unclear how to decide what the upper bound is. I've chosen $\omega$. I brought up this issue earlier in Section 3.4.4.

theory to be valid once quasiparticles attain the Debye energy.[16]

We now return to the original denominator of Eq. (21), which was $\mathcal{D} = E_n^{(0)} - E_m^{(0)} = E_n - E_1 + E_2 + E_3$. Above, we ignored $E_2$ and $E_3$. What happens when we add those back in? Those terms can't be held constant while we change $p+q$ as we did in the previous paragraphs. We have to adjust the concepts of the previous paragraphs. Here's a qualitative approach. The space of variables $(p, q)$ is 6-dimensional and has a density function $\rho$ for converting from sum to integral. Roughly speaking, when there are no constraints, each dimension contributes a factor of $\mathcal{N}^{\frac{1}{3}}$ to the density $\rho$, so that $\int_{-p_F}^{p_F} d^3p d^3q g^2(\epsilon_F) = \mathcal{N}^2$. (In 6-dimensional space, we have $\rho = g^2(\epsilon_F)$.) In the previous example, we imagined a constraint where $E_1 =$ constant – which amounted to setting $|p+q|$ to a constant. The volume of equal-energy hyperplanes (of width $\delta$) was $\beta\mathcal{N}^{\frac{5}{3}}$ where $\beta$ is a constant. The remaining one-dimensional integral also has a density factor (a different one) proportional to $\mathcal{N}^{\frac{1}{3}}$. Altogether, the density per unit of momentum space is constant in a region of width $\mathcal{O}(\sqrt{\omega})$ around $\mathcal{S}_F \times \mathcal{S}_F$ where $\mathcal{S}_F$ is the Fermi surface.[17] and this leads to a total density proportional to $\mathcal{N}^2$. Any further dependence on $\mathcal{N}$ would have to come from the non-density part of the integrand, and as we saw when we pretended that $E_2 = E_3 = 0$, can only come about if there were an energy pole at $E_n^{(0)}$, arising from $E_1 = E_n$. That energy-pole was harmless (i.e., no linear dependence on $\mathcal{N}$) if the denominator at the pole crosses from positive to negative along the line of integration. The only case when the pole could cause a problem, is if it occurs near the edge of the Debye zone. When that happens, the line of integration stops at (or near) the pole, and so the denonimator doesn't change sign along the line of integration (the sign change is necessary in order for the large positive contribution to cancel the large negative contribution). Such

[16]Note that, in some sense, the log behavior is (presumably) a consequence of self-consistency rather than of fundamental physics. As far as the BCS goes, the Debye limit is an arbitrary constraint on the theory (coming from outside the BCS theory). So any breakdown of perturbation theory in the BCS theory, arises from self-consistency applied to the fact that the theory only holds for energies less than the Debye energy.

[17]In this discussion, I treat the Fermi surface as the surface of a sphere. However, for most solids, the Fermi surface is much more complicated. Even so, my qualitative discussion should apply since the results depend on the Debye window being much smaller than the average radius of the Fermi region. Moreover, the Debye window is somewhat complicated for a situation with two or more quasi-particle momenta where interaction terms might depend on, for example, $p+q$. The integration over $\epsilon'_{p+q}$ will cut off when $\epsilon'_{p+q} \leq \omega$.

a case leads to a $\log \mathcal{N}$ result, but as explained above, this is to be expected near the Debye-zone boundaries. The situation is the same with the full denominator that includes $E_1, E_2, E_3$; there is still only one constraint – but it is different. Even so, because there is only one constraint, most of the above discussion holds regarding hyperplanes, lines of integration and the effect of the pole. There is one minor point to consider when we have $E_1 + E_2 + E_3$. This total must be greater than $3\Delta$. But $E_1 < 3\Delta$ when $\epsilon_r'^2 < 8\Delta^2$. So for those values of $\epsilon_r'$, there is no combination of $(p, q)$ when $\mathcal{D} = 0$. In particular, there is no pole in the denominator. That makes no difference to the qualitative aspects of the end-result. We already saw that the pole contributes 0 to the integral.

The final thing to do in evaluating the perturbation of the one-quasiparticle energy, is to estimate the size of the corrections for the cases mentioned above where $\epsilon_r'$ is one of the values $\frac{\Delta}{2}, \frac{\omega}{4}, \delta'$. Use results computed in Section 3.4.7.

- Let $\epsilon_r' = \frac{\Delta}{2}$: Then $E^{(2)} \approx \Lambda^2\omega\left(\sqrt{5}\frac{\Delta}{\omega} + 2\log(2\frac{\omega}{\Delta})\right)$
  Substitute $\frac{\Delta}{\omega} = .02$ and $\Lambda = 0.2$ to get $E^{(2)} \approx 0.09\Delta + 0.37\omega$.The relative correction is obtained by dividing this result by $E_n^{(0)}$ which, in this case is $\frac{\sqrt{5}}{2}\Delta$. This leads to a relative correction of $0.08 + 17$ The second of these summands is way too large (the relative correction should be less than 1). Either this points to a failure of perturbation theory, or a failure to properly account for the coherence factors that we've been ignoring until now.

  Although this is where "the rubber meets the road" – a situation where perturbation theory might fail, even for a relatively small value of $E_n^{(0)}$, I'm not sufficiently motivated to do the hard work of tracking down – in a rigorous way – the effects of the coherence factors. However, I think I can paint a picture that's credible, although non-rigorous. Here are the steps:

  * The large value above comes from $\log \frac{2\omega}{2\Delta}$ which in turn comes from the integral in Eq. (22), $\log \frac{\left(\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'-\delta)^2+\Delta^2}+\epsilon_r'-\delta'\right)}{\left(-\omega+\sqrt{\omega^2+\Delta^2}\right)\left(\sqrt{(\epsilon_r'+\delta')^2+\Delta^2}+\epsilon_r'+\delta'\right)}$
  * The parameter $\omega$ is the upper bound (and $-\omega$ is the lower bound) of the integral in that equation. If we substitute $\Delta$ for $\omega$, then instead of $2\log\frac{2\omega}{\Delta} \approx 9.2$, we get $\log\frac{\left(1+\sqrt{2}\right)}{\left(-1+\sqrt{2}\right)} = 1.8$. If, for some reason, this substitution of $\Delta$ for $\omega$ were valid, that would reduce the relative correction to about 2.8 which

still isn't great.

* So let's see whether there's a good reason why we could cutoff the integral at $\pm\Delta$. Here's the crude argument. Pick any of the quasiparticle energies – for example $r - q$. Then use this energy as the one-dimensional parameter for integration. It intersects constant-energy hypersurfaces. However, it might not intersect them a "right angles". This statement is ambiguous, but I'm picturing something like the following: consider 3-space where the original energy integration is along the z-axis and crosses flat hypersurfaces parallel to the x-y plane at z=0 where the integrand is constant on each of those hypersurfaces. The length of this integration curve (along the $z$-axis is $2\omega$). Now change the integration curve to be along the line $z = x$. The curve no longer intersects the constant-energy hyperplanes at 90°, and it's a longer distance for the curve to go between the $z = -\omega$ and $z = \omega$ boundaries. On the other hand, an adjustment has to made to the measure along the integration curve so that the result comes out the same. This would all work out OK if the integrand is exactly 1 so that suggests we would be doing the right thing by adjusting the upper and lower limits, and then multiplying by an appropriately modified energy-density. The problem with this analysis, is that if we insist on an exact Debye limiting energy, then our various integration limits for both hypersurfaces and the energy curve, are somewhat more complicated.
* For simplicity, but without any rigor, I'm going to pick the energy curve defined by the integration parameter $\epsilon_{|r-q|}$ and I'll set the limits and energy density as I have done above (even though we know that's not right). I'll focus, instead, on the coherence parameters which we ignored. For the particular example we've been focused on, one of the terms includes the factor $u_{r-q}v_{r-q}$. That factor is bounded by $\frac{1}{2}$ and drops off quickly when $|r - q| > \Delta$. Thus the integrand is significantly suppressed beyond $\Delta$. So not only does this effectively change our upper bound from $\omega$ to $\Delta$ but it provides an extra factor of 2 suppression (from the $\frac{1}{2}$) everywhere. Arguably this brings down the relative correction to close to 1.
* Unfortunately, many terms of interest to us don't have factors $u_s v_s$ where $s$ is an integration parameter manifestly related to the energy-integration curve. Nevertheless, all factors include

at least two factors of $v_s$ and two factors of $u_{s'}$ as well as other factors of $u$ or $v$. I believe that some combination of these suppresses integration contributions where the integration parameter is much larger than $\Delta$. However, I don't have even the sketch of a proof ... just a couple of sketchy examples.

So, we've seen that the coherence factors suppress the integrand enough to conceivably reduce the relative correction to something 'perturbative'. There are other coherence factors that weren't discussed and which might also turn out to give additional suppression. That would be helpful (for reducing the size of the correction) since up to now, I've only been focusing on the term $\sum_{p,q} \frac{\langle n^{(0)}|T_5(p',k',q')|m^{(0)}\rangle\langle m^{(0)}|T_3(p,k,q)|n^{(0)}\rangle}{E_n^{(0)}-E_m^{(0)}}$. There are a few other terms that need to be added for the calculation of the corrections – owing to 3 intermediate quasiparticles – to one-quasiparticle energies.

- Recall that we were looking at three different sample energies for $E_n^{(0)}$. Now look at $\epsilon'_r = \frac{\omega}{4}$ and assume (from the argument above) that the energy integrals cutoff at $\pm\Delta$. This leads to $E_n^{(2)} \approx 0.09\omega$. We have $E^{(0)} \approx \frac{\omega}{4}$. Then the relative correction is $\frac{E^{(2)}}{E^{(0)}} \approx 0.4$. This is reasonable.
- Finally look at $\epsilon'_r = \delta'$. Again, cut off the energy integral at $\pm\Delta$. The analysis of Eq. (22) is a bit misleading in this case. Notice the term $-\frac{\sqrt{\epsilon_r'^2+\Delta^2}}{\epsilon'_r}\log\frac{(\omega-\epsilon'_r)\left(\sqrt{(\epsilon'_r-\delta')^2+\Delta^2}-\epsilon'_r\right)}{(\omega+\epsilon'_r)\left(\sqrt{(\epsilon'_r+\delta')^2+\Delta^2}-\epsilon'_r\right)}$ and focus on the factor $\frac{\sqrt{\epsilon_r'^2+\Delta^2}}{\epsilon'_r}$. If we replace $\epsilon'_f$ by $\delta'$ and $\omega$ by $\Delta$, then Taylor-expand, and finally take the ratio $\frac{E^{(2)}}{E^{(0)}}$, the result is something $\mathcal{O}(4)$. Is there some suppression factor that we're missing? Perhaps. Eq. (22) was obtained by ignoring the terms $E_2$ and $E_3$ in the denominator. Since both those terms are greater than $\Delta$, they will play some role in further suppression. However, I can't come up with an easy estimate of the magnitude of that suppression.

Overall, our results suggest that the second-order shift of energies might be large enough – at least for the less energetic quasiparticles – to invalidate perturbation theory. Clearly, the details matter and a more systematic approach would be required. I expect that a Feynman-diagram analysis could deal in a more approachable way with the various integrals encountered. Another approach – if nothing else were available – would be purely numerical.

One important thing we've exposed, is the $\mathcal{N}$-dependence. In Section 3.4.7, where we had not yet introduced the momentum $q$, the perturbative relative corrections were typically $\mathcal{O}\left(\frac{1}{\mathcal{N}}\right)$ or $\mathcal{O}\left(\frac{\log \mathcal{N}}{\mathcal{N}}\right)$. However, in this section we find the relative corrections of $\mathcal{O}(1)$, which makes more sense. Unfortunately, the effective coupling $\Lambda$ is typically not small, and several results require an overall factor of $\frac{\omega}{\Delta}$ which is large. Depending on suppression factors – some of which I have not run to the ground – the final relative corrections may or may not be acceptably small.

So far, I've focused on $T_i$ which have one annihilation operator. But as in Section 3.4.7, we also need to consider $T_4$ which has no annihilation operators. We deal with this more or less how we dealt with the ground state. None of the operators are annihilation operators, so none of them match up to $b_r^\dagger$. Instead, we end up with 5 creation operators including $b_r^\dagger$. Then $\langle m^0| = \langle BCS|b_{-k-q\downarrow}\, b_{p+q\uparrow}\, b_{-p\downarrow}\, b_{k\uparrow} b_{r\uparrow}$. Just as for the ground state, we then need to match $\langle n^{(0)}|T_{10}(p'k'q')|m^{(0)}\rangle$ with $\langle m^{(0)}|$ and the result is just as we had for the ground state **except** that the triple sum needs to exclude $k = r$ and $p + q = r$, since in those 2 cases we would have the square of a creation operator which would give 0. Once we calculate $E_R^{(2)}$ by subtracting the second-order ground state perturbation $E_{gs}^{(2)}$, we end up with a contribution from the 5-quasiparticle intermediate states of $\kappa^4\Delta^4 \sum_{pq} \frac{2}{(E_p+E_{p+q}+E_{r+q})E_pE_{p+q}E_rE_{r+q}}$. We can evaluate this with the usual integration approximation. The double sum becomes a double integral multiplied by $g^2(\epsilon_F)$. This is multiplied with $\kappa^4$, giving $\Lambda^2$. The denominator of the integrand is greater than $3\Delta^5$. However, that analysis is insufficient because if the denominator simply were a constant times $3\Delta^5$, then after integration we would end up with $\Lambda^2 \frac{4\omega^2}{3\Delta^2} = \mathcal{O}(500)$ which is a very large number. I've struggled with this, with partial success. Here's an argument for why the actual integration leads to a reasonable perturbative result. First suppose that $|p_r - p_F|$ is small so that we can set our sum to be $\sum_{pq} \frac{2}{(E_p+E_{p+q}+E_q)E_pE_{p+q}\Delta E_q}$. Next suppose (and this is a stretch!) that the momenta are only one dimensional, the expressions $E_x$ can all be replaced by $x$, , that $p+q \to p-q$ (so that we can explore the most problematic situation) and that the sum becomes an integral $g^2(\epsilon_F) \int_\Delta^\omega dqdp\mathcal{I}(q,p)$ where $\mathcal{I}$ is the integrand. Shortly I'll demonstrate that the integral converges so that as $\omega \to \infty$, the integral approaches a finite number (independent of $\omega$). Since this integral has energy-dimensions $[E]^{-3}$ and since the only energy unit (besides $\omega$ which the integral doesn't depend on) is $\Delta$, this implies the integral is $\mathcal{O}\left(\Delta^{-3}\right)$. Continuing, we multiply $\kappa^4$ by $g(\epsilon_F)^2$ to obtain $\Lambda^2$,

and multiply the above integral by $\Delta^4$ to obtain $E_R^{(2)} = \mathcal{O}(\Lambda^2\Delta)$. Since we're ultimately interested in the ratio of the perturbative term to the $0^{th}$ order term $E_n^{(0)}$, which is greater than $\Delta$, our final conclusion would be (assuming that we can prove convergence of the integral) a relative perturbative effect of order $\Lambda^2$. So now, let's prove convergence. We do this by explicit integration of $\int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q(p-q)}$. The integrand may look rather different than what we started with, but I think it serves as a reasonable model. In the original integrand (after setting $r = 0$), the denominator included a factor $E_p + E_{p+q} + E_q$. The original integration bounds were from $-\omega$ to $\omega$, and the original energies had lower bounds of $\Delta$, and approximated $\epsilon_k$ for large values of $\epsilon_k$ where we take $k$ to be a generic momentum ($q$ or $p$ for example). The potentially problematic convergence would occur when a large negative value of $q$ is added to an equal-magnitude positive value of $p$. This would reduce the large-momentum behavior of the denominator to be of order $k^3$ and since the integral is 2-dimensional, we would expect a logarithmic divergence. So by looking only at positive values of $q$ and $p$, and cutting off the lower bounds at $\Delta$, and replacing $q$ by $-q$, we have a chance to explore whether the potential divergence is actually a problem. Write the integrand as $\frac{1}{p^2q(p-q)} = \frac{1}{p^2q^2} + \left(\frac{1}{p^2q(p-q)} - \frac{1}{p^2q^2}\right) = \frac{1}{p^2q^2} + \frac{2q-p}{p^2q^2(p-q)} = \frac{1}{p^2q^2} + \frac{2}{p^2q(p-q)} - \frac{1}{pq^2(p-q)}$. Now for a trick. When we integrate, we can exchange $p$ and $q$ provided that we remember to multiply by $-1$ since $p - q \to q - p$. In the equation we just derived, the last two summands – if we exchange $q$ and $p$ – become $\frac{3}{p^2q(p-q)}$. But this is just three times the original integrand. That is, we've shown that $\int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q(p-q)} = 3\int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q(p-q)} + \int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q^2}$ or $\int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q(p-q)} = -\frac{1}{2}\int_\Delta^\omega dp \int_\Delta^\omega dq \frac{1}{p^2q^2}$. The RHS can be factored and is easy to integrate. We get $\left(\frac{1}{\omega} - \frac{1}{\Delta}\right)^2$. As $\omega \to \infty$, that becomes $\frac{1}{\Delta^2}$.

Remember that all this was done by setting $r = 0$, as well as making other approximations. I have not been able to figure out any easy tricks for estimating what happens when $r$ is largish (say $> \Delta$). I think one could show fairly easily that the integral converges, but now there are two length scales, $\Delta$ and $r$. So naive dimensional analysis isn't good enough. I suspect that the results will continue to be reasonable, but that will take more work. What about the other approximations, in particular the one-dimensional approximation? Earlier I discussed similar issues in comparing the 1D energy integration to the 3D momentum integration. I think the same kind of considerations apply, where

we look at hypersurfaces of equal energies (or perhaps hypersurfaces of co-dimension 2 which are 'orthogonal' to two 'energy' dimensions). Anyway, that's a lot of work, and there's bound to be better ways to organize all this. It's reasonable to assume that all of this is part of the classical lore in the subject, but most likely done using Feynman diagram formalism (where lots and lots of integration tricks were developed in high energy physics and presumably also in condensed matter physics).

**Here's a brief summary of the highlights of this investigation. When we include $q$-dependence in our Hamiltonian, we go from double sums used in the BCS mean-field theory, to triple sums. Since $\kappa^2$ scales as $\frac{1}{\mathcal{N}}$ and the triple sums might be expected to scale as $\mathcal{N}^3$, one might guess that our results would diverge with $\mathcal{N}$. Indeed, when looking at the ground-state energy in second-order perturbation theory, that grows as $\mathcal{N}$. However, what matters are the energy differences, so once we adjust all energies by subtracting the above-mentioned ground-state energy perturbation, we are left with double-sums. Those have no dominant $\mathcal{N}$ factors but typically they can involve the large ratios of $\omega$ to $\Delta$. Often-times, the integrands can be seen to be suppressed by coherence factors when the integration variable exceeds $\Delta$. In those cases, the factors of $\omega$ don't appear, and perturbative results are $\mathcal{O}(\Lambda^2)$. However, the integrals can be messy and it's not always clear whether there really is this kind of suppression. I've described a few situations of that kind, but am guessing that in the literature, things have been developed much further and probably with satisfactory outcomes.**

I will next explore the second-order perturbations of 2-quasiparticle states including Cooper pairs, and then finally the second-order perturbation of M-quasiparticle states. In Section 3.4.7, we found that the $M$-quasiparticle states had relative corrections of $\mathcal{O}(1)$ rather than $\mathcal{O}(\frac{1}{\mathcal{N}})$, so it will be important to see if that remains the same in this section.

- In this example I'll take our unperturbed state to be the 0-momentum Cooper pair $|n^{(0)}\rangle = b^\dagger_{r\uparrow} b^\dagger_{-r\downarrow}|BCS\rangle$. The analysis is almost identical to that of Section 3.4.7 for Cooper pair states. We look for $T_i$ with two annihilation operators to pair off with the creation operators $b^\dagger_{r\uparrow}, b^\dagger_{-r\downarrow}$. Moreover, since we'll also be evaluating $\langle n^{(0)}|T_i = \langle BCS|b_{-r\downarrow} b_{r\uparrow} T_i$, then

$T_i$ also needs to have two creation operators to pair off with the annihilation operators $b_{-r\downarrow}, b_{r\uparrow}$. The only viable options are $T_1$,$T_6$, $T_{11}$ and $T_{16}$. In Section 3.4.7 we were able to argue that $T_6$ and $T_{11}$ did not contribute, but that isn't the case here. I'll show the matching equations for $\langle n^{(0)}|T_1(p',k',q')|m^{(0)}\rangle\langle m^{(0)}|T_1(p,k,q)|n^{(0)}\rangle$ and will just state results for the other viable combinations. First match the annihilation operators of $T_1(p,k,q)$ to $|n^{(0)}\rangle$. $k+q=r, p+q=r$ so $k=p=r-q$. Then match $|m^{(0)}\rangle$ to the creation operators, leading to $|m^{(0)}\rangle = b^\dagger_{r-q\uparrow}b^\dagger_{-r+q\downarrow}|BCS\rangle$. Next match the creation operators of $T_1(p',k',q')$ to $\langle n^{(0)}\rangle$, leading to $k=r, p=r$. Finally, substitute the values for $k'$ and $p'$ and match the annihilation operators of $T_1(p',k',q')$ to $|m^{(0)}\rangle$. $p'+q'=r+q'=r-q$ and $k'+q'=r+q'=r-q$ from which $q'=-q$. The only free variable is $q$, so we will sum over $q$. Moreover, $(u_1u_2u_3u_4)^2 = u_r^4u_{r-q}^2$. The denominator is $\mathcal{D} = 2\left(E_{r-q}-E_r\right)$. In all, there are 16 pairs of $T_iT_j$ where $i,j = 1,6,11,16$. If we assume[18] that the sums over $r-q, r+q, k$ and $p$ all range from $-\omega$ to $\omega$, then if we work out all the matching and coherence factors, we obtain

$$E_n^{(2)} = \frac{\kappa^2}{2}\sum_{p\neq r}\frac{\left(u_p^2-v_r^2\right)^2}{\left(E_p^{(0)}-E_r^{(0)}\right)}$$

Apart from the coherence factors, this is the same result as we obtained for Cooper-pair second-order correction in Section 3.4.7. In that section, we saw that all corrections (other than those pertaining to pairs whose individual-quasiparticle energies were close to the Debye limit) are suppressed by a factor of $\mathcal{N}$. So these perturbations are well-controlled. If we generalize $|E^{(0)}\rangle$ to be $b^\dagger_{r\uparrow}b^\dagger_{-r'\downarrow}|BCS\rangle$, which is a Cooper-pair of general momentum (usually nonzero), we get the same qualitative behavior.

So far, I've said nothing about the $T_i$ with either one or no annihilation operators to match $|n^{(0)}\rangle$. By now, this should be a familiar story. First look at the $T_i$'s with only one annihilation operator. Those are $T_2$, $T_3$, $T_8$ and $T_{12}$. Consider $T_2 = -\,u_1u_2u_3v_4\; b^\dagger_{k\uparrow}\,b^\dagger_{-r+q\downarrow}\,b^\dagger_{-k-q\downarrow}\,b_{-r\downarrow}$. The annihilation operator will pair off with $b^\dagger_{-r\downarrow}$, leaving 4 creation operators, $b^\dagger_{k\uparrow}\,b^\dagger_{-r+q\downarrow}\,b^\dagger_{-k-q\downarrow}\,b^\dagger_{r\uparrow}$. Those match against the intermediate state $\langle m^{(0)}|b_{r\uparrow}\,b_{-r\downarrow}\,b_{-k\downarrow}\,b_{k\uparrow}$. Then, in order to match $|m^{(0)}\rangle$, we'll take $T_9$ in $\langle n^{(0)}|T_9|m^{(0)}|$. The result will be a double sum over $k,p$ and an overall factor of $\kappa^4$. Similar results are obtained with the combination of

[18]This assumption isn't necessarily justifiable although it seems reasonable

$T_3$ and $T_5$ as well as the other cases involving $T_8$ and $T_{12}$. This analysis is the most complex I've encountered so far, for two reasons. First of all, the coherence factors don't have the simple structure that we've sometimes encountered, where all factors are of the form $u_s v_s$, which becomes $\frac{\Delta}{E_s}$. Moreover, the denominator of Eq. (21) involves a relative minus sign $(-E_r^{(0)} + E_k^{(0)} + ...)$. We've encountered this kind of thing earlier in our analysis of the perturbations to Cooper pairs. In principle, there could be large contributions of $\mathcal{O}\left(\frac{\omega}{\Delta}\right)$ or even $\mathcal{O}\left(\frac{\omega^2}{\Delta^2}\right)$. With luck, all the coherence factors etc. will provide sufficient suppression, but I'm not going to tackle that.

As for $T_i$ with no annihilation operators, that analysis is similar to what we've done before comparing those terms to the terms that perturb the ground state and we expect the contributions to be well-behaved (although as mentioned in the previous section on single-quasiparticle second-order perturbations, that analysis is also not complete).

- Our last example, following Section 3.4.7, will be the one where the unperturbed state $|n^{(0)}\rangle$ consists of $M$ quasiparticles with spin up, and $M$ with opposite spin and opposite (i.e., the negative) momenta. We believe this example will be analyzed as a synthesis of the two-quasiparticle example, and the analysis of $M$-quasiparticle second-order perturbations in Section 3.4.7.

# A State density and spacing

In this section we carefully consider the sum over intermediate states, that appears in Eq. (15). We've simplified our notation to only index states by a 1-D momentum, e.g. $p$. However, we have intended all along that $p$ is actually a 3-D vector $\mathbf{p}$. This matters when taking sums over intermediate states. Whenever we speak of the residual potential action on the intermediate state, we speak of annihilation (or creation) operators matching the values of $k$ in the intermediate state. But what we really mean is that we are summing over all intermediate momenta with a fixed value of $|\mathbf{k}|$, with the assumption (from symmetry) that matrix elements are independent of the direction of $\mathbf{k}$ and that $E_k^{(0)}$ appearing in the denominator of Eq. (15) is the same for all directions.

We will approximate the discrete grid with a continuum and convert sums

to integrals. That requires a bit of care and also requires a dispersion relation between energy and momentum. In the Bloch theory, the dispersion relation[19] is approximately $\epsilon' \approx \frac{1}{m^*}\mathbf{p}_F \cdot (\mathbf{p} - \mathbf{p}_F)$, for momenta $\mathbf{p}$ very close to a momentum vector on the Fermi surface, which we will designate $\mathbf{p}_F$. Strictly speaking, this is a set of dispersion relations dependent on the choice of $\mathbf{p}_F$. However, to first-order , this is independent of that choice. We'll set up our continuum-approximation by separating the discrete momentum-grid into surfaces of approximately[20] constant energy. Then, if we pick a momentum $\mathbf{p}$ on one of those surfaces, we'll compute $\epsilon_{\mathbf{p}'}$ by selecting a $\mathbf{p}_F$ which is parallel to $\mathbf{p}$, in which case the dispersion relation gives us $\epsilon' \approx \frac{p_F}{m^*}(|\mathbf{p}| - p_F)$. If, in the discrete case, the average $p$-distance between adjacent momenta is $\delta$, then the average $p$-distance between adjacent energy-hypersurfaces is also $\delta$.

From the dispersion relationship, the average adjacent energy-difference is $\delta\frac{p_F}{m^*} = \delta\frac{\sqrt{2\epsilon_F}}{\sqrt{m^*}}$. The average number of states between energy-levels (i.e. between constant-energy hypersurfaces) is $4\pi\frac{p^2}{\delta^2} \approx 4\pi\frac{p_F^2}{\delta^2} = \frac{8\pi m^*\epsilon_F}{\delta^2}$. Recall that energies are all within the Debye window $-\omega < \epsilon' < \omega$ so the number of energy hypersurfaces is $\approx \frac{2\omega}{\delta\frac{\sqrt{2\epsilon_F}}{\sqrt{m^*}}}$. If we multiply that number by the number of states between energy-levels, we obtain he total number of states (i.e. particles) in the system, $\mathcal{N}$. From this, we see that the density of states per unit energy is $g(\epsilon_F) \approx \frac{4\pi\sqrt{2\epsilon_F m^{*3}}}{\delta^3}$. Also, for future reference, $\delta = \frac{\left(128\pi^2\epsilon_F m^{*3}\omega^2\right)^{\frac{1}{6}}}{\mathcal{N}^{\frac{1}{3}}}$.

[19]See notes on the semiclassical approximation for metals.

[20]some mathematical rigor must be applied to get a firm grasp of the meaning of 'approximately'.