Contributions to the mathematical study of BCS theory

Andreas Deuchert

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisches Institut
2016

Contributions to the mathematical study of BCS theory

DISSERTATION

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Andreas Deuchert

aus Lich

Tübingen

2016

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	29.09.2016
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Christian Hainzl
2. Berichterstatter:	Prof. Dr. Stefan Teufel

Contents

Acknowledgements

First of all I would like to express my deep gratitude to my supervisor Christian Hainzl for the opportunity to write this thesis and for the constant support he gave me, both professionally and personally. He has a unique point of view towards the mathematics of BCS theory and I have had the fortune to share in his ideas and to profit from his great enthusiasm for research in this field.
I am indebted to Robert Seiringer for the opportunity to spend a period of ten weeks in 2014 at the IST Austria. It was a pleasure to learn from him and to work with him and I am looking forward to future collaborations. It is likewise a pleasure to be able to thank Michael Loss and Alissa Geisinger for the fruitful collaboration that led to the results of Chapter 4.
Special thanks go to Stefan Teufel for his willingness to act as a referee, as well as to Marcel Griesemer and Ruben Jakob for being members of my thesis committee.
Financial support by the DFG through the Graduiertenkolleg 1838 is gratefully acknowledged.
The publication of the results of Chapter 2 is available at Springer via
`http://dx.doi.org/10.1007/s11005-015-0787-5`.
I am truly grateful to all the members of the mathematical physics group in Tuebingen for all the inspiring and stimulating discussions and for the beautiful time we spent together. You know who you are! Special thanks go to Tim Tzaneteas for proofreading large parts of this manuscript, to Wolfgang Gaim for proofreading Chapter 4 and to Mario Laux and Stefan Haag for their support with Latex.
Finally, and most importantly, I have to acknowledge the support and the love of my family and my friends, especially my wife Sabrina, my mother Dagmar and my father Klaus. I could not have done it without you.

Personal Contributions

All results, that is, those of Chapter 2, Chapter 3 and Chapter 4 have been obtained under the supervision of Prof. Dr. Christian Hainzl. The work of Chapter 2 is due to a collaboration with Prof. Dr. Robert Seiringer. The author, however, was responsible for the details of the proofs as well as for the manuscript. In the case of Chapter 4, whose results were obtained in collaboration with M.Sc. Alissa Geisinger and Prof. Dr. Michael Loss, these responsibilities were shared by the author and M.Sc. Alissa Geisinger in equal parts.

Abstract

The results of this thesis contribute to the mathematical study of BCS theory, that is, to the study of the BCS gap equation and the BCS functional. In the first part, we investigate a recent definition of a generalized relative entropy for bounded and not necessarily compact operators, which, in the second part, is used to define and study a non-periodic version of the BCS functional with an external field. In the third part of this work, we consider the BCS functional in two spatial dimensions with a radial pair interaction and show that the translational symmetry of the model is not broken.

The quantum relative entropy plays an important role in statistical mechanics and quantum information theory. Apart from its relevance in quantum physics, it has interesting mathematical properties, such as joint convexity and monotonicity, see e.g. [39]. From a purely mathematical point of view, it can be interpreted as a distance measure between two matrices, or more generally two trace-class operators, A and B. This is because it is positive and it equals zero if and only if $A = B$. In a recent work, Lewin and Sabin [36] considered a family of generalized relative entropies defined for arbitrary bounded and not necessarily compact operators that captures some of the strong properties of the physical relative entropy. Their motivation for introducing this quantity was the study of the Hartree equation for infinitely many particles, and in particular the well-posedness and scattering properties of the equation. Like Hartree states, BCS states are conveniently described in terms of a generalized one-particle density matrix and so it is no surprise that the need to extend the relative entropy to non-compact operators also plays a role in BCS theory. This was our initial motivation in studying this object. The generalized relative entropy of Lewin and Sabin is defined by a limiting procedure that resembles the thermodynamic limit. An important question left open in their work is whether there exists a simple formula that allows one to compute the limit without having to work with the complicated limiting procedure. Such a formula is necessary for example for trial state arguments. In Chapter 2 we answer this question affirmatively and derive such a formula. Assuming some mild regularity conditions for the operators A and B, we show that it takes a particularly simple form, which is suitable for computations. Our proof is based upon a novel integral representation of the generalized relative entropy that is of interest in itself.

Soon after the *microscopic* BCS theory of superconductivity was proposed in 1957 [1], Gorkov realized in which sense it was related to the previously known *macroscopic* Ginzburg-Landau theory of superconductivity [27]. More precisely, Gorkov could

show that close to the critical temperature, the BCS gap equation can be expanded in powers of the gap function Δ, which, to leading order, formally solves a related Ginzburg-Landau equation. In 2012, Frank, Hainzl, Seiringer and Solovej gave the first rigorous proof of this connection [28]. While Gorkov analyzes the relation between the BCS gap equation and the Ginzburg-Landau equation, the proof in [28] is based upon the fact that both equations can be realized as Euler-Lagrange equations for variational problems. The authors could show that in a certain scaling limit, the difference of the free energy of the superconducting state and that of the normal state is to leading order given by the infimum of a suitably chosen Ginzburg-Landau functional. Additionally, minimizers of the BCS functional can be related to the minimizer of this Ginzburg-Landau functional. Frank, Hainzl, Seiringer and Solovej consider in their work a periodic version of the BCS functional where the period is assumed to be large. This implies in particluar that the external fields are assumed to be periodic which may be seen as a disadvantage of the model. Motivated by this, we introduce in Chapter 3 a non-periodic version of the BCS functional, modeling the situation of infinitely many particles, filling all of $\mathbb{R}^3$, that are subject to an adequately localized external potential. We formally derive this BCS functional by taking the difference of the infinite free energy of a state Γ and the infinite free energy of a reference state Γ_0. We regroup the resulting terms and arrive at an energy functional which is well defined. Having this BCS functional at hand, we study it in the scaling limit considered in [28] and derive a lower bound as well as a-priori estimates for states with energy smaller or equal than that of the reference state. While the construction of a lower bound for the energy functional considered in [28] is relatively simple, our functional has a different mathematical structure and proving the existence of a lower bound is already a challanging task. The a-priori estimates for states with energy smaller or equal than the one of the reference state are the first step towards an extension of the derivation of Ginzburg-Landau theory to the set-up considered in this work and should be compared with the a-priori estimates for states with energy smaller or equal than the one of the normal state derived in [28]. Our strategy of proof is motivated by the one used in [28] to derive these a-priori estimates.

In Chapter 4 we consider the two-dimensional BCS functional with a radial pair interaction. We show that the translational symmetry of the system is not broken for temperatures in a certain temperature interval below the critical temperature. In case of vanishing angular momentum, our result carries over to the three-dimensional model. Prior to this work, such a result was only known in the case of $\hat{V} \leq 0$ and not identically zero, see [31]. In the latter situation, the minimizer of the BCS functional is unique, up to a phase in front of the Cooper-pair wave function. Apart from these considerations, we show for the two-dimensional model that Cooper-pairs are in a definite angular momentum state if the temperature lies in the temperature interval mentioned above. A similar result in the three-dimensional case allows us to determine a situation, in which Cooper-pairs are in an s-wave state. The fact that the translational symmetry is not broken in certain situations is of considerable interest, since it allows one to significantly reduce the complexity of the BCS model.

Zusammenfassung

Die Ergebnisse dieser Doktorarbeit liefern einen Beitrag zum mathematischen Studium der BCS Theorie, d.h. zum Studium der BCS Gapgleichung und des BCS Funktionals. Im ersten Teil untersuchen wir eine, vor kurzem eingeführte, verallgemeinerte relative Entropie, die dann im zweiten Teil der Arbeit benutzt wird, um eine nicht periodische Version des BCS Funktionals mit einem äußeren Potential zu definieren und zu untersuchen. Im dritten Teil der Arbeit behandeln wir das BCS Funktional in zwei Raumdimensionen mit einem radialen Wechselwirkungspotential und zeigen, dass die Translationssymmetrie des Systems nicht gebrochen ist.

Die relative Entropie spielt eine wichtige Rolle in der statistischen Physik quantenmechanischer Systeme und in der Quanteninformationstheorie. Darüber hinaus hat sie interessante mathematischen Eigenschaften, wie zum Beispiel die gleichzeitige Konvexität in beiden Argumenten oder ihre Monotonie, siehe z.B. [39]. Von einem rein mathematischen Standpunkt aus betrachtet kann man sie als ein Maß dafür ansehen, wie sich zwei Matrizen, oder allgemeiner zwei Spurklasseoperatoren, A und B voneinander unterscheiden. Der Grund dafür ist die Tatsache, dass die relative Entropie stets positiv ist und genau dann Null wird wenn $A = B$ gilt. Vor kurzem haben Lewin und Sabin in [36] eine Familie verallgemeinerter relativer Entropien für allgemeine beschränkte und nicht notwendigerweise kompakte Operatoren eingeführt, die einige starke Eigenschaften mit der physikalischen relativen Entropie teilt. Ihre Motivation für diese Definition lag im Studium der Hartree Gleichung für unendlich viele Teilchen, wobei insbesondere die Wohlgestelltheit des Anfangswertproblems und die Streueigenschaften der Gleichung untersucht wurden. Da sowohl BCS Zustände wie auch Hartree Zustände typischerweise durch eine verallgemeinerte Einteilchendichtematrix beschrieben werden, ist es naheliegend, dass die relative Entropie für nicht kompakte beschränkte Operatoren auch in der BCS Theorie von Relevanz ist. Dieser Zusammenhang stellt für uns die Hauptmotivation dar, dieses Objekt näher zu untersuchen. Die verallgemeinerte relative Entropie von Lewin und Sabin ist durch eine Limesprozedur definiert, die an einen thermodynamischen Limes erinnert. Eine interessante Frage, die in ihrer Arbeit offen gelassen wurde, ist ob, und wenn ja wie, es möglich ist den zugehörigen Grenzwert mittels einer einfachen Formel direkt zu berechnen, d.h. ohne auf die komplizierte Limesprozedur zurückgreifen zu müssen. Ein solches Ergebnis ist unter anderem wichtig, um Argumente mit explizit konstruierten Testzuständen durchführen zu können. In Kapitel 2 beantworten wir diese Frage positiv und geben eine entsprechende Formel an. Unter sehr schwachen Regularitätsannahmen an die Operatoren A und B zeigen wir zudem, dass diese Formel eine, für explizite

Berechnungen besonders wünschenswerte, Form annimmt. Unser Beweis basiert auf einer neuen Integraldarstellung der verallgemeinerten relativen Entropie, die für sich selbst genommen von Interesse ist.

Bereits kurz nachdem die *mikroskopische* BCS Theorie der Supraleitung im Jahre 1957 vorgeschlagen wurde [1], erkannte Gokov ihren Zusammenhang mit der, schon vorher bekannten, *makroskopischen* Ginzburg-Landau Theorie der Supraleitung [27]. Er stellte fest, dass sich die BCS Gapgleichung für T nahe T_c in Ordnungen der Gapfunktion Δ entwickeln läßt und konnte so formal zeigen, dass Δ in führender Ordnung eine Ginzburg-Landau Gleichung löst. In 2012 gelang es Frank, Hainzl, Seiringer und Solovej in [28] einen ersten rigorosen Beweis dieses Zusammenhangs zu geben. Während Gorkov in seiner Arbeit den Zusammenhang zwischen der Gapgleichung und der Ginzburg-Landau Gleichung untersucht, basiert der Beweis in [28] auf der Tatsache, dass sich beide Gleichungen als Euler-Lagrange Gleichungen von Variationsproblemen ergeben. Die Autoren konnten zeigen, dass in einem gewissen Skalierungslimes die Differenz der freie Energie des supraleitenden Zustandes und die des normalleitenden Zustandes, in führender Ordnung durch das Minimum eines geeignet gewählten Ginzburg-Landau Funktionals gegeben ist. Zusätzlich lassen sich Minimierer des BCS Funktionals mit dem Minimierer dieses Ginzburg-Landau Funktionals in Verbindung bringen. Frank, Hainzl, Seiringer und Solovej betrachten in ihrer Arbeit eine periodische Version des BCS Funktionals, wobei die Periodenlänge groß gewählt wird. Das impliziert insbesondere, dass die äußeren Felder als periodisch angenommen werden, was als Nachteil des Modelles gewertet werden kann. Durch diesen Sachverhalt motiviert, führen wir in Kapitel 3 ein BCS Funktional ein, welches unendlich viele Teilchen im $\mathbb{R}^3$ beschreibt, die sich unter dem Einfluß eines hinreichend lokalisierten äußeren Feldes befinden. Wir leiten dieses BCS Funktional formal her, indem wir die unendliche freie Energie eines Zustandes Γ von der ebenfalls unendlichen freien Energie eines Referenzzustandes Γ_0 abziehen. Nach Umordnung der Terme erhalten wir ein wohldefiniertes Funktional. Das so erhaltene BCS Funktional betrachten wir in dem Skalierungslimes, der in [28] benutzt wurde um Ginzburg-Landau Theorie herzuleiten. Unsere Hauptresultate sind die Konstruktion einer unteren Schranke und die Herleitung von a-priori Abschätzungen für Zustände mit Energie kleiner gleich der des Refrenzzustandes. Die Konstruktion einer unteren Schranke für das, in [28] betrachtete, periodische BCS Funktional ist relativ einfach. Unser Funktional dagegen hat eine andere mathematische Struktur und bereits die Konstruktion einer unteren Schranke ist ein schwieriges Problem. Die a-priori Abschätzungen für Zustände mit Energie kleiner gleich der des Referenzzustandes sollten mit den a-priori Abschätzungen in [28] für Zustände mit Energie kleiner gleich der des Normalzustandes verglichen werden. In diesem Sinne sind sie der erste Schritt einer Herleitung der Ginzburg-Landau Theorie in unserem Setup. Unsere Beweisstrategie ist motiviert durch die Strategie, welche in [28] benutzt wurde um die oben genannten a-priori Abschätzungen zu beweisen.

In Kapitel 4 betrachten wir das BCS Funktional in zwei Raumdimensionen mit einer radialen Paarwechselwirkung. Für Temperaturen in einem gewissen Temperaturintervall unterhalb der kritischen Temperatur zeigen wir, dass die Translationssymmetrie des Systems nicht gebrochen ist. Im Fall des dreidimensionalen Modells gilt dieses Resultat weiterhin wenn die Cooperpaare verschwindenden Drehimpuls haben. Bisher war eine solche Aussage nur für nicht verschwindende Paarwechselwirkungen mit $\hat{V} \leq 0$ bekannt. In dieser Situation ist ein Minimierer des BCS Funktionales, bis auf eine komplexe Phase, eindeutig. Neben diesen Resultaten zeigen wir für das zweidimensional Modell, dass Cooperpaare einen festen Drehimpuls haben wenn die Temperatur in dem, oben erwähnten, Temperaturintervall gewählt wird. Eine ähnliche Betrachtung in drei Raumdimensionen erlaubt es uns eine Situation zu charakterisieren, in der Cooperpaare in einem s-Wellenzustand sind. Ergebnisse, die zeigen, dass die Translationssymmetrie in gewissen Situationen nicht gebrochen ist, sind von besonderem Interesse, da sie eine erhebliche Reduktion der Komplexität des BCS Modells erlauben.

CHAPTER 1

Introduction

In the introduction we first give an overview over the literature on the mathematical properties of the BCS gap equation and the BCS functional. Results that are relevant for this work are discussed to some extend, and therefore this section also serves as a short introduction to mathematical BCS theory. Afterwards, we summarize the main results of this thesis.

1.1 The mathematical literature on the BCS gap equation and the BCS functional

After the phenomenon of superconductivity was discovered by Heike Kamerlingh Onnes in 1911, it took almost 46 years of intensive research before the microscopic origin of this remarkable effect could be explained. In 1957 Bardeen, Cooper and Schrieffer (BCS) published their famous paper with the title "Theory of Superconductivity" [1] in which they introduced the first generally accepted microscopic model of superconductivity based on the idea of electron pairing due to an effective attraction mediated by phonons. In recognition of this work they were awarded the Nobel prize in 1972. Roughly speaking, the idea is that negatively charged electrons can attract each other due to the presence of the positively charged ion cores. If an electron moves away from such an ion core it attracts the core and makes it move a little bit in the same direction. On the other side of the ion core there are several electrons which are attracted by it in a way similar to the way the core is attracted by the original electron, which means that this original electron effectively attracts the electrons on the other side of the ion core. Cooper [2] realized that even a tiny attractive interaction between particles in a Fermi gas leads to pairing between them. The so-called Cooper-pairs that form as a consequence of the attractive interaction should not be thought of as molecules because in metallic superconductors their spatial extension is larger than the average spacing between the particles by a factor of a few hundreds up to 1000. Nevertheless, the pairs behave, at least approximately, like Bosons and form a condensed state which is similar to but not identical with a Bose-Einstein condensate. It turns out that with this strongly correlated quantum state typical properties of superconductors such as the infinite conductivity, the Meissner effect, flux quantization and the isotope effect can be explained, see [1, 3].

From a mathematical point of view, the heart of BCS theory is a nonlinear integral equation for the gap function $\Delta : \mathbb{R}^d \to \mathbb{C}$ which is usually called the BCS gap equation.

In the theory of superconductivity $\Delta(p)$ is interpreted as the energy one has to expend in order to break up a Cooper-pair whose constituents have relative momentum p. The BCS gap equation in its usual form is given by

$$\Delta(p) = -\int_{\mathbb{R}^d} V(p,q) \frac{\tanh\left(\frac{\sqrt{(q^2-\mu)^2+|\Delta(q)|^2}}{2T}\right)}{\sqrt{(q^2-\mu)^2+|\Delta(q)|^2}} \Delta(q) \mathrm{d}q. \tag{1.1}$$

Here $V(p,q)$ denotes the interaction potential, which usually has a dominant negative part, $T \geq 0$ is the temperature of the sample, $\mu \in \mathbb{R}$ the chemical potential and $d \geq 1$ the spatial dimension. As one can directly check, Eq. (1.1) always allows for the solution $\Delta = 0$. The main question motivated from physics is whether the gap equation has a non-trivial solution, that is, a solution with $\Delta \neq 0$. If this is the case, the system is said to be in the superconducting state; if not, it is said to be in the normal state. The answer to this question usually depends strongly on the parameter T and of course on V. As we will see below, one can show under very general assumptions that there exists a critical temperature $T_c \geq 0$ such that for $T < T_c$ the BCS gap equation has a non-trivial solution, whereas for $T \geq T_c$ it allows only for the trivial solution. In principle, this behavior is a consequence of the fact that the expression $\tanh\left(x/(2T)\right)/x$ is monotone increasing in x for fixed T as well as in T for fixed x.

As Legget points out in [4], the BCS gap equation can be realized as the Euler-Lagrange equation of a functional which is usually called the BCS functional. At least heuristically, this functional can be derived from quantum mechanics in three steps, see [16]: One considers the free energy $F = E - \mu N - TS$, where E is the internal energy, μ the chemical potential, N the particle number, T the temperature and S the entropy of the system. To be more precise, $E - \mu N - TS$ is called the thermodynamic potential in the physics literature and the free energy is given by $E - TS$. Nevertheless, we stick to our convention because it has become common in the mathematical physics literature. In order to compute F in the framework of BCS theory, one restricts attention to what are called BCS states or quasi-free states. A state is called quasi-free if it fulfils Wick's theorem, see [8], that is, expectation values of a finite number of creation and annihilation operators reduce to products of expectation values of only two creation and/or annihilation operators. In the second step, one additionally claims that the states are translation and $SU(2)$ rotation invariant, whereas in the third step one neglects two terms in the free energy called the direct and the exchange term. The first term is a density-density correlation energy, whereas the second term takes the correlation energy due to the fermionic nature of the particles into account. What one finally obtains is called the BCS functional. The striking feature of the BCS energy functional and of the BCS gap equation is that on the one hand, they are able to describe the phase transition from the normal state (Fermi liquid) with finite electric resistance to a paired superconducting state, and, on the other hand, they are simple enough to allow for a mathematical treatment. The BCS functional depends on the reduced one-particle density matrix γ and on the Cooper-pair wave function α. The basic question is whether for the minimizer of the BCS functional one has $\alpha \neq 0$, which is, as we will see later, equivalent to having a non-trivial solution of the BCS gap equation (α and Δ are related). From a physics point of view, the BCS functional and the BCS gap equation have been studied in a vast number of situations and any attempt to give an overview of the existing literature would be hopeless. Accordingly, we will focus in the following on works devoted to the mathematical properties of the BCS gap equation and/or the BCS functional. The reader who is interested in the physics of superconductivity and BCS theory is refered to [3, 4, 5, 6, 7] and the references therein.

The first mathematical works on BCS theory [9, 10, 11, 12, 13, 14] considered the gap equation as the basic object and did not use the fact that it can be realized as the Euler-Lagrange equation of a functional. The interaction kernels treated by these authors are non-local in nature and are therefore adequate for describing the typical phonon-mediated effective interactions between electrons in metals and alloys. To our knowledge the first existence theorem for the BCS gap equation in the spherically symmetric case is given in [9]. Other works studying this special situation are [10] and [11]. In all these works, the authors give conditions on the kernel on the right hand side of the gap equations, for example a Lipschitz condition or a decay at infinity, which allow them to prove that the related integral operator is compact. In a second step they use a suitable fixed point theorem to show the existence of solutions. The first works treating the case where the gap function cannot be expected to be radial are [12, 13]. The authors of these works consider the situation of a negative interaction kernel $V(p, q)$, where one can expect that a solution of the gap equation is unique (up to a phase) and positive (here a choice of the phase has been made). Besides other things they establish the existence of a critical temperature T_c such that the BCS gap equation has a non-trivial solution for $T < T_c$ and allows only for the trivial solution for $T \geq T_c$. In [13] the authors show that unique solutions exist in a class of functions that can be approximated by solutions on bounded domains, which is helpful since it also suggests a numerical scheme for computing solutions. In this work, the authors also study the case of an interaction kernel without a fixed sign and present examples where the solution is not unique. Finally, a multiband version of the equation, which, apart from other applications, is also relevant in the theory of high-temperature superconductivity, has been studied in [14]. Also for this model, the authors can prove the existence of a critical temperature with the above-mentioned properties.

After the first Bose-Einstein condensate had been realized in an experiment in 1995, for which were the experimentalists Cornell, Ketterle and Wieman awarded the Nobel prize in 2001, the door to the study of quantum phenomena in cold bosonic and fermionic gases was opened. From a many-body physics point of view such systems are of great interest because they are less complex than most of the systems one has to deal with in solid state physics. This is because in quantum gases the range of typical interactions originating from polarization effects between the atoms is small compared to the average distance between them. Additionally, there exist techniques based on so-called Feshbach resonances that allow experimentalists to tune the strength of the effective interactions almost arbitrarily. Due to this freedom, experiments with cold atomic gases are often referred to as quantum physics simulators. This has to be compared with the complicated situation in solids, where one always has an interplay between the coulombic and phononic interactions of the electrons, where the latter depend strongly on the geometry of the lattice as well as on the chemistry of the ion cores forming the lattice. When interactions typical for quantum gases are considered in the BCS gap equation, it is a widely used theoretical tool to study the normal-to-superfluid phase transition in cold and dilute Fermi gases, see [15] and references therein. We note that superconductivity and superfluidity in principle refer to the same effect, except that the word superconductivity is only used in the special situation where the particles carry an electric charge.

Motivated by these developments, Hainzl, Hamza, Seiringer and Solovej considered in [16] the BCS functional for general local pair interactions suitable for describing the interactions in cold Fermi gases. The key observation in this work is that although the BCS functional and the BCS gap equation are highly nonlinear, the question of whether the gap equation has a non-trivial solution can be answered by a linear criterion. Using this linear criterion, Hainzl, Seiringer and co-authors could investigate the dependence of the critical temperature on parameters of the system such as the interaction potential V and the chemical potential μ in great detail. The BCS functional and the results of [16] are the basis for our work and so we discuss these results to some extent. In [16, Definition 1] the BCS functional in three space dimensions is defined as follows:

Definition 1.1 *Let $\mathcal{D}$ denote the set of pairs (γ, α) with $\gamma \in L^1(\mathbb{R}^3, (1+p^2)dp)$, $0 \leq \gamma(p) \leq 1$, and $\alpha \in H^1(\mathbb{R}^3, dx)$, satisfying $|\hat{\alpha}(p)|^2 \leq \gamma(p)(1-\gamma(p))$. Let $V \in L^{3/2}(\mathbb{R}^3, dx)$ and $\mu \in \mathbb{R}$. For $T = 1/\beta \geq 0$ and $(\gamma, \alpha) \in \mathcal{D}$, the energy functional $\mathcal{F}_\beta$ is defined to be*

$$\mathcal{F}_\beta(\gamma, \alpha) = \int_{\mathbb{R}^3} \left(p^2 - \mu\right) \gamma(p)\, dp + \int_{\mathbb{R}^3} V(x)\, |\alpha(x)|^2\, dx - \frac{1}{\beta} S(\gamma, \alpha), \tag{1.2}$$

where

$$S(\gamma, \alpha) = -\int_{\mathbb{R}^3} \left[s(p) \ln s(p) + (1 - s(p)) \ln(1 - s(p))\right] dp, \tag{1.3}$$

with $s(p)$ determined by $s(1-s) = \gamma(1-\gamma) - |\hat{\alpha}|^2$.

We note that it is natural to require $V \in L^{3/2}(\mathbb{R}^3)$ because it guarantees the relative form boundedness of the potential V with respect to the Laplacian. Using standard techniques, it can be shown that the BCS functional is bounded from below and always attains its infimum, see [16, Proposition 2]. If $V = 0$, its minimizer can be computed explicitly and is given by the pair $(\gamma_0, 0)$ with $\gamma_0(p) = [1 + \exp(\beta(p^2 - \mu))]^{-1}$. We refer to this state as the normal state, that is the state in which no superfluidity (or superconductivity) is present. As we will see in a few lines, this naming is justified because $(0, \gamma_0)$ minimizes the BCS functional if and only if the BCS gap equation has only the trivial solution $\Delta = 0$. Accordingly, the main question in the study of the BCS functional is whether a non-vanishing α can lower the energy for a given $V \neq 0$. If this is so, the minimizer of the BCS functional will be given by a pair (γ, α) with $\alpha \neq 0$ and one says that the system is in the superfluid phase. As we have already mentioned, the above question can be answered with the help of a linear criterion which is captured in [16, Theorem 1]:

Theorem 1.1 *Let $V \in L^{3/2}(\mathbb{R}^3)$, $\mu \in \mathbb{R}$, and $\infty > T = 1/\beta \geq 0$. Then the following three statements are equivalent:*

(i) The normal state $(\gamma_0, 0)$ is unstable under pair formation, i.e.,

$$\inf_{(\gamma,\alpha)\in\mathcal{D}} \mathcal{F}_\beta(\gamma, \alpha) < \mathcal{F}_\beta(\gamma_0, 0). \tag{1.4}$$

(ii) There exists a pair $(\gamma, \alpha) \in \mathcal{D}$, with $\alpha \neq 0$, such that

$$\Delta(p) = -\frac{p^2 - \mu}{\gamma(p) - \frac{1}{2}} \hat{\alpha}(p) \tag{1.5}$$

satisfies the BCS gap equation

$$\Delta = -\frac{1}{(2\pi)^{3/2}} \hat{V} * \left(\frac{\tanh\left(\frac{\beta E}{2}\right)}{E} \Delta \right), \quad \textit{with } E(p) = \sqrt{(p^2 - \mu)^2 + |\Delta(p)|^2}. \tag{1.6}$$

(iii) The linear operator

$$K_T + V, \qquad K_T = \frac{p^2 - \mu}{\tanh\left(\frac{p^2 - \mu}{2T}\right)}, \tag{1.7}$$

has at least one negative eigenvalue.

It should be mentioned that the operator in part three of the above theorem appears as the second variation of the BCS functional at the normal state $(\gamma_0, 0)$. If the second variation is negative, the normal state is unstable and hence cannot be the minimizer of the BCS functional. Accordingly, *(iii)* implies *(i)*. On the other hand, the theorem tells us that if $(\gamma_0, 0)$ is not the minimizer, it cannot be a local minimum which is a non-trivial statement. The theorem also tells us that from any non-trivial solution of the BCS gap equation, one can construct a pair (γ, α) whose energy is strictly smaller than that of the normal state. Hence, if one is interested in the question of whether the system is superfluid or not, one can either study the BCS gap equation or the BCS functional, because both contain the relevant information.

Since the operator K_T+V is monotone in T, the above theorem motivates the following definition of the critical temperature:

$$T_c = \inf \{T | K_T + V \geq 0\} . \quad (1.8)$$

It can easily be seen that for $T < T_c$ the system is in the superfluid phase while for $T \geq T_c$ it is in the normal phase. Note that the definition of T_c is entirely in terms of the linear operator $K_T + V$ and does not involve the BCS gap equation or the BCS functional. In [16, Theorem 3] it has been shown that $\mu > 0$ and V such that $V(x) \leq 0$ for almost all $x \in \mathbb{R}^3$ but not identically zero imply $T_c > 0$. Additionally, one can add a small positive part V_+ to the potential V and still guarantee $T_c > 0$ as long as V_+ is small in a suitable sense. Based on the definition Eq. (1.8), the behavior of the critical temperature in the weak coupling limit and in the low density limit has, among other things, been studied in [16, 17, 18, 20, 19]. The study of the weak coupling limit is based upon the fact that the Birman-Schwinger operator related to $K_T + V$ becomes singular on the Fermi sphere as T tends to zero. This fact can be used to derive a characterization of the critical temperature in terms of an appropriate operator acting on L^2-functions on the Fermi sphere which encodes the singularity. Having this characterization at hand, the authors apply spectral theoretic methods to analyse this operator. Their analysis yields a generalization of a formula for T_c in the small-coupling limit, previously known in the physics literature for special interactions, to a class of general interaction potentials. Using a novel expression of the scattering length a and assuming that $a < 0$, they also generalize a formula for the critical temperature in the low density limit to general interaction potentials. Finally, in [25] finally the translation-invariant BCS functional has been studied in the situation where the spin-up and spin-down states of the fermions in the gas are not equally occupied, which results in a richer phase diagram of the system depending also on the imbalance of the spin occupation numbers.

After these works on the translation-invariant BCS functional, several works on the BCS functional with external electric and magnetic fields followed, the most important one being [28]. Before the microscopic mechanism of superconductivity had been discovered, several aspects of the phenomenon could be described by means of phenomenological theories such as, for example, the London equations, which model the electrodynamics of a superconductor [5]. Another such theory is based on the Ginzburg-Landau equations, which, among many other things, can be used to describe vortex lattices in superconductors exposed to magnetic fields [26]. Only two years after BCS published their microscopic theory of superconductivity, Gorkov [27] realized that close to the critical temperature (in the superconducting phase) the Ginzburg-Landau equations can be formally derived from the BCS model. It is the contribution of [28] to make this formal derivation rigorous. To implement the closeness to the critical temperature, the authors consider a scaling limit where $T = T_c(1 - h^2)$ and $h \ll 1$. Close to T_c all but small fields destroy superconductivity and so the external fields have to be chosen appropriately. Important for the rigorous derivation of Ginzburg-Landau theory is also the assumption that the external fields vary on a macroscopic scale whereas the interaction between the particles takes place on a microscopic scale. Under these assumptions, the authors of [28], roughly speaking, show that the energy of any near minimizer (γ, α) is to leading order in h given by the minimum of a suitably chosen Ginzburg-Landau functional whose coefficients can be computed entirely from the translation-invariant BCS functional. That is, they only depend on microscopic properties of the system. Additionally, the Cooper-pair wave function α of the just mentioned near minimizer is to leading order given by

$$\alpha(x, y) = \alpha_0(x - y)\psi\left(\frac{x + y}{2}\right) \tag{1.9}$$

where α_0 is some universal function related to the the translation-invariant BCS functional and ψ is the minimizer of the above mentioned Ginzburg-Landau functional. Thus the translation-invariant BCS functional describes the relative coordinate of the Cooper-pair wave function and Ginzburg-Landau theory the center of mass coordinate. Easily readable summaries of this work can be found in [30] and [31]. In [29] the authors use the machinery developed in [28] to study the influence of the external electric and magnetic fields on the critical temperature. They find that the leading order of the critical temperature is given by the critical temperature of the translation-invariant BCS functional while the next-to-leading order in $(T_c - T)/T_c$ is determined by the lowest eigenvalue of the linearization of the Ginzburg-Landau equation that arises as the Euler-Lagrange equation of the Ginzburg-Landau functional that already appeared above. Other works studying the BCS functional in the presence of external fields but at $T = 0$ are [22, 24]. An extension of the derivation of Ginzburg-Landau theory from the BCS functional in the case of several order parameters (p-wave, d-wave pairing) but without external fields has been given in [32].

As we mentioned at the beginning of this section, the BCS functional can be derived from quantum mechanics in a formal way. One step in this derivation is to neglect the direct and exchange terms in the energy. This approximation has been investigated in [21]. It was that for interactions which are sufficiently short-range, the direct and exchange terms only give rise to a renormalization of the chemical potential and leave all other properties of the BCS functional unchanged. Therefore, it can be interpreted as a justification of this approximation for systems with sufficiently short-range interactions. The BCS functional with direct and exchange terms is sometimes referred to as the Bogoliubov-Hartree-Fock functional. Another important contribution of [21] is the first proof of pairing in the Bogoliubov-Hartree-Fock model in the continuum. In [23] the same model at $T = 0$ but in the presence of an external electric potential was investigated. It was shown that in the low-density limit, the ground state of the model consists of a Bose-Einstein condensate of tightly bound fermion pairs which can be described by a suitably chosen Gross-Pitaevskii energy functional. The result in [23] is an extension of [22], where a similar derivation of the Gross-pitaevskii functional from the BCS functional, that is from the model without direct and exchange terms, was carried out.

Ginzburg-Landau theory arises from BCS theory in a natural way in the case of equilibrium states. Less natural and highly debated in the physics literature are the related questions in the case of dynamics. The time-dependent equations governing the dynamics of a state in BCS theory are called the Bogoliubov-de-Gennes (BdG) equations and like their time-independent counterpart, they can formally be derived from quantum mechanics. In [33] it could be shown that for an explicitly given initial state (γ_0, α_0) with $\alpha_0 \neq 0$, the Cooper-pair wave function does not decay in time as the corresponding diffusive Ginzburg-Landau equation predicts. Only the linearized part of the BdG equations shows this long time behavior. That is, although the leading order in a formal expansion of the equations predicts a diffusive behavior, the full nonlinear BdG equations behave differently. The reason for this effect is that small denominators invalidate the approximation. Hence, the non-decaying order parameter is a purely non-linear effect. In case of a one-dimensional sample with contact interaction this behavior has been confirmed numerically in [34].

1.2 Summary of main results

1.2.1 Note on a family of monotone quantum relative entropies

Let $\varphi : [0,1] \to \mathbb{R}$ be a convex function with $\varphi \in \mathcal{C}([0,1]) \cap \mathcal{C}^1((0,1))$. For two hermitian matrices A, B with $0 \leq A, B \leq 1$, define

$$\mathcal{H}_\varphi(A,B) = \operatorname{Tr}\left[\varphi(A) - \varphi(B) - \varphi'(B)\,(A-B)\right]. \tag{1.10}$$

Following [36], we call $\mathcal{H}_\varphi(A,B)$ the generalized relative entropy of A with respect to B. Eq. (1.10) is a priori well-defined if $0 < B < 1$ or for arbitrary B if $\varphi \in \mathcal{C}^1([0,1])$. If 0 and/or 1 is contained in the spectrum of B and φ' has a discontinuity at one or both of these points, we define $\mathcal{H}_\varphi(A,B) = \infty$ except if $A = B$ on the corresponding subspaces. In this case, the trace is taken on their orthogonal complement. Examples for functions that one may choose for φ are

$$\begin{aligned}\varphi_1(x) &= x\ln(x), \\ \varphi_2(x) &= x\ln(x) + (1-x)\ln(1-x), \\ \varphi_3(x) &= x\ln(x) - (1+x)\ln(1+x).\end{aligned} \tag{1.11}$$

Insertion of φ_1 yields the usual relative entropy used in quantum statistical mechanics. The function φ_2 on the other hand is used in case of fermionic quasi-free states and φ_3 when bosonic quasi-free states are considered. From a more mathematical point of view, the generalized relative entropy of A with respect to B can be interpreted as a distance measure between the matrices A and B. This is because $\mathcal{H}_\varphi(A,B)$ is positive and if φ is strictly convex it equals zero if and only if $A = B$. Both statements can easily be seen with the help of Klein's inequality [35], which we state here in the following form.

Lemma 1.1 *Let A, B be two hermitian matrices and for $k = 1, ..., N$, let $f_k : \sigma(A) \to \mathbb{R}$, $g_k : \sigma(B) \to \mathbb{R}$, where $\sigma(A)$ and $\sigma(B)$ denote the spectra of A and B, respectively. Assume*

$$\sum_{k=1}^{N} f_k(\alpha) g_k(\beta) \geq 0 \quad \forall \alpha \in \sigma(A), \beta \in \sigma(B). \tag{1.12}$$

Then

$$Tr\left[\sum_{k=1}^{N} f_k(A) g_k(B)\right] \geq 0 \tag{1.13}$$

holds.

Hence, the question whether $\mathcal{H}_\varphi(A, B)$ is positive can be reduced to the question of whether a related inequality for numbers holds, which is assured by the convexity of the function φ. The same reasoning works for the second claim.

Distance measures between operators of the above kind are used extensively in many areas of mathematics, for example in quantum statistical mechanics, quantum information theory or random matrix theory, and so it is a natural question to ask whether the above notion of generalized relative entropy can be extended to a larger class of operators. The extension to trace-class operators is more or less straight forward but the related question for bounded and not necessarily compact operators on a separable Hilbert space X turns out to be more interesting. The simplest idea that one can possibly come up with is to consider Eq. (1.10) also for bounded operators. Unfortunately, this leads to several difficulties. First of all, Eq. (1.10) yields something well-defined only if the operator under the trace is trace-class. Since φ' is possibly unbounded, it is not easy to decide when this is the case, especially because most operators which are of interest from a physics point of view have continuous spectrum at 0 and/or 1. On the other hand, one would like to be able to derive upper and lower bounds for the generalized relative entropy in order to have an object that one can actually work with. But this also turns out to be not so easy since the operators A and B may have continuous spectrum, in which case Klein's inequality is not applicable. Hence, there is a need to consider alternatives.

In [36] Lewin and Sabin proposed the following idea inspired by the thermodynamic limit. Let X be a separable Hilbert space and $A, B \in \mathcal{L}(X)$ with $0 \leq A, B \leq 1$. Choose an increasing sequence $\{P_n\}_{n=1}^{\infty}$ of finite rank projections with $P_n \to 1$ in the strong operator topology. Here increasing means that the range of P_n is included in the range of P_m for $m \geq n$. Lewin and Sabin now consider the object $\mathcal{H}_\varphi(P_n A P_n, P_n B P_n)$ and investigate its behavior as a function of n. Among other things, they show in [36, Theorem 2]:

Theorem 2.2 *Let φ, A, B and $\{P_n\}_{n=1}^{\infty}$ be as above and assume additionally that φ' is operator monotone on $(0, 1)$. Then the sequence $\mathcal{H}_\varphi(P_n A P_n, P_n B P_n)$ is monotone* *increasing and possesses a limit in* $\mathbb{R}_+ \cup \{+\infty\}$. *This limit does not depend on the chosen sequence* $\{P_n\}_{n=1}^\infty$ *and hence*

$$\mathcal{H}_\varphi(A, B) := \lim_{n\to\infty} \mathcal{H}_\varphi(P_n A P_n, P_n B P_n) \tag{1.14}$$

is well defined.

Let us note that a function $f : \mathbb{D} \subseteq \mathbb{R} \to \mathbb{R}$ is called operator monotone if for all hermitian matrices A, B with $\sigma(A), \sigma(B) \subseteq \mathbb{D}$ the relations $A \leq B$ implies $f(A) \leq f(B)$. Operator monotone functions coincide with a certain class of analytic functions called Pick functions, see [38]. These functions have an analytic extension to the upper half plane H_+ that maps H_+ into itself. They also have an analytic extension to the lower half plane H_- which can be obtained by reflecting across the real interval that they were defined on initially. Let us also note that the three functions in Eq. (1.11) are such that their derivative is operator monotone, that is the most important examples used in physics are captured by the above theorem. Apart from the fact that it is based upon a more restrictive class of functions, the definition of Lewin and Sabin has two advantages and one disadvantage compared with the approach to use Eq. (1.10) as the definition of the generalized relative entropy. The first advantage is that we do not have to worry about $\mathcal{H}_\varphi(A, B)$ being well-defined any more. We cannot tell a-priori whether the generalized relative entropy is finite but we know that no other difficulties occur. The second advantage is that there is a natural way to derive upper and lower bounds. This is because bounds for $\mathcal{H}_\varphi(P_n A P_n, P_n B P_n)$ can be proven with the help of Klein's inequality. What remains to be done afterwards, is to find a way to treat the limits of these bounds as n tends to infinity which is often possible, see Chapter 3 and [36]. The disadvantage of the above definition of the generalized relative entropy is that it is hard to compute $\lim_{n\to\infty} \mathcal{H}_\varphi(P_n A P_n, P_n B P_n)$ for explicitly given operators A and B, which is often necessary in applications, especially if one wants to do arguments based on trial states.

This leads to a natural question left open in [36], namely whether it is true that

$$\lim_{n\to\infty} \mathcal{H}_\varphi(P_n A P_n, P_n B P_n) \stackrel{?}{=} \mathrm{Tr}\left[\varphi(A) - \varphi(B) - \varphi'(B)(A - B)\right] \tag{1.15}$$

holds. The purpose of Chapter 2 is to answer this question and we find that Eq. (1.15) in principle holds. Here in principle stands for the fact that the right hand side of Eq. (1.15) turns out to be not the correct limit in general but only in a special situation. Our main result consists of two theorems, one of which is the following statement:

Theorem 2.3 *Let* $\varphi \in C^0([0,1], \mathbb{R})$ *be such that* φ' *is operator monotone on* $(0, 1)$ *and let* $\{P_n\}_{n=1}^\infty$ *be defined as above. Then*

$$\lim_{n\to\infty} \mathcal{H}(P_n A P_n, P_n B P_n) = \mathrm{Tr}\left[\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)\Big|_{\alpha=0}\right] \tag{1.16}$$

with the understanding that either both sides are finite and equal to each other, or both sides are infinite.

The idea why the right-hand side of Eq. (1.16) is a more reasonable candidate for the limit than the right hand side of Eq. (1.15) is the following: The function φ' is operator monotone by assumption. Any primitive φ is an operator convex function which can easily be seen with the help of the integral representations of operator monotone and operator convex functions given for example in [38, Corollary V.4.5, Theorem V.4.6]. A function is called operator convex if the usual convexity inequality holds for hermitian matrices of arbitrary rank. Using this convexity inequality, one can check that the operator under the trace on the right hand side of Eq. (1.16) is, as long as it is well defined, positive. This lets us define its trace for arbitrary A and B, which is necessary because the left-hand side of Eq. (1.16) is well defined for arbitrary A and B. On the other hand, a straightforward computation that exploits the cyclicity of the trace shows that the right-hand side of Eq. (1.15) and the right-hand side of Eq. (1.16) coincide if A and B are matrices. Because of this and since the operator under the trace on the right-hand side of Eq. (1.16) is symmetric, its trace can be seen as a symmetrised extension of the right-hand side of Eq. (1.15) suitable for working with non-compact operators. Theorem 2.3 also tells us that the right-hand side of Eq. (1.16) is a more natural definition of the generalized relative entropy than the one given in Eq. (1.10).

When it comes to explicit computations, the right-hand side of Eq. (1.16) is already much more appealing than the limiting procedure we started with but still not satisfactory due to the non-commutative derivative with respect to α. As one expects after the discussion in the previous paragraph, it is possible to show that the right-hand side of Eq. (1.15) and the right-hand side of Eq. (1.16) equal each other if we impose some mild regularity conditions on the operators A and B. This is captured in the following theorem:

Theorem 2.4 *Let $\varphi \in C^0([0,1],\mathbb{R})$ be such that φ' is operator monotone on (0,1). Assume in addition that $(A-B)$, $\varphi(A)-\varphi(B)$ and $\varphi'(B)(A-B)$ are trace-class. Then $\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ is trace-class and the identity*

$$\begin{aligned}\mathrm{Tr}\left[\varphi(A)-\varphi(B)-\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)\Big|_{\alpha=0}\right] \\ =\mathrm{Tr}\left[\varphi(A)-\varphi(B)-\varphi'(B)(A-B)\right]\end{aligned} \tag{1.17}$$

holds.

This finally is a satisfactory result since it allows us to combine the advantages of the two different definitions of the generalized relative entropy without their disadvantages. The prize one has to pay is the restriction to a smaller class of functions from which φ can be chosen but since all examples from physics fall into this class there is no reason to worry about this too much. The correct viewpoint is rather that the definition of a generalized relative entropy with φ' operator monotone captures many of the beautiful and strong properties that the relative entropies from physics share.

The proofs of the above statements are based upon a novel integral representation of the generalized relative entropy that is of interest in itself. It can be derived from an integral representation of operator monotone functions and is captured in the following lemma:

Lemma 2.1 *Let* $\varphi \in C([0,1],\mathbb{R})$ *be such that* φ' *is operator monotone on* $(0,1)$*. Assume further that* B *has no eigenvalues at points of discontinuity of* φ' *with* $(A-B) \neq 0$ *on the corresponding eigenspaces. Then there exists a constant* $b \geq 0$ *and a unique Borel probability measure* μ *on* $[-1,1]$ *such that*

$$\operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi(\alpha A + (1-\alpha)B)\Big|_{\alpha=0}\right] = \tag{1.18}$$
$$2b \int_{-1}^{1}\int_{0}^{\infty} \operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t} Q \frac{1}{1+\lambda(1-2A)+t} Q \frac{1}{1+\lambda(1-2B)+t}\right] dt\ d\mu(\lambda),$$

where $Q = (A-B)$*. (To be precise,* μ *is unique only if* $b > 0$*.)*

Let us note that the requirement for A and B is not really a restriction because if it is not fulfilled the relative entropy cannot be finite. The formula on the right hand side of Eq. (1.18) has very appealing properties. To compute it, one first takes the trace of a bounded positive operator, which yields a positive measurable function of λ and t. Afterwards, this positive function is integrated against two positive measures. It is this structure that finally lets us take the limit $n \to \infty$, which is the key step in the proof of Theorem 2.3.

1.2.2 A non-periodic version of the BCS functional in an external field

If one wants to model a fermionic many-particle system in the thermodynamic limit which is exposed to external fields within the framework of BCS theory, there are several things one can do. One option is to consider the situation of periodic external fields, where the notion of a free energy per unit volume makes sense. Another one is to assume that the external fields are sufficiently localized, which allows one to measure the free energy of the sample with respect to that of a suitable reference state. This reference state could, for example, be the lowest-energy state of the system without external fields. The first approach has been pursued by Frank, Hainzl, Seiringer and Solovej in [28]. In this work, the authors considered a periodic sample and additionally assumed that the period is large in order to avoid finite size effects. Although this approach is very reasonable, it may be seen as a disadvantage that the enforced periodicity of the sample is somewhat artificial. It is the goal of Chapter 3 to follow the second approach and to define and investigate a version of the BCS functional without artificial boundary conditions.

Let us briefly motivate the definition of our BCS functional. As already mentioned above, it is natural to measure the free energy of the sample with respect to the free energy of the system where the external fields are absent. Hence, it is reasonable to consider the minimizer of the translation-invariant BCS functional as a reference state. It can be written in the form

$$\Gamma_0(p) = \begin{pmatrix} \gamma(p) & \hat{\alpha}_0(p) \\ \overline{\hat{\alpha}_0(p)} & 1 - \overline{\gamma_0(p)} \end{pmatrix} = \frac{1}{1 + e^{\beta H_{\Delta_0}(p)}}, \tag{1.19}$$
$$H_{\Delta_0}(p) = \begin{pmatrix} p^2 - \mu & \hat{\Delta}_0(p) \\ \overline{\hat{\Delta}_0(p)} & -(p^2 - \mu) \end{pmatrix}$$

where $\Delta_0(x) = 2V(x)\alpha_0(x)$. Eq. (1.19) is nothing but the Euler-Lagrange equation of the translation-invariant BCS functional and the gap equation can be derived from it. Although Γ_0 is the natural candidate for a reference state, we found it more convenient to work instead with the state

$$\Gamma_0^w = \frac{1}{1 + e^{\beta H_0^w}}, \tag{1.20}$$
$$H_0^w = \begin{pmatrix} (-i\nabla)^2 - \mu + W & \hat{\Delta}_0(-i\nabla) \\ \overline{\hat{\Delta}_0(-i\nabla)} & -\left((-i\nabla)^2 - \mu + W\right) \end{pmatrix}$$

where W denotes the external potential. We note that in the above formula all operators except for W are pseudo-differential operators. In contrast, W denotes a multiplication operator in position space, that is, $(W\psi)(x) = W(x)\psi(x)$ for all $\psi \in L^2(\mathbb{R}^3)$. Of course, the function W should be thought of as suitably localized. Heuristically, Γ_0^w is a better choice than Γ_0 because it already incorporates the external potential W. It is also convenient to write a general state (γ, α) in a matrix form similar to the one used in Eq. (1.19) and Eq. (1.20). In the following we denote

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \overline{\alpha} & 1 - \overline{\gamma} \end{pmatrix} \in \mathcal{L}(L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)). \tag{1.21}$$

The operator Γ is called the generalized one-particle density matrix and obviously contains the same information as the pair (γ, α).

Having a suitable reference state at hand, we have to give a meaning to $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) = \mathcal{F}_\beta(\Gamma) - \mathcal{F}_\beta(\Gamma_0^w)$ where

$$\text{“}\mathcal{F}_\beta(\Gamma) = \mathrm{Tr}_{L^2(\mathbb{R}^3)} \left[\left((-i\nabla)^2 - \mu + W \right) \gamma \right] + \int_{\mathbb{R}^6} V(x-y) \, |\alpha(x,y)|^2 \, \mathrm{d}(x,y) - TS(\Gamma)\text{”},$$
$$\text{“}S(\Gamma) = -\frac{1}{2} \mathrm{Tr}_{L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)} \left[\Gamma \ln(\Gamma) + (1 - \Gamma) \ln(1 - \Gamma) \right] \text{”}. \tag{1.22}$$

We use quotation marks in Eq. (1.22) to highlight that everything is meant only formally. If one inserted the minimizer of the translation-invariant BCS functional into

Eq. (1.22) each term of the functional would be infinite or possibly even ill defined. When we do some formal algebraic manipulations, it is possible to see that $\mathcal{F}_\beta(\Gamma, \Gamma_0^w)$ can be written as

$$\mathcal{F}_\beta(\Gamma, \Gamma_0^w) = \frac{1}{2\beta}\mathcal{H}(\Gamma, \Gamma_0^w) + \int_{\mathbb{R}^6} V(x-y)\,|\alpha(x,y) - \alpha_0^w(x,y)|^2\,\mathrm{d}(x,y) \\ + 2\mathrm{Re}\int_{\mathbb{R}^6} V(x-y)\,(\alpha(x,y) - \alpha_0^w(x,y))\,\overline{(\alpha_0^w(x,y) - \alpha_0(x-y))}\mathrm{d}(x,y), \tag{1.23}$$

where the relative entropy $\mathcal{H}(\Gamma, \Gamma_0^w)$ of the state Γ with respect to the state Γ_0^w is given by

$$\mathcal{H}(\Gamma, \Gamma_0^w) = \mathrm{Tr}\left[\varphi(\Gamma) - \varphi(\Gamma_0^w) - \frac{\mathrm{d}}{\mathrm{d}s}\varphi(s\Gamma + (1-s)\Gamma_0^w)\Big|_{s=0}\right], \tag{1.24}$$
$$\varphi(x) = x\ln(x) + (1-x)\ln(1-x).$$

The fact that we write the relative entropy in the above form and not as

$$\mathrm{Tr}\left[\varphi(\Gamma) - \varphi(\Gamma_0^w) - \varphi'(\Gamma_0^w)(\Gamma - \Gamma_0^w)\right] \tag{1.25}$$

is motivated by the analysis of Chapter 2. The advantage of the form of the BCS functional given in Eq. (1.23) is that only differences of Γ and Γ_0^w as well as differences of α and α_0^w appear. Hence, there is a reasonable chance that Eq. (1.24) is a good definition for the BCS functional.

As in [28], we consider the physical situation of a system in the superfluid phase close to its critical temperature. In such a system all but small fields living on the energy scale $T_c - T$ would destroy the superfluid state immediately. Hence, we have to consider W to be small. Additionally, we assume that the external field varies on a macroscopic scale while the interaction potential varies on a microscopic scale. If we choose macroscopic coordinates, this situation is implemented mathematically by choosing $T = T_c(1 - h^2 D)$ for $D > 0$ and $h \ll 1$. The interaction potential is given by $V(x/h)$ and the external potential reads $h^2 W(x)$. Our BCS functional in this scaling is given by

$$\mathcal{F}_\beta(\Gamma, \Gamma_0^w) = \frac{1}{2\beta}\mathcal{H}(\Gamma, \Gamma_0^w) + \int_{\mathbb{R}^6} V\left(\frac{x-y}{h}\right)|\alpha(x,y) - \alpha_0^w(x,y)|^2\,\mathrm{d}(x,y) \\ + 2\mathrm{Re}\int_{\mathbb{R}^6} V\left(\frac{x-y}{h}\right)(\alpha(x,y) - \alpha_0^w(x,y))\,\overline{\left(\alpha_0^w(x,y) - h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\right)}\mathrm{d}(x,y), \tag{1.26}$$

where

$$\Gamma_0^w = \begin{pmatrix} \gamma_0^w & \alpha_0^w \\ \overline{\alpha_0^w} & 1 - \overline{\gamma_0^w} \end{pmatrix} = \frac{1}{1 + e^{\beta H_0^w}}, \tag{1.27}$$
$$H_0^w = \begin{pmatrix} (-ih\nabla)^2 - \mu + h^2 W(x) & \hat{\Delta}_0(-ih\nabla) \\ \hat{\Delta}_0(-ih\nabla) & -((-ih\nabla)^2 - \mu + h^2 W(x)) \end{pmatrix}.$$

We consider the BCS functional defined by Eq. (1.26) for states Γ such that $\mathcal{H}(\Gamma, \Gamma_0^w) < \infty$. Such states are called *admissible*.

Our main result is a lower bound for the BCS functional from which a-priori estimates for states with energy smaller or equal than the one of the reference state can be derived. Apart from some mild regularity assumptions on V and W, our main assumption is that $\hat{V}(p) \leq 0$ holds for almost all $p \in \mathbb{R}^3$. This condition implies in particular that the minimizer of the translation-invariant BCS functional is unique up to a phase. To state our result, we have to introduce some more notation. If $\alpha_0(x)$ is the Cooper-pair wave function of the minimizer of the translation-invariant BCS functional, we can write the Cooper-pair wave function $\alpha(x, y)$ of any *admissible* state as

$$\alpha(x,y) = h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\psi(y) + \xi_0(x,y). \tag{1.28}$$

This decomposition plays an important role in the construction of the lower bound for the BCS functional. Our result is captured in the following theorem.

Theorem 3.1 *Let Γ be an admissible state with $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \leq 0$. Then for $r > 0$ large enough and $h > 0$ small enough, there exist constants $C_1, C_2 > 0$ (depending on r) such that*

$$\begin{aligned}\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \geq C_1\Bigg(h\,\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + h\left\|\widehat{|\psi|^2 - 1}\right\|^2_{L^2(B_r)} + \|\xi_0\|^2_{H_1(\mathbb{R}^6)} \qquad (1.29)\\ + \|\gamma - \gamma_0^w\|^2_{H^1(\mathbb{R}^6)}\Bigg) - C_2 h.\end{aligned}$$

In the above equation, B_r denotes the ball of radius r centered around zero and the $H^1(\mathbb{R}^6)$-norms are, according to our choice of coordinates, given by $\|f\|^2_{H^1(\mathbb{R}^6)} = \|f\|^2_{L^2(\mathbb{R}^6)} + \|h\nabla_x f\|^2_{L^2(\mathbb{R}^6)} + \|h\nabla_y f\|^2_{L^2(\mathbb{R}^6)}$. Eq. (1.29) implies the a-priori bounds

$$\begin{aligned}\|\nabla\psi\|_{L^2(\mathbb{R}^3)} + \big\||\psi|^2 - 1\big\|_{L^2(\mathbb{R}^3)} &\leq C, \qquad (1.30)\\ \|\xi_0\|_{H^1(\mathbb{R}^6)} + \|\gamma - \gamma_0^w\|_{H^1(\mathbb{R}^6)} &\leq Ch^{1/2},\end{aligned}$$

for an appropriately chosen constant $C > 0$.

Let us note that although the construction of a lower bound for the periodic version of the BCS functional treated in [28] is very easy, this is not the case for our functional. It turns out that the BCS functional we treat here has a completely different mathematical structure, which already makes it a challenging task to construct a lower bound. The strategy of proof used to prove Theorem 3.1 has to be compared with the proof of the a-priori estimates carried out in [28, Chapter 5]. Nevertheless, it has to be noted that additional severe difficulties arise from the fact that on the one hand one has to use the fact that W localizes possible deviations from the translation-invariant system and on the other hand that one positive term has a more complicated structure than the corresponding term in [28, Chapter 5].

Although our proof is carried out in the scaling suitable to derive Ginzburg-Landau theory, we believe that our strategy carries over to the construction of a lower bound for the unscaled functional. This is because the leading order of the Cooper-pair wave function is scaled such, that the leading order contributions of the three terms in the BCS functional are of the same order in h. Only the terms proportional to the next-to-leading order of the Cooper-pair wavefunction ξ_0 are easier to treat in our scaling, which makes us believe that we already capture the main difficulties of the general situation. Since we work in the special scaling mentioned above, the a-priori estimates in Theorem 3.1 are the first step towards an extension of the derivation of Ginzburg-Landau theory to the present set-up.

1.2.3 No translational symmetry breaking in the BCS model with radial pair interaction

In Chapter 4 we consider a two-dimensional version of the BCS functional with a radial pair interaction. Our main result is that the translational symmetry of the system is not broken in a certain temperature interval below the critical temperature. In the case of vanishing angular momentum, our results carry over to the three-dimensional case.

To be more precise, we consider a periodic version of the BCS functional that has been introduced and investigated in [28, 29]. We describe BCS states in $d \geq 1$ spatial dimensions by their generalized one-particle density matrix

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \overline{\alpha} & 1-\overline{\gamma} \end{pmatrix} \in \mathcal{L}\left(L^2(\mathbb{R}^d) \oplus L^2(\mathbb{R}^d)\right), \tag{1.31}$$

which obeys $0 \leq \Gamma \leq 1$. In case of vanishing external fields, the periodic BCS functional is given by

$$\mathcal{F}^{per}(\Gamma) = \mathrm{Tr}_\Omega\left[\left(-\nabla^2 - \mu\right)\gamma\right] + \int_{\Omega\times\mathbb{R}^d} V(x-y)\,|\alpha(x,y)|^2\,\mathrm{d}(x,y) - TS(\Gamma), \tag{1.32}$$
$$S(\Gamma) = -\frac{1}{2}\mathrm{Tr}_\Omega\left[\Gamma\ln(\Gamma) + (1-\Gamma)\ln(1-\Gamma)\right].$$

The operators γ and α in Eq. (1.31) are assumed to be periodic with period one. In terms of kernels, this means $\gamma(x+v, y+v) = \gamma(x,y)$ for all $v \in \mathbb{Z}^d$ and the same for $\alpha(x,y)$. By $\mathrm{Tr}_\Omega[A] = \mathrm{Tr}[\chi_\Omega(x)A\chi_\Omega(x)]$ we denote the usual trace per unit volume with χ_Ω being the characteristic function of the cube $\Omega = [0,1]^d$. The complex conjugate operators in the formula for Γ are defined by $\overline{\gamma} = C\gamma C$, where C denotes complex conjugation. In terms of kernels this reads $\overline{\gamma}(x,y) = \overline{\gamma(x,y)}$. The property $0 \leq \Gamma \leq 1$

implies $0 \leq \gamma \leq 1$, $\alpha\overline{\alpha} \leq \gamma(1-\gamma)$ as well as $\alpha^* = \overline{\alpha}$, where again in terms of integral kernels the last statement reads $\alpha(x,y) = \alpha(y,x)$. A periodic BCS state is called *admissible* if $\mathrm{Tr}_\Omega\left[(1-\Delta)\gamma\right] < \infty$ holds. Let us note that the translation-invariant BCS functional $\mathcal{F}_\beta$ is obtained from $\mathcal{F}^{per}$ by restricting the set of possible states to translation-invariant ones. For the sake of convenience, we suppress the index β and write $\mathcal{F} \equiv \mathcal{F}_\beta$ for the translation-invariant BCS functional in the following.

To state our result, we have to introduce a little bit of notation. Let us fix $d = 2$ and remind that the critical temperature of the translation-invariant BCS functional is defined by $T_c = \inf\{T \geq 0 \mid K_T + V|_{\mathrm{sym}} \geq 0\}$, see Chapter 1.1. In contrast to the definition in this Chapter we view $K_T + V$ as an operator on $L^2_{sym}(\mathbb{R}^2)$, that is on functions f with $f(x) = f(-x)$ for a.e. $x \in \mathbb{R}^2$. This is because our Cooper-pair wave functions are symmetric. If we assume that V is a radial function, $K_T + V$ commutes with rotations in $\mathbb{R}^2$ and all its eigenfunctions are of the form $e^{i\ell\theta}\alpha_*(|p|)$ for $\ell \in 2\mathbb{Z}$, where θ is the angle of p in polar coordinates. In the following, we denote $\mathcal{H}_\ell = \left\{\alpha \in H^1_{sym}(\mathbb{R}^2) \mid \alpha(p) = e^{i\ell\theta}\sigma(|p|)\right\}$. Motivated by this, we introduce a sector-wise critical temperature by

$$T_c(\ell) = \inf\left\{T \geq 0 \mid (K_T + V)|_{\mathcal{H}_\ell} \geq 0\right\}. \tag{1.33}$$

One can easily check that $T_c(-\ell) = T_c(\ell)$ for all $\ell \in 2\mathbb{Z}$, as well as $T_c = \max_{\ell \in 2\mathbb{Z}} T_c(\ell)$. In the same spirit, we introduce a sector-wise translation-invariant BCS functional. Let

$$\mathcal{D}_\ell = \left\{(\gamma, \alpha) \in \mathcal{D} \mid \gamma(p) = \tilde{\gamma}(|p|), \hat{\alpha}(p) = e^{i\ell\theta}\sigma_\ell(|p|)\right\}, \tag{1.34}$$

where $\mathcal{D}$ is the domain of the translation-invariant BCS functional, see Chapter 1.1. The translation-invariant BCS functional restricted to $\mathcal{D}_\ell$ is denoted by $\mathcal{F}_\ell$, that is, $\mathcal{F}_\ell = \mathcal{F}|_{\mathcal{D}_\ell}$. We note that $T_c(\ell)$ is the critical temperature of the restricted BCS functional $\mathcal{F}_\ell$. Having these definitions at hand, we state our main result:

Theorem 4.1 *Let $V \in L^2(\mathbb{R}^2)$ with $\hat{V} \in L^r(\mathbb{R}^2)$, $r \in [1,2)$, be radial and assume that $T_c > 0$. Suppose that there exist $\ell_0, \ell_1 \in 2\mathbb{Z}$ such that*

$$T_c(\ell_0) > T_c(\ell_1) \geq T_c(\ell) \tag{1.35}$$

for all $\ell \in 2\mathbb{Z} \setminus \{\pm\ell_0\}$. If $(\gamma_{\ell_0}, \alpha_{\ell_0}) \in \mathcal{D}_{\ell_0}$ is a minimizer of $\mathcal{F}_{\ell_0}$ (which always exists), then $(\gamma_{\ell_0}, \alpha_{\ell_0})$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0})$ minimize the full BCS functional $\mathcal{F}^{per}$ for $T \in [T_c(\ell_1), T_c)$. Moreover, $\sigma_{\ell_0} = \sigma_{-\ell_0}$ up to phases. For $T \in (T_c(\ell_0), T_c)$ these are the only minimizers of $\mathcal{F}^{per}$.

Apart from the fact that the translational symmetry of the system is not broken, the above theorem tells us that Cooper-pairs are in a definite angular momentum eigenstate. In case of $\ell_0 = 0$, this means that the minimizing Cooper-pair wave function is radial and that the minimizer is unique, up to a phase in front of α_{ℓ_0}. For of a small interaction potential, the methods of [17, 19] can be used to decide in which angular momentum sector the ground state of $K_{T_c} + V$ lies. In this case, it is sufficient to consider a certain operator acting on functions on the Fermi sphere, whose eigenvalues can be computed explicitly.

In the special case where $\ell_0 = 0$, our results carry over to three spatial dimensions. Let $\mathcal{H}_r = \left\{\alpha \in H^1_{sym}(\mathbb{R}^3) \mid \alpha(p) = \sigma_0(|p|)\right\}$ as well as

$$T_c' = \inf\left\{T \geq 0 \mid (K_T + V)|_{\mathcal{H}_r^\perp} \geq 0\right\}. \tag{1.36}$$

It should be compared to $T_c(\ell_1)$ in the two-dimensional case. The following statement holds for the three-dimensional model:

Theorem 4.2 *Let $V \in L^2(\mathbb{R}^3)$ with $\hat{V} \in L^r(\mathbb{R}^3)$, for some $r \in [1, 12/7)$, be radial and assume that $T_c > 0$. Assume further that zero is a non-degenerate eigenvalue of $K_{T_c} + V$. Then, for $T \in [T', T_c)$, there exists a pair (γ_0, α_0) with γ_0 and α_0 being radial functions, that minimizes the BCS functional $\mathcal{F}^{per}$. Moreover, up to a phase in front of α_0, (γ_0, α_0) is the only minimizer of $\mathcal{F}^{per}$ for $T \in (T_c', T_c)$.*

Prior to this work, it was only known that the translational symmetry of the BCS model is not broken if $\hat{V} \leq 0$ and not identically zero, see [31]. Under these assumptions the minimizer is, up to a phase in front of the Cooper-pair wave function, unique. This is also the only situation, in which it was known that Cooper-pairs are in a definite angular momentum state.

CHAPTER 2

Note on a family of monotone quantum relative entropies

Abstract. Given a convex function φ and two hermitian matrices A and B, Lewin and Sabin study in [36] the relative entropy defined by $\mathcal{H}(A,B) = \operatorname{Tr}[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)]$. Amongst other things, they prove that the so-defined quantity is monotone if and only if φ' is operator monotone. The monotonicity is then used to properly define $\mathcal{H}(A,B)$ for bounded self-adjoint operators acting on an infinite-dimensional Hilbert space by a limiting procedure. More precisely, for an increasing sequence of finite-dimensional projections $\{P_n\}_{n=1}^{\infty}$ with $P_n \to 1$ strongly, the limit $\lim_{n\to\infty} \mathcal{H}(P_nAP_n, P_nBP_n)$ is shown to exist and to be independent of the sequence of projections $\{P_n\}_{n=1}^{\infty}$. The question whether this sequence converges to its "obvious" limit, namely $\operatorname{Tr}[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)]$, has been left open. We answer this question in principle affirmatively and show that $\lim_{n\to\infty} \mathcal{H}(P_nAP_n, P_nBP_n) = \operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi(\alpha A + (1-\alpha)B)\,|_{\alpha=0}\right]$. If the operators A and B are regular enough, that is $(A-B)$, $\varphi(A) - \varphi(B)$ and $\varphi'(B)(A-B)$ are trace-class, the identity $\operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi(\alpha A + (1-\alpha)B)\,|_{\alpha=0}\right] = \operatorname{Tr}[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)]$ holds.

2.1 Introduction and main results

We start with a quick review of the setting and the results of [36] that are of interest for us. Let $\varphi \in C^0([0,1],\mathbb{R})$ be a continuous, convex function such that φ' is continously differentiable on $(0,1)$ and let A and B be two hermitian matrices with $0 \le A, B \le 1$. Lewin and Sabin define a family of relative entropies of A with respect to B by the formula

$$\mathcal{H}(A,B) = \operatorname{Tr}[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)]\,. \tag{2.1}$$

As long as $0 < B < 1$, the above expression is well defined. If 0 and/or 1 are contained in the spectrum of B and if φ is not differentiable at these points this is still true if $A = B$ on $\operatorname{Ker}(B)$, $\operatorname{Ker}(1-B)$ or $\operatorname{Ker}(B) \oplus \operatorname{Ker}(1-B)$, respectively (the trace is taken on the complement of these subspaces). Are the just mentioned conditions not fulfilled, they define $\mathcal{H}(A,B) = \infty$.

In [36, Theorem 1], the authors show that the so-defined relative entropy is monotone if and only if φ' is operator monotone. We quote:

Theorem 2.1 *(Monotonicity). Under the above conditions, the following are equivalent*

1. *φ' is operator monotone on $(0,1)$;*

2. *For any linear map $X : h_1 \to h_2$ on finite-dimensional spaces h_1 and h_2 with $X^*X \leq 1$, and for any $0 \leq A, B \leq 1$ on h_1, we have*

$$\mathcal{H}(XAX^*, XBX^*) \leq \mathcal{H}(A, B), \tag{2.2}$$

with $\mathcal{H}(A,B)$ defined in Eq. (2.1).

In a second step, this result is used to extend the definition of the relative entropy to self-adjoint operators acting on an infinite-dimensional separable Hilbert space h via the formula

$$\mathcal{H}(A, B) := \lim_{n\to\infty} \mathcal{H}(P_nAP_n, P_nBP_n), \tag{2.3}$$

where $\{P_n\}_{n=1}^{\infty}$ is an increasing sequence of finite-dimensional projections with $P_n \to 1$ in the strong operator topology. By $\mathcal{L}(h)$ we denote the set of bounded linear operators on h and h_1, h_2 denote infinite-dimensional separable Hilbert spaces. We quote again:

Theorem 2.2 *(Generalized relative entropy in infinite dimension). We assume that $\varphi \in C^0([0,1], \mathbb{R})$ and that φ' is operator monotone on $(0,1)$.*

1. *($\mathcal{H}$ is well defined). For an increasing sequence P_n of finite-dimensional projections on h such that $P_n \to 1$ strongly, the sequence $\mathcal{H}(P_nAP_n, P_nBP_n)$ is monotone and possesses a limit in $\mathbb{R}^+ \cup \{+\infty\}$. This limit does not depend on the chosen sequence P_n and hence $\mathcal{H}(A,B)$ is well-defined in $\mathbb{R}^+ \cup \{+\infty\}$.*

2. *(Approximation). If $X_n : h_1 \to h_2$ is a sequence such that $X_n^*X_n \leq 1$ and $X_n^*X_n \to 1$ strongly in h_1, then*

$$\mathcal{H}(A, B) = \lim_{n\to\infty} \mathcal{H}(X_nAX_n, X_nBX_n). \tag{2.4}$$

3. *(Weak lower semi-continuity). The relative entropy is weakly lower semi-continuous: if $0 \leq A_n, B_n \leq 1$ are two sequences such that $A_n \rightharpoonup A$ and $B_n \rightharpoonup B$ weakly-$*$ in $\mathcal{L}(h)$, then*

$$\mathcal{H}(A, B) \leq \liminf_{n\to\infty} \mathcal{H}(A_n, B_n). \tag{2.5}$$

As one would expect, $\mathcal{H}(A,B)$ can take finite values when A and B themselves are not compact, as the following upper bound shows [36, Theorem 3]:

$$\mathcal{H}(A, B) \leq C \operatorname{Tr}\left(\frac{1}{B^2} + \frac{1}{(1-B)^2}\right)(A-B)^2. \tag{2.6}$$

Note that the only dependence on φ on the right hand side of Eq. (2.6) is in the constant C. The question whether their notion of relative entropy in infinite dimensions is related to $\mathrm{Tr}\,[\varphi(A)-\varphi(B)-\varphi'(B)(A-B)]$, which is a-priori well-defined when the operator under the trace is trace-class, has been left open by the authors.

We answer this question in principle affirmatively, where "in principle" stands for the fact that $\mathrm{Tr}\,[\varphi(A)-\varphi(B)-\varphi'(B)(A-B)]$ turns out not to be the correct limit, in general.

Theorem 2.3 *Let $\varphi \in C^0([0,1],\mathbb{R})$ be such that φ' is operator monotone on (0,1) and let $\{P_n\}_{n=1}^{\infty}$ be defined as in Theorem 2.2. Then*

$$\lim_{n\to\infty} \mathcal{H}(P_nAP_n, P_nBP_n) = \mathrm{Tr}\left[\varphi(A)-\varphi(B)-\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)\Big|_{\alpha=0}\right], \quad (2.7)$$

with the understanding that either both sides are finite and equal each other, or both sides are infinite.

Remark 2.1 *We define the differential in Eq. (2.7) by the formula*

$$\frac{d}{d\alpha}\left(\psi,\varphi\left(\alpha A+(1-\alpha)B\right)\psi\right)\Big|_{\alpha=0} = \lim_{\alpha\to 0}\alpha^{-1}\left(\psi,\left[\varphi(\alpha A+(1-\alpha)B)-\varphi(B)\right]\psi\right). \quad (2.8)$$

In case φ' is continuous on $[0,1]$, this limit exists for all $\psi \in h$. If φ' is not continuous on the whole interval, it has singularities at 0 and/or 1 (we remind that φ' is monotone increasing and continuous on $(0,1)$ by assumption) and we have to distinguish between three cases. First, assume B has no eigenvalues at the points of discontinuity of φ'. Then the above limit exists for all ψ in a suitably chosen dense set $D \subset h$ (see Section 2.2, Lemma 2.3 for more details). Second, if B has an eigenvalue at a point of discontinuity of φ', ψ is the corresponding eigenvector to the just mentioned eigenvalue and $(A-B)\psi \neq 0$, then the above limit equals $-\infty$. Third, if $(A-B)\psi = 0$ in the just mentioned situation, the above limit is equal to zero.

Remark 2.2 *The operator monotonicity of the function φ' implies the operator convexity of its primitives which in turn implies that $\varphi(A)-\varphi(B)-\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ is positive (see Section 2.2 for more details). This property can now be used to define a notion of trace that is applicable on the right hand side of Eq. (2.7). Assume for the moment that B has no eigenvalues at points of discontinuity of φ' with $(A-B)\neq 0$ on the corresponding eigenspaces. Then the symmetric operator $\varphi(A)-\varphi(B)-\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ can be defined on the dense set D mentioned in Remark 2.1 and since it is positive it has a Friedrichs extension $(T,\mathcal{D}(T))$. By restricting attention to bases $\{e_\beta\}_{\beta=1}^{\infty}$ where all e_β lie in the form domain of T, we can define the trace of the operator on the right hand side of Eq. (2.7) to be $\sum_{\beta=1}^{\infty}(e_\beta, Te_\beta)$, see Section 2.2 for more details. The so-defined trace equals the usual trace whenever $\varphi(A)-\varphi(B)-\frac{d}{d\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ is trace class and $+\infty$ otherwise. Now if B has an eigenvalue at a point of discontinuity of φ' and $(A-B)\neq 0$ on the corresponding eigenspace Remark 2.1 suggest to define the trace on the right hand side of Eq. (2.7)* *to be $+\infty$. This goes hand in hand with the definition of the relative entropy for hermitian matrices of Lewin and Sabin which has been explained in the beginning of the introduction.*

Remark 2.3 *The idea to define the relative entropy as a trace over a manifestly positive operator in order to make it well-defined on a larger set, has already been used in [40]. The formula in the just mentioned reference equals the trace over $\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)|_{\alpha=0}$ with $\varphi(x) = x\ln(x) + (1-x)\ln(1-x)$ and resembles Eq. (2.9).*

Remark 2.4 *Assuming matrices, the equality $\operatorname{Tr}\left[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)\right] = \operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)|_{\alpha=0}\right]$ holds as one can see with a direct computation that exploits the cyclicity of the trace (see [38, Theorem V.3.3] for a simple way to compute the derivative). In general however, this cannot be expected.*

A crucial ingredient of our proof of Theorem 2.3 is the following Lemma which we state here because it is of interest in itself.

Lemma 2.1 *Let $\varphi \in C\left([0,1],\mathbb{R}\right)$ be such that φ' is operator monotone on $(0,1)$. Assume further that B has no eigenvalues at points of discontinuity of φ' with $(A-B) \neq 0$ on the corresponding eigenspaces. Then there exists a constant $b \geq 0$ and a unique Borel probability measure μ on $[-1,1]$ such that*

$$\operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi(\alpha A + (1-\alpha)B)\Big|_{\alpha=0}\right] = \tag{2.9}$$

$$2b\int_{-1}^{1}\int_{0}^{\infty}\operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2A)+t}Q\frac{1}{1+\lambda(1-2B)+t}\right]dt\ d\mu(\lambda),$$

where $Q = (A-B)$. (To be precise, μ is unique only if $b > 0$.)

Remark 2.5 *The formula on the right hand side of Eq. (2.9) is in many circumstances easier to handle than the formula on the left hand side. This is because it is the integral (with a positive measure) of a positive function which is the trace of a bounded positive operator. In particular, the operator under the trace has a simpler form than the one on the left hand side of Eq. (2.9).*

Assuming more regular operators A and B, the equality mentioned in Remark 2.4 is still true in infinite dimensions as the following statement shows.

Theorem 2.4 *Let $\varphi \in C^0([0,1],\mathbb{R})$ be such that φ' is operator monotone on (0,1). Assume in addition that $(A-B)$, $\varphi(A) - \varphi(B)$ and $\varphi'(B)(A-B)$ are trace-class. Then $\frac{d}{d\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)|_{\alpha=0}$ is trace-class and the identity*

$$\begin{aligned}\operatorname{Tr}\left[\varphi(A) - \varphi(B) - \frac{d}{d\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)\Big|_{\alpha=0}\right]& \\ = \operatorname{Tr}\left[\varphi(A) - \varphi(B) - \varphi'(B)(A-B)\right]&\end{aligned} \tag{2.10}$$

holds.

Remark 2.6 *In mathematical physics one encounters applications where the state of a physical system is defined to be a minimizer of a nonlinear functional in which the physical relative entropy appears, see e.g. [40, 28]. For a fermionic many-particle system the function* $\varphi(x) = x\ln(x) + (1-x)\ln(1-x)$ *is the right choice to define the physical relative entropy, while for bosons it is* $\varphi(x) = x\ln(x) - (1+x)\ln(1+x)$. *We note that both functions fulfill the requirements of Theorems 2.2-2.4. Since the right hand side of Eq.* (2.10) *is in practice much easier to evaluate explicitly than the left hand side when given a trial state* A, *Theorem 2.4 becomes important if one wants to derive an upper bound for the minimal energy of such a functional. The left hand side of Eq.* (2.7) *in contrast is important since it allows one to prove upper or lower bounds for the relative entropy with the help of Klein's inequality, see [36].*

2.2 Proof of Theorem 2.3

The main ingredient of our proof is the derivation of the formula stated in Lemma 2.1. Having this identity at hand, we show the convergence of the relative entropy by first showing it for the trace under the integral. In a second step, we argue why the limit can be interchanged with the integrals over λ and t. At this point Theorem 2.1 enters the analysis in a crucial way.

Since φ' is operator monotone on $(0,1)$ there exists a unique Borel probability measure μ on $[-1,1]$ such that (see [38, Corollary V.4.5])

$$\varphi'(x) = a + b\int_{-1}^{1} \frac{2x-1}{1-\lambda(2x-1)}\mathrm{d}\mu(\lambda), \tag{2.11}$$

with $b \geq 0$ (To be precise, μ is unique only if $b > 0$.). When integrating the above expression, one obtains a primitive for φ' which is of the form

$$\varphi(x) = ax + c - \frac{b}{2}\int_{-1}^{1}\left(\frac{2x-1}{\lambda} + \frac{\ln\left(1+\lambda(1-2x)\right)}{\lambda^2}\right)\mathrm{d}\mu(\lambda). \tag{2.12}$$

Since $x \mapsto -\ln(x)$ is an operator convex function, the same holds true for φ.

To keep the main argumentation straight, we first prove two technical Lemmata. The first concerns the relation between the regularity of φ at the endpoints of the interval $[0,1]$ and the behavior of the measure μ in the vicinity of -1 and 1.

Lemma 2.2 *Assume* $\varphi \in C^0([0,1],\mathbb{R})$ *such that* φ' *is operator monotone on* $(0,1)$. *Then* $\mu(\{-1\}) = 0 = \mu(\{1\})$,

$$\int_{1/2}^{1} -\ln(1-\lambda)\,d\mu(\lambda) < \infty \qquad \textit{and} \qquad \int_{-1}^{-1/2} -\ln(1+\lambda)\,d\mu(\lambda) < \infty. \tag{2.13}$$

If in addition $\varphi' \in C^0([0,1],\mathbb{R})$, the stronger implications

$$\int_{1/2}^{1} \frac{1}{1-\lambda} d\mu(\lambda) < \infty \qquad \textit{and} \qquad \int_{-1}^{-1/2} \frac{1}{1+\lambda} d\mu(\lambda) < \infty \tag{2.14}$$

hold. In case φ' is not continuous at 1 the first integral in Eq. (2.14) equals $+\infty$ and if it is not continuous at 0 this is true for the second integral.

Proof. We start with the first case, hence we assume that only φ is continuous on $[0,1]$. Since the limits $\lim_{x\to 0}\varphi(x)$ and $\lim_{x\to 1}\varphi(x)$ exist we can conclude that $\mu(\{-1\}) = 0 = \mu(\{1\})$ holds. We further conclude that the following limit exists (see Eq. (2.12))

$$\infty > \lim_{x\to 1} \int_{1/2}^{1} -\frac{\ln(1+\lambda(1-2x))}{\lambda^2} \mathrm{d}\mu(\lambda) \geq \frac{1}{4}\int_{1/2}^{1} -\ln(1-\lambda)\mathrm{d}\mu(\lambda). \tag{2.15}$$

To come to the expression on the right hand side, we have applied Fatou's Lemma. Doing the same argumentation again, this time with the limit $x \to 0$, yields $\int_{-1}^{-1/2} \ln(1+\lambda)\mathrm{d}\mu(\lambda) < \infty$. If also φ' is continuous on $[0,1]$, we compute

$$\lim_{x\to 1} \varphi'(x) = a + b \lim_{x\to 1} \int_{-1}^{1} \frac{2x-1}{1-\lambda(2x-1)} \mathrm{d}\mu(\lambda) \geq a + \int_{-1}^{1} \frac{1}{1-\lambda} \mathrm{d}\mu(\lambda), \tag{2.16}$$

where, as before, we have applied Fatou's Lemma. The same procedure with $\lim_{x\to 0}\varphi'(x)$ yields the other bound. Using the monotonicity of the integrand, one easily shows that the just discussed integrals diverge to $+\infty$ in case φ' is not continuous at 0 and/or 1, respectively. □

In order to obtain a handy formula for the operator $\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)|_{\alpha=0}$, we explicitly compute the directional derivative.

Lemma 2.3 *Assume $\varphi \in C^0([0,1],\mathbb{R})$ such that φ' is operator monotone on $(0,1)$. If φ' is continuous on $[0,1]$ or if it is discontinuous at 0 and/or 1 and B has no eigenvalue at these points then*

$$\begin{aligned} \frac{d}{d\alpha}\left(\psi, \varphi(\alpha A + (1-\alpha)B)\psi\right)\Big|_{\alpha=0} &= a\left(\psi,(A-B)\psi\right) - \frac{b}{2}\int_{-1}^{1}\left[\left(\psi, \frac{2(A-B)}{\lambda}\psi\right)\right. \\ - \frac{2}{\lambda}\int_0^\infty &\left.\left(\psi, \frac{1}{1+\lambda(1-2B)+t}(A-B)\frac{1}{1+\lambda(1-2B)+t}\psi\right) dt\right] d\mu(\lambda), \end{aligned} \tag{2.17}$$

where a,b and μ are defined by Eq. (2.11). In case of the first scenario (φ' continuous on $[0,1]$), the derivative is taken for all $\psi \in h$ while in the second scenario it is taken only for all ψ in a dense set $D \subset h$. Explicitly, the set D is given by $D = \cup_{\epsilon>0}\mathbb{1}(\epsilon < B < 1-\epsilon)h$ in case 0 and 1 are points of discontinuity of φ' and by the obvious *generalization when φ' is discontinuous only at one of these points. This accounts for the fact that the limiting operator may be unbounded. In case φ' has discontinuities and B has eigenvalues at at least one of these points, we have to treat the above limit with ψ being one of the eigenvectors to the just mentioned eigenvalues separately. We distinguish between two cases. If $(A-B)\psi \neq 0$ we have*

$$-\lim_{\alpha\to 0}\left(\psi, \left[\frac{\varphi(\alpha A+(1-\alpha B))-\varphi(B)}{\alpha}\right]\psi\right) = \infty. \tag{2.18}$$

If $(A-B)\psi = 0$ instead, the limit in Eq. (2.18) equals zero.

Proof. Using Eq. (2.12), one can easily check the identity

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\psi, \varphi\left(\alpha A+(1-\alpha)B\right)\psi\right)\Big|_{\alpha=0} = a\left(\psi,(A-B)\psi\right) - \frac{b}{2}\lim_{\alpha\to 0}\int_{-1}^{1}\left[\left(\psi, \frac{2(A-B)}{\lambda}\psi\right)\right.$$
$$\left. + \left(\psi, \frac{\ln\left(1+\lambda(1-2(B+\alpha(A-B)))\right)-\ln\left(1+\lambda(1-2B)\right)}{\alpha\lambda^2}\psi\right)\right]\mathrm{d}\mu(\lambda). \tag{2.19}$$

The second term is just the difference quotient defining the directional derivative of the second term in Eq. (2.12). Let us have a closer look at the term with the logarithms. We use the formula $\ln(x) = \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t}\right)\mathrm{d}t$ and apply the resolvent identity once, to see that it can be written as

$$\left(\psi, \frac{\ln\left(1+\lambda(1-2(B+\alpha(A-B)))\right)-\ln\left(1+\lambda(1-2B)\right)}{\alpha\lambda^2}\psi\right) = \tag{2.20}$$
$$-\frac{2}{\lambda}\int_0^\infty\left(\psi, \frac{1}{1+\lambda(1-2B)+t}(A-B)\frac{1}{1+\lambda(1-2(B+\alpha(A-B)))+t}\psi\right)\mathrm{d}t.$$

In order to explicitly compute the limit $\alpha\to 0$, it needs to be interchanged in a first step with the integral over λ and in a second step with the integral over t. The second step will follow easily from the estimates used to show the first step since $\lambda\in(-1,1)$ is then fixed which implies that all resolvents are uniformly bounded. We therefore focus on the interchange of the limit $\alpha\to 0$ with the integral over λ. In order to be able to apply dominated convergence, we have to find a positive function $g\in L^1(\mu)$ with

$$\left|\left(\psi, \frac{2(A-B)}{\lambda}\psi\right) - \frac{2}{\lambda}\int_0^\infty\left(\psi, R(B)(A-B)R(B+\alpha(A-B))\psi\right)\mathrm{d}t\right| \leq g(\lambda) \tag{2.21}$$

for all ψ at least in a dense subset of h (The case where B has eigenvalues at points of discontinuity of φ' will be treated at the end.). To shorten the writing, we have introduced the notation $R(B) = (1+\lambda(1-2B)+t)^{-1}$.

Let us first investigate the behavior of our integrand for $\lambda\in(-1+\epsilon, 1-\epsilon)$. We write $R(B) = \frac{1}{1+t} - \frac{\lambda}{1+t}(1-2B)R(B)$ (and the same for $R(B+\alpha(A-B))$) and evaluate the contribution of the first term which reads

$$-\frac{2}{\lambda}\int_0^\infty \frac{1}{1+t}\left(\psi,(A-B)\psi\right)\frac{1}{1+t}\mathrm{d}t = -\frac{2}{\lambda}\left(\psi,(A-B)\psi\right). \tag{2.22}$$

It cancels the first term under the integral on the right hand side of Eq. (2.19). The three remaining terms have no singularity and can be bounded by a constant.

In the vicinity of $\lambda=-1$ and $\lambda=1$ the situation is a little different and one needs to argue more carefully. We will distinguish three cases depending on the regularity of φ' at 0 and 1 and on the spectrum of B. First let us assume that φ' is not continuous at 0 and 1 and that B has no eigenvalues at these points. Let $D_\epsilon = \mathbb{1}\left(\epsilon < B < 1-\epsilon\right)h$ and define $D=\cup_{\epsilon>0}D_\epsilon$. Due to our assumptions on B, the set D is dense in h. For $\psi\in D$, we investigate

$$\begin{aligned}\int_0^\infty &\left|\left(\psi,\frac{1}{1+\lambda(1-2B)+t}(A-B)\frac{1}{1+\lambda(1-2(\alpha A+(1-\alpha)B))+t}\psi\right)\right|\mathrm{d}t \quad (2.23)\\ &\le \int_0^\infty \left\|\frac{1}{1+\lambda(1-2B)+t}\psi\right\|\,\|A-B\|\,\left\|\frac{1}{1+\lambda(1-2(\alpha A+(1-\alpha)B))+t}\psi\right\|\mathrm{d}t\end{aligned}$$

which is the relevant contribution from Eq. (2.21). The part of the integral over t from say 1 to ∞ is easy to control. One just bounds the resolvents in operator norm by $1/t$. After the evaluation of the integral, we end up with a constant. To bound the other part of the integral over t (the one from 0 to 1), we use the fact that $\psi\in D$ which implies that $\left\|\frac{1}{1+\lambda(1-2B)+t}\psi\right\|\le 1/\epsilon$ for an $\epsilon>0$ that depends on ψ. On the other hand $\left\|\frac{1}{1+\lambda(1-2(\alpha A+(1-\alpha)B))+t}\psi\right\|\le\frac{1}{1+\lambda+t}$ for λ close to -1. Putting this together, we obtain

$$\begin{aligned}\int_0^\infty \left\|\frac{1}{1+\lambda(1-2B)+t}\psi\right\|\left\|A-B\right\|_\infty\left\|\frac{1}{1+\lambda(1-2(\alpha A+(1-\alpha)B))+t}\psi\right\|\mathrm{d}t&\\ \le \|A-B\|\left(\frac{1}{\epsilon}\int_0^1\frac{1}{1+\lambda+t}\mathrm{d}t + C\right)& \quad (2.24)\\ \le C(\epsilon)\left(-\ln(1+\lambda)+1\right).&\end{aligned}$$

A similar bound can be obtained for λ close to 1. There the function $-\ln(1-\lambda)$ enters the analysis. Hence, there exists a constant $C(\epsilon)$ depending on ψ such that

$$\begin{aligned}\left|\left(\psi,\frac{2(A-B)}{\lambda}\psi\right)-\frac{2}{\lambda}\int_0^\infty\left(\psi,R(B)(A-B)R(B+\alpha(A-B))\psi\right)\mathrm{d}t\right|& \quad (2.25)\\ \le C(\epsilon)\left(-\ln(1-|\lambda|)+1\right).&\end{aligned}$$

Because of Lemma 2.2, the bound allows us to take the limit inside the integral and proves the claim in this situation.

Nearly the same argumentation goes through when B has spectrum at 0 and/or 1 and if φ' is continuous at these points. By bounding both resolvents like we did with the second in the previous step, that is $\|R(B)\psi\| \leq (1-|\lambda|+t)^{-1}$ and the same with $\|R(\alpha A+(1-\alpha)B)\psi\|$, one obtains

$$\left|\left(\psi, \frac{2(A-B)}{\lambda}\psi\right) - \frac{2}{\lambda}\int_0^\infty \Big(\psi, R(B)(A-B)R(B+\alpha(A-B))\psi\Big)\,\mathrm{d}t\right| \leq \frac{C}{1-|\lambda|}. \tag{2.26}$$

Again due to Lemma 2.2, this is enough to interchange the limit and the integral. The case where φ' is discontinuous only at one point is treated in the obvious way.

For the last case we have to assume that φ' is not continuous at 0 and/or 1 and that B has an eigenvalue at at least one of these points. We only investigate the relevant contribution. Let ψ be the eigenvector of B to the eigenvalue 0 for example (the other cases go the same way). We will show that

$$\lim_{\alpha\to 0}\int_{-1}^{-1/2} \frac{-1}{\lambda}\int_0^\infty (\psi, R(B)(A-B)R(B+\alpha(A-B))\psi)\,\mathrm{d}t\;\mathrm{d}\mu(\lambda) = \infty, \tag{2.27}$$

if $(A-B)\psi \neq 0$ and that the above limit equals zero in case $(A-B)\psi = 0$. Using $R(B+\alpha(A-B)) = R(B) + 2\alpha\lambda R(B+\alpha(A-B))(A-B)R(B)$, the integrand can be written as

$$\begin{aligned} &\frac{-1}{\lambda}(\psi, R(B)(A-B)R(B)\psi) \\ &\qquad - 2\alpha\,(\psi, R(B)(A-B)R(B+\alpha(A-B))(A-B)R(B)\psi) \\ &\qquad = \left(\frac{1}{1+\lambda+t}\right)^2\left[\frac{-1}{\lambda}(\psi, A\psi) - 2\alpha\,(\psi, AR(B+\alpha(A-B))A\psi)\right]. \end{aligned} \tag{2.28}$$

Let us first assume that $(A-B)\psi \neq 0$ which implies that $(\psi, A\psi) > 0$. Since the function $t \mapsto \frac{1}{t}$ is operator convex on the interval $(0,\infty)$, see [38, Exercise V.2.11], we know that $R(\alpha A+(1-\alpha)B) \leq \alpha R(A) + (1-\alpha)R(B)$. If we apply this inequality on the right hand side of Eq. (2.28) and discard all positive terms in order to obtain a lower bound, we find

$$\begin{aligned} &\left(\frac{1}{1+\lambda+t}\right)^2\left[\frac{-1}{\lambda}(\psi, A\psi) - 2\alpha\,(\psi, AR(B+\alpha(A-B))A\psi)\right] \\ &\quad \geq \left(\frac{1}{1+\lambda+t}\right)^2\left[\frac{-1}{\lambda}(\psi, A\psi) - 2\alpha^2\,(\psi, AR(A)A\psi) - 2\alpha\,(\psi, AR(B)A\psi)\right]. \end{aligned} \tag{2.29}$$

The right hand side of this equation, viewed as a function of α, is certainly monotone and so we can use monotone convergence to show that

$$\lim_{\alpha\to 0}\int_{-1}^{-1/2}\int_0^\infty \left(\frac{1}{1+\lambda+t}\right)^2\left[\frac{-1}{\lambda}(\psi, A\psi) - 2\alpha^2(\psi, AR(A)A\psi)\right.$$
$$\left. - 2\alpha(\psi, AR(B)A\psi)\right]\mathrm{d}t\ \mathrm{d}\mu(\lambda) \tag{2.30}$$
$$= (\psi, A\psi)\int_{-1}^{-1/2}\frac{-1}{\lambda}\int_0^\infty\left(\frac{1}{1+\lambda+t}\right)^2\mathrm{d}t\ \mathrm{d}\mu(\lambda)$$
$$\geq (\psi, A\psi)\int_{-1}^{-1/2}\frac{1}{1+\lambda}\mathrm{d}\mu(\lambda) = \infty.$$

The last equality is achieved with the help of Lemma 2.2. Now assume that $(A-B)\psi = 0$ which means that $A\psi = 0$. Hence, $[\alpha A + (1-\alpha)B]\,\psi = 0$ and $\varphi(\alpha A+(1-\alpha)B)\psi = \varphi(0)\psi$. Since this expression is a constant the derivative with respect to α vanishes. A similar argument can be done when B has 1 as an eigenvalue. This concludes the proof of Lemma 2.3. □

Before we come to the main part of the proof, we have to argue how the trace on the right hand side of Eq. (2.7) can be defined. Let us for the moment assume that B has no eigenvalues at points of discontinuity of φ' with $(A-B)\neq 0$ on the corresponding eigenspaces. Then by Lemma 2.3, we can define the quadratic form

$$q(\psi,\eta) = \lim_{\alpha\to 0}\left(\psi, \left[\varphi(A) - \varphi(B) - \frac{\varphi(B+\alpha(A-B)) - \varphi(B)}{\alpha}\right]\eta\right) \tag{2.31}$$

on the dense set $D\subset h$ (The set D has been defined in Lemma 2.3.). The operator convexity of φ implies that $\frac{\varphi(B+\alpha(A-B))-\varphi(B)}{\alpha}\leq \varphi(A)-\varphi(B)$ holds for all $0<\alpha\leq 1$. Since the inequality is preserved by the limiting procedure $\alpha\to 0$ we conclude that q is positive. It is an easy exercise to check with the methods used in the proof of Lemma 2.3, that on D, the operator

$$\varphi(A) - \varphi(B) - a(A-B) \tag{2.32}$$
$$+\frac{b}{2}\int_{-1}^{1}\left[\frac{2(A-B)}{\lambda} - \frac{2}{\lambda}\int_0^\infty \frac{1}{1+\lambda(1-2B)+t}(A-B)\frac{1}{1+\lambda(1-2B)+t}\mathrm{d}t\right]\mathrm{d}\mu(\lambda),$$

is well-defined, symmetric and due to the previous reasoning also positive [compare with Eq. (2.17)]. Again by Lemma 2.3, its associated quadratic form is q. The theorem on the Friedrichs extension tells us that q is closable and that its closure $(\hat{q}, \mathcal{Q}(\hat{q}))$ is the quadratic form of a unique self-adjoint operator $(T, \mathcal{D}(T))$ whose domain $\mathcal{D}(T)$ is contained in the form domain $\mathcal{Q}(\hat{q})$ of $\hat{q}$, see [42, Theorem X.23]. Additionally, T is positive. Having the Friedrichs extension at hand, we can define the right hand side of Eq. (2.7) to be the trace of T. To that end, we restrict our attention to bases $\{e_\beta\}_{\beta=1}^\infty$ of h with $e_\beta\in\mathcal{Q}(\hat{q})$ for all $\beta\in\mathbb{N}$ and define $\mathrm{Tr}(T) = \sum_{\beta=1}^\infty \hat{q}(e_\beta, e_\beta)$. Of course, this definition does not depend on the choice of the basis. It yields the usual notion of trace when T is trace-class and gives $\mathrm{Tr}(T) = +\infty$ otherwise. If B has an eigenvalue at a point of discontinuity of φ' with $A - B \neq 0$ on the corresponding eigenspace, Lemma 2.3 suggest to define the trace of the right hand side of Eq. (2.7) to be $+\infty$. This goes hand in hand with the definition of Lewin and Sabin mentioned in the beginning of the introduction.

Having these prerequisites at hand, we come to the main part of our proof. If B has an eigenvalue at a point of discontinuity of φ' and $(A - B) \neq 0$ on the corresponding eigenspace then both sides of Eq. (2.7) equal $+\infty$. For the right hand side this has been discussed in the previous paragraph while for the left hand side, this can be seen by choosing P_1 such that the eigenspace of the just mentioned eigenvalue lies in its range. Hence, we can exclude this case in what follows. The key point of our proof is the more explicit formula Eq. (2.9) for the trace of the operator $\varphi(A) - \varphi(B) - \frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi(\alpha A + (1-\alpha)B)|_{\alpha=0}$ which we derive now. Using Eq. (2.12), the operator $\varphi(A) - \varphi(B)$ can be written as

$$\begin{aligned}\varphi(A) - \varphi(B) &= a(A - B) \\ &\quad - \frac{b}{2}\int_{-1}^{1}\left[\frac{2(A-B)}{\lambda} + \frac{\ln(1+\lambda(1-2A)) - \ln(1+\lambda(1-2B))}{\lambda^2}\right]\mathrm{d}\mu(\lambda).\end{aligned} \tag{2.33}$$

In the next step, we write the difference of the two logarithms in Eq. (2.33) with the help of the formula $\ln(x) = \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t}\right)\mathrm{d}t$ as an integral over resolvents. When we add the explicit representation for $\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi(\alpha A + (1-\alpha)B)|_{\alpha=0}$ that has been derived in Lemma 2.3 and apply the resolvent identity twice, we arrive at the formula

$$\begin{aligned}&\varphi(A) - \varphi(B) - \frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi(\alpha A + (1-\alpha)B)\big|_{\alpha=0} = \\ &\quad 2b\int_{-1}^{1}\int_0^\infty \frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2A)+t}Q\frac{1}{1+\lambda(1-2B)+t}\mathrm{d}t\ \mathrm{d}\mu(\lambda),\end{aligned} \tag{2.34}$$

where we have introduced the shortcut $Q = (A - B)$. Taking the trace on both sides, we can commute it with the integrals because the integrand is a positive operator and obtain Eq. (2.9). Hence, we have proved Lemma 2.1.

Now let $\{P_n\}_{n=1}^\infty$ be an increasing sequence of finite-dimensional projections that converges to 1 in the strong operator topology. Because for matrices the two ways of writing the relative entropy are the same (see Remark 2.4) we have the formula

$$\mathcal{H}(A_n, B_n) = 2b\int_{-1}^{1}\int_0^\infty \mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right]\mathrm{d}t\ \mathrm{d}\mu(\lambda) \tag{2.35}$$

with $A_n = P_nAP_n$ and so on. We will first show that

$$\lim_{n\to\infty}\mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right] = \mathrm{Tr}\left[R(B)QR(A)QR(B)\right] \tag{2.36}$$

and then argue why we can interchange the limit with the two integrals.

Let $m \geq 1$. In order to be able to restrict the trace on the right hand side of Eq. (2.35) to a finite-dimensional subspace, we first investigate

$$\begin{aligned}
&\operatorname{Tr}\left[(1-P_m)R(B_n)Q_nR(A_n)Q_nR(B_n)(1-P_m)\right] \\
&\qquad \leq \operatorname{Tr}\left[(1-P_m)R(B_n)Q_n^2R(B_n)(1-P_m)\right]\frac{1}{1-|\lambda|+t} \\
&\qquad \leq \operatorname{Tr}\left[(1-P_m)R(B_n)P_nQ^2P_nR(B_n)(1-P_m)\right]\frac{1}{1-|\lambda|+t}.
\end{aligned} \tag{2.37}$$

Let us for the moment assume that Q is Hilbert-Schmidt which implies that it can be written as $Q = \sum_{\beta=1}^{\infty} q_\beta |\psi_\beta\rangle\langle\psi_\beta|$ with $\sum_{\beta=1}^{\infty} q_\beta^2 < \infty$. The case when this does not hold true is taken care of at the end. Using the cyclicity of the trace, we write

$$\begin{aligned}
\operatorname{Tr}\left[(1-P_m)R(B_n)P_nQ^2P_nR(B_n)(1-P_m)\right] &= \operatorname{Tr}\left[QP_nR(B_n)(1-P_m)R(B_n)P_nQ\right] \\
&= \sum_{\alpha=1}^{k}(\psi_\alpha, QP_nR(B_n)(1-P_m)R(B_n)P_nQ\psi_\alpha) \\
&\quad + \sum_{\alpha=k+1}^{\infty}(\psi_\alpha, QP_nR(B_n)(1-P_m)R(B_n)P_nQ\psi_\alpha)\,.
\end{aligned} \tag{2.38}$$

The term in the last line on the right hand side of Eq. (2.38) can be bounded uniformly in n as the next calculation shows,

$$\begin{aligned}
&\left|\sum_{\alpha=k+1}^{\infty}\left(\psi_\alpha, QP_nR(B_n)(1-P_m)R(B_n)P_nQ\psi_\alpha\right)\right| \\
&\qquad = \left|\sum_{\alpha=k+1}^{\infty} q_\alpha^2\,(\psi_\alpha, P_nR(B_n)(1-P_m)R(B_n)P_n\psi_\alpha)\right| \\
&\qquad \leq \left(\frac{1}{1-|\lambda|+t}\right)^2 \sum_{\alpha=k+1}^{\infty} q_\alpha^2.
\end{aligned} \tag{2.39}$$

The right hand side of Eq. (2.39) goes to zero as k tends to infinity for all $-1 < \lambda < 1$ and $t \geq 0$ due to the assumptions on Q. On the other hand,

$$\begin{aligned}
&\sum_{\alpha=1}^{k}(\psi_\alpha, QP_nR(B_n)(1-P_m)R(B_n)P_nQ\psi_\alpha) \\
&\qquad \to \sum_{\alpha=1}^{k}(\psi_\alpha, QPR(B)(1-P_m)R(B)Q\psi_\alpha)
\end{aligned} \tag{2.40}$$

for $n \to \infty$ because the sum is finite and the operator in the middle is convergent in the strong operator topology, see [41, Theorem VIII.20]. When we consider Eq. (2.37) again and take the limit $n \to \infty$ followed by the limit $k \to \infty$, we arrive at

$$\lim_{n\to\infty} \mathrm{Tr}\left[(1-P_m)R(B_n)Q^2R(B_n)(1-P_m)\right] \\ = \mathrm{Tr}\left[(1-P_m)R(B)Q^2R(B)(1-P_m)\right]. \tag{2.41}$$

Let us denote the left hand side of this equation by $\delta(n,m)$ and the right hand side by $\delta(m)$. By construction, $\lim_{m\to\infty}\delta(m)=0$ holds. Using this result, we easily get the following two inequalities

$$\begin{aligned} \mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right] &\le \mathrm{Tr}\left[P_mR(B_n)Q_nR(A_n)Q_nR(B_n)P_m\right]+\tilde{\delta}(n,m), \\ \mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right] &\ge \mathrm{Tr}\left[P_mR(B_n)Q_nR(A_n)Q_nR(B_n)P_m\right], \end{aligned} \tag{2.42}$$

where $\tilde{\delta}(n,m)=\delta(n,m)(1-|\lambda|+t)^{-1}$. Taking first the limit $n\to\infty$ and then the limit $m\to\infty$ in the above equations, we conclude that

$$\lim_{n\to\infty} \mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right] = \mathrm{Tr}\left[R(B)QR(A)QR(B)\right] \tag{2.43}$$

for all $-1<\lambda<1$.

The next step in the proof is to interchange the limit $n\to\infty$ and the integrals. Let us start with the integral over t. Since we only need a bound for almost every λ to apply dominated convergence we can assume that $-1<\lambda<1$. Under these conditions a dominating function is easily constructed because

$$\mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right] \le \left(\frac{1}{1-|\lambda|+t}\right)^3 \|Q\|_2^2\,. \tag{2.44}$$

Hence, we have shown that

$$\int_{-1}^{1}\int_{0}^{\infty} \mathrm{Tr}\left[R(B)QR(A)QR(B)\right] \mathrm{d}t\mathrm{d}\mu(\lambda) \\ = \int_{-1}^{1} \lim_{n\to\infty}\left(\int_0^\infty \mathrm{Tr}\left[R(B_n)Q_nR(A_n)Q_nR(B_n)\right]\mathrm{d}t\right)\mathrm{d}\mu(\lambda). \tag{2.45}$$

To interchange the limit with the first integral, we have to argue more carefully and use the monotonicity of the relative entropy. With similar but somewhat easier arguments than the ones used to prove Lemma 2.3, we can show that

$$\frac{1}{\lambda^2}\Big(-\ln(1+\lambda(1-2A_n))+\ln(1+\lambda(1-2B_n)) \\ +\frac{\mathrm{d}}{\mathrm{d}\alpha}\ln(1+\lambda(1-2(\alpha A_n+(1-\alpha B_n))))|_{\alpha=0}\Big) \\ = \int_0^\infty R(B_n)Q_nR(A_n)Q_nR(B_n)\mathrm{d}t. \tag{2.46}$$

Now we take the trace on both sides of the above equation. On the right hand side, we interchange the trace with the integral over t and use the result from Eq. (2.45) to arrive at

$$\begin{aligned}\int_{-1}^{1}\int_{0}^{\infty} \mathrm{Tr}\left[R(B)QR(A)QR(B)\right] \mathrm{d}t\mathrm{d}\mu(\lambda) \qquad (2.47)\\ = \int_{-1}^{1} \frac{1}{\lambda^2}\Big(\lim_{n\to\infty} \mathrm{Tr}\Big[-\ln(1+\lambda(1-2A_n)) + \ln(1+\lambda(1-2B_n)) \\ + \frac{\mathrm{d}}{\mathrm{d}\alpha} \ln(1+\lambda(1-2(\alpha A_n + (1-\alpha B_n))))|_{\alpha=0}\Big]\Big) \mathrm{d}\mu(\lambda).\end{aligned}$$

Since $x \mapsto (-\ln(x))' = -\frac{1}{x}$ is operator monotone, the integrand on the right hand side of Eq. (2.47) is monotone in n by Theorem 2.1. On the other hand, from what we said above, we know that it converges pointwise for all $-1 < \lambda < 1$ as n tends to infinity. Therefore, the interchange of the limit $n \to \infty$ and the integral over λ is justified by monotone convergence. This completes the proof for the case when $Q = (A - B)$ is Hilbert-Schmidt.

Now assume that $(A - B)$ is not Hilbert-Schmidt. From [36, Theorem 3], we conclude that there is a constant $C > 0$ such that

$$\lim_{n\to\infty} \mathcal{H}(A_n, B_n) \geq C\, \|(A-B)\|_2^2 = \infty. \qquad (2.48)$$

On the other hand

$$\mathrm{Tr}\left[R(B)QR(A)QR(B)\right] \geq \mathrm{Tr}\left[R(B)Q^2R(B)\right] \frac{1}{4+t} = \infty, \qquad (2.49)$$

where the equality on the right hand side is justified by the fact that $R(B)$ is bounded and invertible for all $-1 < \lambda < 1$. Hence, the right hand side of Eq. (2.7) equals $+\infty$ as well. This completes the proof of Theorem 2.3.

2.3 Proof of Theorem 2.4

As in the proof of Theorem 2.3, we start with a Lemma in order not to interrupt the main argumentation. Throughout the whole section we assume that the b in Eq. (2.11) is strictly positive. This is reasonable because otherwise the relative entropy equals zero.

Lemma 2.4 *Assume that $(A - B)$ and $\varphi'(B)(A - B)$ are trace-class. Then*

$$\sum_{\beta=1}^{\infty} \int_{-1}^{1} |q_\beta| \left(\psi_\beta, \left|\frac{2B-1}{1-\lambda(2B-1)}\right| \psi_\beta\right) d\mu(\lambda) < \infty, \qquad (2.50)$$

where $(A-B) = \sum_{\beta=1}^{\infty} q_\beta |\psi_\beta\rangle\langle\psi_\beta|$ with $\sum_{\beta=1}^{\infty} |q_\beta| < \infty$.

Proof. The integral representation of φ', Eq. (2.11), tells us that

$$\varphi'(B)(A-B) = a(A-B) + b\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)} \mathrm{d}\mu(\lambda)(A-B). \tag{2.51}$$

Because $(A-B)$ is trace-class by assumption we know that the second term on the right hand side of the above equation is trace-class as well. And due to the polar decomposition, there exist two partial isometries U and V such that

$$\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)} \mathrm{d}\mu(\lambda)(A-B) = U\left|\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)} \mathrm{d}\mu(\lambda)\right| |A-B|\, V. \tag{2.52}$$

Since the set of all trace-class operators is a two-sided ideal in the algebra of bounded operators $\mathcal{L}(h)$ we conclude that the term on the right hand side of Eq. (2.52) without U and V is trace-class as well. We decompose the operator B in the way $B = B\,\mathbb{1}(B < 1/2) + B\,\mathbb{1}(B \geq 1/2)$ to see that the absolute value of the integral on the right hand side of Eq. (2.52) is given by $\left|\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)} \mathrm{d}\mu(\lambda)\right| = \int_{-1}^{1} \left|\frac{2B-1}{1-\lambda(2B-1)}\right| \mathrm{d}\mu(\lambda)$. Therefore,

$$\begin{aligned}\infty > \mathrm{Tr}\left|\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)} \mathrm{d}\mu(\lambda)\right| |A-B| \\ = \sum_{\beta=1}^{\infty} \int_{-1}^{1} |q_\beta| \left(\psi_\beta, \left|\frac{2B-1}{1-\lambda(2B-1)}\right| \psi_\beta\right) \mathrm{d}\mu(\lambda).\end{aligned} \tag{2.53}$$

This is what we intended to show. □

Having Lemma 2.4 at hand, the proof of Theorem 2.4, that is the proof of the identity $\mathrm{Tr}\left[\varphi'(B)(A-B)\right] = \mathrm{Tr}\left[\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi\left(\alpha A + (1-\alpha)B\right)\big|_{\alpha=0}\right]$, is in principle a straightforward computation that exploits the cyclicity of the trace. We start by inserting the integral representation of φ' [Eq. (2.11)] into $\mathrm{Tr}\left[\varphi'(B)(A-B)\right]$ to obtain

$$\begin{aligned}\mathrm{Tr}\left[\varphi'(B)(A-B)\right] &= a\,\mathrm{Tr}(A-B) + b\,\mathrm{Tr}\left[\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)}(A-B)\mathrm{d}\mu(\lambda)\right] \\ &= a\,\mathrm{Tr}(A-B) + b\sum_{\beta=1}^{\infty} \int_{-1}^{1} q_\beta \left(\psi_\beta, \frac{2B-1}{1-\lambda(2B-1)}\psi_\beta\right) \mathrm{d}\mu(\lambda).\end{aligned} \tag{2.54}$$

Here, $\{\psi_\beta\}_{\beta=1}^{\infty}$ denotes the complete set of eigenfunctions of the self-adjoint operator $(A-B)$. We wish to interchange the sum over β and the integral over λ on the right hand side of the above equation. Using the bound

$$\left|q_\beta\left(\psi_\beta, \frac{2B-1}{1-\lambda(2B-1)}\psi_\beta\right)\right| \leq |q_\beta| \left(\psi_\beta, \left|\frac{2B-1}{1-\lambda(2B-1)}\right| \psi_\beta\right) \tag{2.55}$$

and Lemma 2.4, this is justified by an application of Fubini's theorem. On the other hand, the operator $\frac{2B-1}{1-\lambda(2B-1)}(A-B)$ is trace-class as long as $-1 < \lambda < 1$ because

$(A-B)$ is trace-class and $\frac{2B-1}{1-\lambda(2B-1)}$ is bounded. We conclude that

$$\operatorname{Tr}\left[\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)}(A-B)\mathrm{d}\mu(\lambda)\right] \\ = \int_{-1}^{1} \operatorname{Tr}\left[\frac{2B-1}{1-\lambda(2B-1)}(A-B)\right] \mathrm{d}\mu(\lambda). \tag{2.56}$$

Using the identity

$$\frac{1-2B}{1+\lambda(1-2B)} = \frac{1}{\lambda} - \frac{1}{\lambda}\int_0^\infty \left(\frac{1}{1+\lambda(1-2B)+t}\right)^2 \mathrm{d}t, \tag{2.57}$$

Eq. (2.56) can be written as

$$\operatorname{Tr}\left[\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)}(A-B)\mathrm{d}\mu(\lambda)\right] = \\ = -\frac{1}{2}\int_{-1}^{1} \operatorname{Tr}\left[\left\{\frac{2}{\lambda} - \frac{2}{\lambda}\int_0^\infty \left(\frac{1}{1+\lambda(1-2B)+t}\right)^2 \mathrm{d}t\right\}(A-B)\right] \mathrm{d}\mu(\lambda). \tag{2.58}$$

With the bound

$$\left|\sum_{\beta=1}^{\infty} \left(\psi_\beta, \left(\frac{1}{1+\lambda(1-2B)+t}\right)^2 (A-B)\psi_\beta\right)\right| \le \left(\frac{1}{1-|\lambda|+t}\right)^2 \|A-B\|_1 \tag{2.59}$$

which holds for all $-1<\lambda<1$, we argue like above with Fubini that the trace can be interchanged with the integral over t. Now we can use the cyclicity of the trace to arrive at

$$\operatorname{Tr}\left[b\int_{-1}^{1} \frac{2B-1}{1-\lambda(2B-1)}Q\mathrm{d}\mu(\lambda)\right] = \\ = \frac{-b}{2}\int_{-1}^{1} \left(\frac{2}{\lambda}\operatorname{Tr} Q - \frac{2}{\lambda}\int_0^\infty \operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2B)+t}\right] \mathrm{d}t\right) \mathrm{d}\mu(\lambda). \tag{2.60}$$

To shorten the writing, we have used the shortcut $Q=(A-B)$. Except for the fact that the trace is inside the integral, this is what we wanted to obtain (compare with the result of Lemma 2.3).

Now we have to argue why we can take the trace out of the integral again which would complete the proof. By Q_+ and Q_- we denote the positive and the negative part of the operator $Q=(A-B)$, respectively. First, we want to show that the above term with Q replaced by Q_+ or by Q_-, that is

$$\int_{-1}^{1} \left(\frac{2}{\lambda}\operatorname{Tr} Q_\pm - \frac{2}{\lambda}\int_0^\infty \operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t}Q_\pm\frac{1}{1+\lambda(1-2B)+t}\right] \mathrm{d}t\right) \mathrm{d}\mu(\lambda), \tag{2.61}$$

is finite. To that end, we use the cyclicity of the trace to bring the two resolvents $(1+\lambda(1-2B)+t)^{-1}$ together again, the bound from Eq. (2.55) [with $Q=(A-B)$ replaced by $Q_\pm$ on the left hand side] and Lemma 2.4 another time. In other words, we go from Eq. (2.60) to Eq. (2.55) in backward order with $Q=(A-B)$ replaced by $Q_\pm$. This shows the finiteness of Eq. (2.61).

Next, we go back to Eq. (2.61) and split the integral over λ into three parts, one from -1 to $-1/2$, one from $-1/2$ to $1/2$ and a last one from $1/2$ to 1. The integral from $-1/2$ to $1/2$ is easy to treat. We look at Eq. (2.61) again, adjust the boundaries of the integral over λ to run from $-1/2$ to $1/2$ and evaluate the trace in an arbitrary basis. Like in the proof of Lemma 2.3, we show that there is no singularity at $\lambda=0$. Together with the standard estimates used in the proof of Theorem 2.3, this implies that the expression inside the integral over λ can be bounded by a constant. Since μ is a probability measure this is enough to apply dominated convergence and interchange the sum coming from the trace and the integral over λ. The fact that this works for any basis, shows that

$$\int_{-1/2}^{1/2}\left(\frac{2}{\lambda}Q_\pm-\frac{2}{\lambda}\int_0^\infty\frac{1}{1+\lambda(1-2B)+t}Q_\pm\frac{1}{1+\lambda(1-2B)+t}\mathrm{d}t\right)\mathrm{d}\mu(\lambda),\tag{2.62}$$

is trace-class and that for this term the trace and the integral can be interchanged.

In the next step, we investigate the integral from 1/2 to 1, that is Eq. (2.61) with the adjusted integral boundaries. Since

$$\int_{1/2}^{1}\frac{2}{\lambda}\operatorname{Tr}Q_\pm\mathrm{d}\mu(\lambda)\leq 4\left\|Q\right\|_1\tag{2.63}$$

the first term inside the integral over λ can be integrated separately. Additionally, the trace and the integral over λ can be interchanged for this term as well. This implies that also

$$\int_{1/2}^{1}\frac{2}{\lambda}\int_0^\infty\operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t}Q_\pm\frac{1}{1+\lambda(1-2B)+t}\right]\mathrm{d}t\mathrm{d}\mu(\lambda)\tag{2.64}$$

is finite. Since the operator inside the trace is positive we can apply Fubini to interchange the trace with the integral over t and afterwards with the integral over λ. The same arguments work for the integral from -1 to $-1/2$. Putting all this together, we have shown that

$$\begin{aligned}&\int_{-1}^{1}\left(\frac{2}{\lambda}\operatorname{Tr}Q-\frac{2}{\lambda}\int_0^\infty\operatorname{Tr}\left[\frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2B)+t}\right]\mathrm{d}t\right)\mathrm{d}\mu(\lambda)\qquad(2.65)\\&\quad=\operatorname{Tr}\left[\int_{-1}^{1}\left(\frac{2}{\lambda}Q-\frac{2}{\lambda}\int_0^\infty\frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2B)+t}\mathrm{d}t\right)\mathrm{d}\mu(\lambda)\right],\end{aligned}$$

which together with Eq. (2.54) and Eq. (2.60) implies that

$$\begin{aligned}
&\mathrm{Tr}\left[\varphi'(B)(A-B)\right] &(2.66)\\
&=\mathrm{Tr}\left[aQ-\frac{b}{2}\int_{-1}^{1}\left\{\frac{2Q}{\lambda}-\frac{2}{\lambda}\int_{0}^{\infty}\frac{1}{1+\lambda(1-2B)+t}Q\frac{1}{1+\lambda(1-2B)+t}\mathrm{d}t\right\}\mathrm{d}\mu(\lambda)\right]
\end{aligned}$$

holds. In particular, the operator on the right hand side of Eq. (2.66) is trace-class. Since $\varphi'(B)(A-B)$ is trace-class by assumption we know that B cannot have eigenvalues at points of discontinuity of φ' with $(A-B)\neq 0$ on the corresponding eigenspaces. From this we conclude with the help of Lemma 2.3 that $\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ can be defined as a semibounded quadratic form on D. Also on D, it is the associated quadratic form of the operator under the trace on the right hand side of Eq. (2.66), see again Lemma 3. This operator is bounded and hence we can extend $\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi\left(\alpha A+(1-\alpha)B\right)|_{\alpha=0}$ to a bounded and symmetric quadratic form on all of h whose associated self-adjoint operator is the operator under the trace on the right hand side of Eq. (2.66). Hence, we have shown that

$$\mathrm{Tr}\left[\varphi'(B)(A-B)\right]=\mathrm{Tr}\left[\frac{\mathrm{d}}{\mathrm{d}\alpha}\varphi\left(\alpha A+(a-\alpha)B\right)\big|_{\alpha=0}\right]. \tag{2.67}$$

This concludes the proof of Theorem 2.4.

CHAPTER 3

A non-periodic version of the BCS functional in an external field

Abstract. We consider a many-body system of fermionic atoms interacting via a local pair potential V and subject to an external potential W. In order to describe such a gas in the thermodynamic limit within the framework of BCS theory, one has several possibilities: One is to consider a periodic sample where the notion of free energy per unit volume makes sense, and another is to measure the free energy of the whole sample with respect to the free energy of a reference state. In this work we take the second approach. We define a version of the BCS functional in the superconducting phase close to the critical temperature with a slowly varying weak external field and show that this energy functional is bounded from below. Using the lower bound, we find a-priori estimates for states with energy less than or equal to that of the reference state. Our results are the starting point for an extension of the derivation of Ginzburg-Landau theory in the sense of [28] to the present setting.

3.1 Introduction, set-up and main results

3.1.1 Introduction

In 1950, Ginzburg and Landau introduced a phenomenological model of superconductivity that has been extremely successful and is widely used in physics [43, 7]. On the one hand, it is possible to derive the previously known London equations, that model the electrodynamics of a superconductor, from Ginzburg-Landau theory and thereby reproduce their predictions. On the other hand, it captures many more phenomena such as the creation of vortices and the distinction between type-I and type-II superconductors. As with the London equations, the Ginzburg-Landau equations are a *macroscopic* theory of superconductivity and do not intend to give an explanation for the *microscopic* mechanism behind the phenomenon. Apart from their success in physics, they show a rich mathematical structure, which has been well recognized, see e.g. [44, 45, 46, 47, 48] and references therein.

Only a few years after the famous paper of Ginzburg and Landau, in 1957, Bardeen, Cooper and Schrieffer (BCS) proposed the first accepted *microscopic* theory of superconductivity [1], which was a major breakthrough and awarded them the Nobel prize in 1972. From a physics point of view, the work of BCS has three main ingredients. First of all, the observation that an attractive interaction that cannot bind two particles in free space leads to pairing in a Fermi gas, that is, in the presence of a Fermi sea. Secondly, the fact that the phonon-mediated effective interaction between conducting electrons in a metal may be attractive despite their natural Coulombic repulsion. And thirdly, the construction of a cleverly chosen class of trial states, called BCS states or quasi-free states, that capture the relevant properties of the normal state (Fermi liquid) as well as the ones of the superconducting state (pairing) and that are nevertheless simple enough to allow for a mathematical analysis of the model. From a mathematical point of view, the BCS theory of superconductivity (without external fields) has been studied by Hainzl, Seiringer and co-authors in [17, 16, 18, 19, 20, 21] and by other authors in [11, 12, 13, 14]. With external fields the BCS functional has been studied in the temperature zero case in [22, 23].

The connection between the *microscopic* BCS theory and the *macroscopic* Ginzburg-Landau theory was established in 1959 by Gor'kov [27]. He showed that close to the critical temperature it is possible to expand the equation for the position-dependent gap function Δ in orders of Δ. In doing so, he found that the center of mass part of Δ solves, to leading order, a Ginzburg-Landau equation whose coefficients depend on *microscopic* parameters of the system. A rigorous version of Gorkov's formal derivation was achieved in 2012 by Frank, Hainzl, Seiringer and Solovej in [28]. They show that close to the critical temperature and under the assumption of slowly varying weak external fields, the difference of the free energy of the superconducting state and that of the normal state is to leading order given by the minimum of a Ginzburg-Landau functional. The coefficients of this Ginzburg-Landau functional can be computed entirely from the minimizer of a translation-invariant version of the BCS functional, that is, from *microscopic* quantities. While Gorkov considers the particular non-local rank one interaction used in the original work of BCS, the authors of [28] allowed for a general class of local pair interactions that are relevant for the description of typical interactions in cold atomic gases. Since the atoms in such a gas do not carry a charge, the corresponding pairing mechanism is relevant for superfluidity rather than for superconductivity. Later, in 2014, the same authors studied in [29] the influence of external fields on the critical temperature. Under the same assumptions needed for the derivation of Ginzburg-Landau theory, they show that the next to leading order of the critical temperature (the leading order is determined by the translation-invariant BCS functional) is determined by the lowest eigenvalue of the linearization of the Ginzburg–Landau equation that arises as the Euler-Lagrange equation of the Ginzburg-Landau functional.

In order to model a fermionic gas in the thermodynamic limit that is exposed to external fields within the framework of BCS theory one has several options. One is to consider the situation of a periodic sample, that is, the external fields are assumed to be periodic, see [28]. This set-up allows for the definition of a free energy per unit volume and through the usual heuristic derivation [16] also for the definition of a BCS functional that describes this system. On the other hand, one could equally well assume a situation where the external fields are localized to some finite area in space and measure the free energy of the system with respect to the free energy of a suitably chosen reference state. Since it is an extensive quantity, the free energy of each state is, of course, infinite, but the difference of their two free energies has a reasonable chance to be finite. The functional describing a periodic sample has the advantage of being easier to handle from a mathematical point of view. However, it may be seen as a disadvantage that, especially for macroscopic external fields, the periodicity of the system is somewhat artificial. It is the aim of this work to follow the second approach.

Having defined the BCS functional, we study it in the physical setting in which Ginzburg-Landau theory is assumed to be valid, that is, close to the critical temperature and with weak and slowly varying external fields. In more mathematical terms, this means we investigate the BCS functional in the scaling that has been introduced in [28] to derive Ginzburg-Landau theory and that we will call the Ginzburg-Landau scaling in the following. Our main results are the construction of a lower bound for our BCS functional and the derivation of a-priori estimates for states with energy less than or equal to that of the reference state. To put things into perspective, we have to note that the construction of a lower bound for the BCS functional considered in [28] is fairly easy. This situation changes drastically if one considers the version of the BCS functional investigated in this work, where the construction of a lower bound is a challenging task. Since we investigate our functional in the Ginzburg-Landau scaling, the a-priori estimates for states with free energy less than or equal to that of the reference state should be compared to the a-priori estimates derived in [28, Chapter 5] for states with energy less than or equal to that of the normal state. They are the starting point for an extension of the derivation of Ginzburg-Landau theory to the set-up considered in this work.

3.1.2 Set-up

BCS states and the translation-invariant BCS functional

Consider a gas of fermionic atoms in three spatial dimensions that interact via a local two-body potential V and that are exposed to an external potential W. In BCS theory the quantum mechanical state of a system is most conveniently described by its generalized one-particle density matrix [8]. We call an operator $\Gamma \in \mathcal{L}\left(L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)\right)$ a BCS state if it is of the form

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \overline{\alpha} & 1-\overline{\gamma} \end{pmatrix} \qquad \text{with} \qquad 0 \leq \Gamma \leq 1. \tag{3.1}$$

Here, $\overline{\alpha} = C\alpha C$ with C denoting complex conjugation. The operators γ and α are usually called the one-particle density matrix and the Cooper-pair wave function of the state Γ, respectively. With the above definitions we cannot conclude that γ and α have integral kernels. Nevertheless, admissible states, which we are going to define below, will have this property. In terms of integral kernels, the definition $\overline{\alpha} = C\alpha C$ reads $\overline{\alpha}(x,y) = \overline{\alpha(x,y)}$. For the operators γ and α, Eq. (3.1) implies $\gamma^* = \gamma$ and $\alpha^* = \overline{\alpha}$, where again in terms of kernels the last condition can be rephrased as $\alpha(x,y) = \alpha(y,x)$. The spatial Cooper-pair wave function $\alpha(x,y)$ is symmetric because we do not include spin variables and always assume that Cooper-pairs are in a spin singlet state. This makes the overall Cooper-pair wave function antisymmetric under a combined exchange of position and spin variables.

In the absence of an external field, it is reasonable to restrict attention to translation-invariant states, that is, states with $\alpha(x,y) = \alpha(x-y)$ and $\gamma(x,y) = \gamma(x-y)$. A translation-invariant version of the BCS functional has been introduced and studied in detail in [16]. It reads

$$\begin{aligned} \mathcal{F}_{TI}(\Gamma) &= \int_{\mathbb{R}^3} (p^2-\mu)\gamma(p)\mathrm{d}p + \int_{\mathbb{R}^3} V(x)\,|\alpha(x)|^2\,\mathrm{d}x - TS(\Gamma), \\ S(\Gamma) &= -\frac{1}{2}\int_{\mathbb{R}^3} \mathrm{Tr}_{\mathbb{C}^2}\left[\Gamma(p)\ln(\Gamma(p)) + (1-\Gamma(p))\ln(1-\Gamma(p))\right]\mathrm{d}p, \\ \Gamma(p) &= \begin{pmatrix} \gamma(p) & \hat{\alpha}(p) \\ \overline{\hat{\alpha}(p)} & 1-\overline{\gamma(p)} \end{pmatrix}. \end{aligned} \tag{3.2}$$

Here, the parameters $T \geq 0$ and $\mu \in \mathbb{R}$ denote the temperature and the chemical potential of the fermionic gas, respectively. For $V \in L^{3/2}(\mathbb{R}^3)$, the natural domain of the above functional consists of all translation-invariant BCS states Γ such that for the corresponding pair (α, γ) one has $\alpha \in H^1(\mathbb{R}^3)$ and $\gamma \in L^1(\mathbb{R}^3, (1+p^2)\mathrm{d}p)$. The condition $0 \leq \Gamma \leq 1$ leads in the case of translation-invariant states to $0 \leq \gamma(p) \leq 1$ and $|\hat{\alpha}(p)|^2 \leq \gamma(p)(1-\gamma(p))$. The functional $\mathcal{F}_{TI}$ is coercive on its domain and attains its infimum. Additionally, it has been shown in [16] that there exists a critical temperature $T_c \geq 0$ such that for all $T \geq T_c$ the functional is minimized by the pair $\alpha = 0$ and $\gamma_n(p) = \left(1 + e^{(p^2-\mu)/(2T)}\right)^{-1}$, that is, no Cooper-pairs are present and γ_n is the one-particle density matrix of a free Fermi gas. We refer to this state as the normal state. In contrast, if $T < T_c$ (possible if $T_c > 0$) the minimizer will be of the form

$$\begin{aligned} \Gamma_0(p) &= \begin{pmatrix} \gamma_0(p) & \hat{\alpha}_0(p) \\ \overline{\hat{\alpha}_0(p)} & 1-\overline{\gamma_0(p)} \end{pmatrix} = \frac{1}{1+e^{\beta H_0(p)}}, \\ H_0(p) &= \begin{pmatrix} p^2-\mu & \hat{\Delta}_0(p) \\ \overline{\hat{\Delta}_0(p)} & -(p^2-\mu) \end{pmatrix}, \\ \Delta_0(x) &= 2V(x)\alpha_0(x) \end{aligned} \tag{3.3}$$

with $\alpha_0 \neq 0$. The three equations Eq. (3.3) are the Euler-Lagrange equations of the translation-invariant BCS functional $\mathcal{F}_{TI}$. States with $\alpha_0 \neq 0$ are, depending on the context, referred to as superconducting or superfluid states. The first of the above set of equations can be rewritten in two equations, one for α_0 alone and another one that allows us to compute γ_0 if α_0 is given, see [16]. The equation for α_0 reads

$$\left(K_T^{\Delta_0} + V\right) \alpha_0 = 0, \tag{3.4}$$

where

$$K_T^{\Delta_0} = \frac{E(-i\nabla)}{\tanh\left(\frac{E(-i\nabla)}{2T}\right)} \quad \text{and} \quad E(p) = \sqrt{(p^2-\mu)^2 + \left|\hat{\Delta}_0(p)\right|^2}, \tag{3.5}$$

that is, $K_T^{\Delta_0}$ is a pseudo-differential operator. Eq. (3.4) holds in Fourier space almost everywhere. The function $E(p)$ is often referred to as an effective dispersion relation. A justification for this naming and a heuristic derivation of $\mathcal{F}_{TI}$ from Quantum Mechanics can be found in [16].

The BCS functional with an external potential and the Ginzburg-Landau scaling

Having these prerequisites at hand, we come to the definition of our BCS functional. The goal is to give a meaning to the formal expression

$$\text{“}\mathcal{F}_\beta(\Gamma) = \mathrm{Tr}_{L^2(\mathbb{R}^3)}\left[(-\Delta - \mu + W)\,\gamma\right] + \int_{\mathbb{R}^6} V(x-y)\left|\alpha(x,y)\right|^2 \mathrm{d}(x,y) - TS(\Gamma)\text{”},$$
$$\text{“}S(\Gamma) = -\frac{1}{2}\mathrm{Tr}_{L^2(\mathbb{R}^3)\oplus L^2(\mathbb{R}^3)}\left[\Gamma\ln(\Gamma) + (1-\Gamma)\ln(1-\Gamma)\right]\text{”}, \tag{3.6}$$

or more precisely to

$$\text{“}\mathcal{F}_\beta(\Gamma,\Gamma') = \mathcal{F}_\beta(\Gamma) - \mathcal{F}_\beta(\Gamma')\text{”}, \tag{3.7}$$

where Γ' is a reasonably chosen reference state and β denotes the inverse temperature. In the following, we call $\mathcal{F}_\beta(\Gamma,\Gamma')$ the relative free energy of Γ with respect to Γ'. The most natural candidate for a reference state is certainly Γ_0, the minimizer of the translation-invariant BCS functional. Nevertheless, we decided instead to work with the state Γ_0^w defined by

$$\Gamma_0^w = \begin{pmatrix} \gamma_0^w & \alpha_0^w \\ \overline{\alpha_0^w} & 1-\overline{\gamma_0^w} \end{pmatrix} = \frac{1}{1+e^{\beta H_0^w}} \quad \text{where} \tag{3.8}$$
$$H_0^w = \begin{pmatrix} k(-i\nabla) + W & \hat{\Delta}_0(-i\nabla) \\ \overline{\hat{\Delta}_0}(-i\nabla) & -(k(-i\nabla)+W) \end{pmatrix}$$

and $k(p) = p^2 - \mu$. The advantage of Γ_0^w compared to Γ_0 is that it already includes the external potential W. If we insert Γ_0^w into Eq. (3.7) and rearrange the terms a little, we obtain

$$\mathcal{F}_\beta\left(\Gamma,\Gamma_0^w\right) = \frac{1}{2\beta}\mathcal{H}\left(\Gamma,\Gamma_0^w\right) + \int_{\mathbb{R}6} V(x-y)\left|\alpha(x,y) - \alpha_0^w(x,y)\right|^2 \mathrm{d}(x,y) \tag{3.9}$$
$$+ 2\mathrm{Re}\int_{\mathbb{R}^6} V(x-y)\left(\alpha(x,y) - \alpha_0^w(x,y)\right)\overline{\left(\alpha_0^w(x,y) - \alpha_0(x-y)\right)}\mathrm{d}(x,y)$$

where the relative entropy $\mathcal{H}(\Gamma, \Gamma_0^w)$ of the state Γ with respect to the state Γ_0^w is given by

$$\begin{aligned} \mathcal{H}(\Gamma, \Gamma_0^w) &= \operatorname{Tr}\left[\varphi(\Gamma) - \varphi(\Gamma_0^w) - \varphi'(\Gamma_0^w)(\Gamma - \Gamma_0^w)\right], \\ \varphi(x) &= x\ln(x) + (1-x)\ln(1-x). \end{aligned} \tag{3.10}$$

It includes the kinetic energy and the potential energy coming from W because $\varphi'(\Gamma_0^w) = -\beta H_0^w$. If we had instead chosen the state Γ_0, the potential W would not be included in the relative entropy but we would have one term of the form $\int_{\mathbb{R}^3} W(x)\gamma(x,x)\mathrm{d}x$, which turns out to be more difficult to handle. This is because the relative entropy is needed to dominate the interaction term, that is the second term on the right-hand side of (3.9). What remains from the relative entropy after this step is not strong enough to dominate a term of the form $\int_{\mathbb{R}^3} W(x)\gamma(x,x)\mathrm{d}x$. This problem does not occur if we choose Γ_0^w as the reference state. On the other hand, it is very reasonable that $\mathcal{F}_\beta(\Gamma_0, \Gamma_0^w) < \infty$. Since formally "$\mathcal{F}_\beta(\Gamma, \Gamma_0) = \mathcal{F}_\beta(\Gamma, \Gamma_0^w) - \mathcal{F}_\beta(\Gamma_0, \Gamma_0^w)$" this indicates that both choices of reference state should be possible. The third term in Eq. (3.9) appears because Γ_0^w does not solve an equation similar to the one that is solved by Γ_0, see Eq. (3.3). Since it is only linear in the variable α it is easy to handle in most circumstances. For $\mathcal{F}_\beta(\Gamma, \Gamma_0^w)$ to be a finite quantity, we need that $\alpha - \alpha_0^w$ has sufficient decay at infinity and that Γ is "close enough" to Γ_0^w so that the relative entropy is finite. This is a reasonable assumption because $\Gamma - \Gamma_0^w$ describes how the localized external potential W disturbs the formerly translation-invariant system. A class of BCS states for which the BCS functional yields a finite value will be introduced in the next section.

The goal of this work is to study the BCS functional in the physical setting in which Ginzburg-Landau theory is assumed to be valid, that is, in the superconducting phase but close to the critical temperature and with a weak and slowly varying external field W. The weakness assumption for the external field is natural because the closeness of T to the critical temperature defines an energy scale on which W has to live in order to contribute to the energy in a non-trivial way. External fields on a larger energy scale immediately destroy the superconducting state while fields on a smaller scale do not contribute to the energy to leading order. We implement these conditions as in [28] and choose $T = T_c(1 - Dh^2)$ for $h \ll 1$ where $T_c > 0$ is the critical temperature of the translation-invariant BCS functional and $D > 0$. The external potential is of the form $h^2 W(hx)$ and the interaction potential reads $V(x)$, that is, the interaction takes place on the *microscopic* scale. In the following, we will refer to this scaling as the Ginzburg-Landau scaling. As do the authors of [28], we find it convenient to work with macroscopic coordinates instead of microscopic ones. The change between the two coordinate systems is given by the unitary transformation $U : L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ defined by $(U\psi)(x) = h^{-3/2}\psi(x/h)$. In the new coordinates the external potential and the interaction potential are given by $h^2 W(x)$ and $V(x/h)$, respectively. We also have to transform the kinetic energy which in macroscopic coordinates is given by $(-ih\nabla)^2$.

The scaled version of the BCS functional in macroscopic coordinates reads

$$\begin{aligned}\mathcal{F}_\beta(\Gamma,\Gamma_0^w) &= \frac{1}{2\beta}\mathcal{H}(\Gamma,\Gamma_0^w) + \int_{\mathbb{R}6} V\left(\frac{x-y}{h}\right)|\alpha(x,y)-\alpha_0^w(x,y)|^2\,\mathrm{d}(x,y) \\ &+ 2\mathrm{Re}\int_{\mathbb{R}^6} V\left(\frac{x-y}{h}\right)(\alpha(x,y)-\alpha_0^w(x,y))\overline{\left(\alpha_0^w(x,y)-h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\right)}\mathrm{d}(x,y),\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}\Gamma_0^w &= \begin{pmatrix}\gamma_0^w & \alpha_0^w \\ \overline{\alpha_0^w} & 1-\overline{\gamma_0^w}\end{pmatrix} = \frac{1}{1+e^{\beta H_0^w}}, \\ H_0^w &= \begin{pmatrix} k(-ih\nabla)+h^2W(x) & \hat{\Delta}_0(-ih\nabla) \\ \hat{\Delta}_0(-ih\nabla) & -\left(k(-ih\nabla)+h^2W(x)\right)\end{pmatrix}.\end{aligned} \tag{3.12}$$

We highlight the particular scaling of the gap function $\hat{\Delta}_0(p)$ in the definition of H_0^w, that is in the definition of our reference state, which has to be compared to its unscaled version in macroscopic coordinates $h^{3/2}\hat{\Delta}_0(hp)$. It is chosen in such a way that all three terms of the BCS functional are comparable to leading order in h which is motivated by [28, Theorem 1] and in particular by Eq. (1.15) in this reference. It holds under the assumption that $\alpha_0 \sim h$ for h small enough, which is what one would expect. Lemma 3.7 tells us that $\|\alpha_0\|_{L^2(\mathbb{R}^3)} \lesssim h$ in this regime, which is a weaker statement. If α_0/h goes to zero for $h \to 0$ we adjust the scaling of $\hat{\Delta}_0$ by choosing a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that the $L^2(\mathbb{R}^3)$-norm of $f(h)\hat{\Delta}_0(hp)$ is proportional to $h^{-1/2}$ for small h. When we compare our setting with the one in [28], we have to note two things. First, since $\Delta_0(x) = 2V(x)\alpha_0(x)$, the decomposition for $\alpha(x,y)$ in [28, Eq. (1.15)] implies a similar one for $\Delta(x,y) = 2V((x-y)/h)\alpha(x,y)$. Second, we have to caution the reader that $\hat{\alpha}_0(hp)$ is of order h in our setting, while the same symbol denotes a function that does not depend on h in [28]. Hence, we have an additional factor of h^{-1} in front of $\hat{\Delta}_0(hp)$ compared to [28, Eq. (1.15)]. This difference occurs because in our work α_0 denotes the minimizer of the translation-invariant BCS functional which behaves as $\sqrt{T_c - T} \sim h$ for $T \to T_c$. On the other hand, in [28] it is the unique solution of the equation $(K_{T_c} + V)\,\alpha_0 = 0$, that is of the linearized Euler-Lagrange equation of $\mathcal{F}_{TI}$ at $T = T_c$ and hence does not depend on h except through the coordinates. Eq. (3.11) is the version of the BCS functional that we study in the rest of the text.

The relative entropy and admissible states

Let us have a closer look at the relative entropy defined in Eq. (3.10). A priori, it is well-defined if the operator under the trace is trace-class. Unfortunately, since the function $\varphi(x) = x\ln(x) + (1-x)\ln(1-x)$ does not have a bounded derivative for all $x \in [0,1]$, it is not easy to decide when this is the case. Additionally, since states in our set-up are generically non-compact it turns out to be quite hard to derive upper and lower bounds for $\mathcal{H}(\Gamma, \Gamma_0^w)$. When we consider matrices instead this is much easier because then we can use Klein's inequality, see e.g. [35], which reduces the problem of proving an inequality for the relative entropy to proving an inequality for numbers. This strategy can be extended to operators whose spectra solely consist of eigenvalues, which cannot be expected of Γ and Γ_0^w. For small h, the operator Γ_0^w for example is close to Γ_0, which is a pseudo-differential operator and has no eigenvalues at all. In [28] this problem can be circumvented by the definition of a trace per unit volume, which naturally leads to the definition of a relative entropy per unit volume that does not suffer from these problems.

In order to understand how to treat these difficulties, we first discuss how the relative entropy behaves for hermitian matrices A and B with $0 \leq A, B \leq 1$. Since φ is continuously differentiable on $(0,1)$ the relative entropy $\mathcal{H}(A,B)$ is certainly well-defined as long as $0 < B < 1$. This is still true when 0 and/or 1 are contained in the spectrum of B and $A = B$ on $\mathrm{Ker}(B)$, $\mathrm{Ker}(1-B)$ or $\mathrm{Ker}(B) \oplus \mathrm{Ker}(1-B)$. The trace is then understood to be taken on the complement of these subspaces. If these requirements are not fulfilled, we define $\mathcal{H}(A,B) = \infty$. This is reasonable because as long as $\mathcal{H}(A,B)$ is finite, it is also positive, which easily follows from the convexity of φ and an application of Klein's inequality. In order to obtain a well-behaved definition for the relative entropy for general operators, we do the following computation. If we exploit the cyclicity of the trace, it can easily be seen that

$$\mathcal{H}(A,B) = \mathrm{Tr}\left[\varphi(A) - \varphi(B) - \frac{\mathrm{d}}{\mathrm{d}s}\varphi\left(sA + (1-s)B\right)\Big|_{s=0}\right]. \tag{3.13}$$

Consult Chapter 2 for how to compute the derivative in Eq. (3.13). The above formula for the relative entropy has one important advantage compared to the original one: Since the function φ is not only convex but also operator convex, see Chapter 2, this implies that the operator under the trace is positive. This can be seen, at least on a formal level, like in the case of numbers. Since the trace of a positive operator is always well-defined, Eq. (3.13) has a chance to yield a reasonable definition for a relative entropy for all bounded operators A, B with $0 \leq A, B \leq 1$ which takes values in the positive extended real numbers. From now on, we will use Eq. (3.13) instead of Eq. (3.10) as the defining equation for the relative entropy.

The details of the definition of the operator under the trace on the right-hand side of Eq. (3.13) as well as the definition of its trace can be found in Chapter 2. This is motivated by a recent paper of Lewin and Sabin. In [36] they define a family of generalized relative entropies for bounded operators by a limiting procedure that can be interpreted as a thermodynamic limit. The analysis in [36] together with the results from Chapter 2 indicate that Eq. (3.13) is the correct definition for the relative entropy. On the one hand it equals the previous definition if matrices are considered and on the other hand, it naturally appears as a limit of relative entropies when the thermodynamic limit is taken. In that sense Eq. (3.13) extends the usual definition of the relative entropy to general bounded operators. The fact that this relative entropy can be approximated as mentioned above also yields a strategy for constructing upper and lower bounds. If one chooses a sequence $\{P_n\}_{n=1}^{\infty}$ of finite dimensional projections with $P_n \to 1$ in the strong operator topology, Theorem 2.3 tells us that $\mathcal{H}(P_nAP_n, P_nBP_n) \to \mathcal{H}(A,B)$ as n tends to infinity. A bound for the approximate relative entropies $\mathcal{H}(P_nAP_n, P_nBP_n)$ can be constructed with the help of Klein's inequality because P_nAP_n and P_nBP_n have finite rank. What remains to be shown afterwards is that this bound behaves nicely when the limit $n \to \infty$ is taken. This strategy has been applied in the proof of Lemma 3.1 to construct a lower bound for the relative entropy.

The relative entropy defined above is a positive quantity which makes it natural to call a BCS state Γ *admissible* if $\mathcal{H}(\Gamma, \Gamma_0^w) < \infty$ holds. To see what kind of properties *admissible* states have, we apply the inequality for the relative entropy that has been proven in Lemma 3.1 to $\mathcal{H}(\Gamma, \Gamma_0^w)$. When we neglect the second term on the right-hand side of of Eq. (3.1), we obtain

$$\mathcal{H}(\Gamma, \Gamma_0^w) \geq \mathrm{Tr}\left[(\Gamma - \Gamma_0^w)\frac{\beta H_0^w}{\tanh\left(\frac{\beta H_0^w}{2}\right)}(\Gamma - \Gamma_0^w)\right] \geq 2\mathrm{Tr}\left[(\Gamma - \Gamma_0^w)^2\right]. \tag{3.14}$$

To come to the right-hand side of the above equation, we used that $x/\tanh\left(\frac{x}{2T}\right) \geq 2T$ holds for all $x \in \mathbb{R}$. Hence, for any *admissible* state Γ the operator $\Gamma - \Gamma_0^w$ is Hilbert-Schmidt. Since Γ_0^w has an integral kernel, see Section 3.3.2, this implies that Γ, and therefore γ and α, also have integral kernels. In particular, we have $\alpha - \alpha_0^w \in L^2(\mathbb{R}^6)$ and $\gamma - \gamma_0^w \in L^2(\mathbb{R}^6)$. When we assume that V is a bounded function and apply Lemma 3.8 to control the third term in Eq. (3.11), those properties guarantee $|\mathcal{F}_\beta(\Gamma, \Gamma_0^w)| < \infty$ for any *admissible* state Γ.

3.1.3 Main results

Before we state our results, let us make a few assumptions. In order to be able to carry out computations in a convenient way, we assume some regularity conditions for our potentials V and W.

Assumption 1 *We assume for the interaction potential V that $V \in H^1(\mathbb{R}^3) \cap W^{1,\infty}(\mathbb{R}^3)$ together with $\hat{V} \in L^1(\mathbb{R}^3) \cap H^4(\mathbb{R}^3) \cap W^{2,\infty}(\mathbb{R}^3)$. Additionally, we assume that V is a symmetric function, that is $V(-x) = V(x)$ for almost all $x \in \mathbb{R}^3$. The external potential W obeys $W \in H^1(\mathbb{R}^3) \cap W^{1,\infty}(\mathbb{R}^3)$ with $\hat{W} \in L^1(\mathbb{R}^3) \cap W^{4,1}(\mathbb{R}^3) \cap W^{4,\infty}(\mathbb{R}^3)$ and $(1 + (\cdot)^2)\hat{W} \in L^\infty(\mathbb{R}^3)$.*

By $H^n(\mathbb{R}^3)$ and $W^{n,p}(\mathbb{R}^3)$ we denote the usual Sobolev spaces equipped with their natural norms. Most of the above assumptions could be relaxed, but we rather prefer to keep the proofs to a reasonable length. On the other hand, the following assumptions for the interaction potential V are crucial.

Assumption 2 *The potential V is such that the following two statements are true: (i) $T_c > 0$, (ii) $\hat{V} \leq 0$ and not identically zero.*

Remark 3.1 *Property (i) holds for example if $\mu > 0$ and $V \in L^{3/2}(\mathbb{R}^3)$ is negative and not identically zero, see [16, Theorem 3]. The same Theorem also tells us that one can add a positive part V_+ to the potential V and thereby keep the property $T_c > 0$ if V_+ is small enough in a suitable sense.*

Remark 3.2 *Let us assume $W = 0$ for the moment and consider the unscaled version of the BCS functional. The condition $\hat{V} \leq 0$ and not identically zero implies that the minimizer of the translation-invariant BCS functional (which in this situation is unique up to the choice of a phase, see Lemma 3.7) is also the unique minimizer of $\mathcal{F}_\beta(\Gamma, \Gamma_0)$. In particular, the translational symmetry of the system is not broken. This can be seen in the following way: When we apply the inequality for the relative entropy that has been proven in Lemma 3.1 and discard the second term on the right-hand side of Eq.* (3.1)*, we obtain*

$$\begin{aligned}
\mathcal{F}_\beta(\Gamma, \Gamma_0) &\geq \frac{1}{2} Tr \left[(\Gamma - \Gamma_0) \frac{H_0}{\tanh\left(\frac{H_0}{2T}\right)} (\Gamma - \Gamma_0) \right] \\
&\qquad + \int_{\mathbb{R}^6} V(x-y) \left| \alpha(x,y) - \alpha_0(x-y) \right|^2 d(x,y) \\
&= \int_{\mathbb{R}^3} \left(\alpha - \alpha_0, (K_{T,x}^{\Delta_0} + V_y(x))(\alpha - \alpha_0) \right)_{L^2(\mathbb{R}^3)} dy \\
&\qquad + 2T\, Tr\left[(\gamma - \gamma_0)^2 \right] \\
&\geq 0.
\end{aligned} \tag{3.15}$$

The second line has to be understood so that the operator $K_{T,x}^{\Delta_0} + V(x-y)$ acts on the x-variable of the function $\alpha(x,y) - \alpha_0(x-y)$. After the $L^2(\mathbb{R}^3)$-inner product in the x-variable is evaluated, one integrates over the y-variable. To come from the first to the second line, we used that $x \mapsto x/\tanh(x/(2T))$ is an even function and that $H_0(p)^2 = \mathbb{1}_{\mathbb{C}^2} E(p)^2$. The expression in the second line is due to our assumption for V nonnegative, see Lemma 3.7. It equals zero if and only if α is of the form $\alpha(x,y) = \alpha_0(x-y)\psi(y)$ for some measurable function ψ. Since α is symmetric, that is $\alpha(x,y) = \alpha(y,x)$, we conclude that ψ is a constant. From Eq. (3.15) *we also conclude that $\gamma = \gamma_0$ holds. If we look for a minimizer, we can restrict attention to states of this kind. But the translation-invariant BCS functional has a unique minimizer (that is a state with lowest energy per unit volume) in this set, which tells us that this state also minimizes $\mathcal{F}_\beta(\Gamma, \Gamma_0)$, that is, $\psi = 1$.*

Remark 3.3 *If the interaction potential V is rotationally symmetric the fact that Γ_0 is the unique minimizer of the BCS functional also tells us that Cooper-pairs are necessarily in an s-wave state, that is, they have angular momentum zero. This is because if α_0 was not a radial function we could easily construct new minimizers of the BCS functional by simply rotating α_0 about some axis. This cannot be true and hence α_0 is radial which means that it lives in the zero angular momentum sector.*

During the construction of the lower bound for the BCS functional, it turns out to be natural to decompose $\alpha(x, y)$ into two mutually orthogonal parts. A similar orthogonal decomposition also plays a crucial role in the derivation of the Ginzburg-Landau functional in [28]. To any Cooper-pair wave function α, we associate a measurable function ψ by

$$\psi(y) = \frac{\int_{\mathbb{R}^3} \alpha_0\left(\frac{x-y}{h}\right) \alpha(x, y) \mathrm{d}x}{\int_{\mathbb{R}^3} |\alpha_0(x)|^2 \, \mathrm{d}x}. \tag{3.16}$$

The function $\xi_0(x, y)$ is defined by

$$\alpha(x, y) = h^{-3} \alpha_0\left(\frac{x-y}{h}\right) \psi(y) + \xi_0(x, y). \tag{3.17}$$

The particular scaling of the first term on the right-hand side of Eq. (3.17) is chosen under the assumption that $\alpha_0 \sim h$ for small h and has to be adjusted if α_0 goes to zero more quickly. For more details, see the discussion at the end of Chapter 3.1.2. Note that $h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\psi(y)$ lies in the kernel of the operator $K_{T,x}^{\Delta_0} + V(x-y)$, where as above the operator $K_{T,x}^{\Delta_0}$ is understood to act on the x-component. Because of our assumption that the kernel of $K_T^{\Delta_0} + V$ is one dimensional, the orthogonality relation Eq. (3.16) implies that $\xi_0(x, y)$ lies in the orthogonal complement of this kernel.

Our main theorem is a lower bound for the BCS functional defined in Eq. (3.11) which implies a-priori estimates for states Γ with energy less than or equal to that of the reference state. To state our result, we need the decomposition of α introduced in Eqs. (3.16) and (3.17).

Theorem 3.1 *Let Γ be an admissible state with $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \leq 0$. Then for $r > 0$ large enough and $h > 0$ small enough, there exist constants $C_1, C_2 > 0$ depending on r such that*

$$\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \geq C_1 \Bigg(h \|\nabla\psi\|_{L^2(\mathbb{R}^3)}^2 + h \left\| \widehat{|\psi|^2 - 1} \right\|_{L^2(B_r)}^2 + \|\xi_0\|_{H_1(\mathbb{R}^6)}^2 \tag{3.18}$$
$$+ \|\gamma - \gamma_0^w\|_{H^1(\mathbb{R}^6)}^2 \Bigg) - C_2 h.$$

In the above equation, B_r denotes the ball of radius r centered at zero and the $H^1(\mathbb{R}^6)$-norms are, according to our choice of coordinates, given by $\|f\|_{H^1(\mathbb{R}^6)}^2 = \|f\|_{L^2(\mathbb{R}^6)}^2 + \|h\nabla_x f\|_{L^2(\mathbb{R}^6)}^2 + \|h\nabla_y f\|_{L^2(\mathbb{R}^6)}^2$. Eq. (3.18) implies the a-priori bounds

$$\|\nabla\psi\|_{L^2(\mathbb{R}^3)} + \left\| |\psi|^2 - 1 \right\|_{L^2(\mathbb{R}^3)} \leq C, \tag{3.19}$$
$$\|\xi_0\|_{H^1(\mathbb{R}^6)} + \|\gamma - \gamma_0^w\|_{H^1(\mathbb{R}^6)} \leq C h^{1/2},$$

for an appropriately chosen constant $C > 0$.

Remark 3.4 *Although the construction of a lower bound for the periodic version of the BCS functional treated in [28] is an easy task, the same problem for the BCS functional investigated in this work is a much harder problem.*

Remark 3.5 *The leading order of the Cooper-pair wave function is scaled such that the leading order contributions of the three terms in the BCS functional are of the same order in h. Because of this, we believe that the construction of the lower bound in the Ginzburg-Landau scaling already captures the main difficulties of the construction of a lower bound in the general situation. In order to derive a lower bound for the unscaled version of this BCS functional, one would have to explicitly use that W not only localizes* ψ *but also* ξ_0 *and* $\gamma - \gamma_0^w$ *which changes some estimates in Section 3.2.2. Afterwards, the analysis in Section 3.2.3 would have to be adjusted.*

Remark 3.6 *The a-priori estimates for states with energy less than or equal to that of the reference state in Theorem 3.1 should be compared with [28, Chapter 5], where related estimates for states with energy less than or equal to that of the normal state are derived. As the estimates in [28, Chapter 5] are the first step in the derivation of Ginzburg-Landau theory, the a-priori estimates stated in Theorem 3.1 are the starting point for an extension of the derivation of Ginzburg-Landau theory to our set-up.*

3.2 Proof of Theorem 3.1

Our construction of the lower bound for the BCS functional and the proof of the a-priori estimates will be done in three steps, which we will describe very briefly. In the first step, we prove an inequality for the relative entropy that has been introduced in [28] in the setting of periodic states and that will enable us to control the quadratic interaction term in the BCS functional, that is, the second term on the right-hand side of Eq. (3.11). This inequality is harder to prove in our case because our states are in general non-compact and hence Klein's inequality is not directly applicable. What remains to be controlled after this step are the terms proportional to the external potential $h^2 W$ and the linear interaction term, which is the third term on the right-hand side of Eq. (3.11). In the second step, we derive bounds on these non-positive terms. Whenever $h^2 W$ is acting on $\gamma - \gamma_0^w$, it is sufficient to use the fact that $h^2 \|W\|_{L^\infty(\mathbb{R}^3)}$ is small. In contrast, if $h^2 W$ is acting on $\alpha - \alpha_0^w$ we explicitly have to use that W is a localized function. In the third and final step, we use the bounds derived in step 1 and step 2 to show that the BCS functional is bounded from below. Having this bound at hand, the a-priori estimates can be proven.

3.2.1 Step 1: Lower bound for the relative entropy and domination of the quadratic interaction term

We start our discussion with a version of an inequality for the relative entropy that has for the first time been shown in [28] in the setting of periodic states.

Lemma 3.1 *Let Γ and Γ' be BCS states such that $\Gamma' = \left(1+e^H\right)^{-1}$ where H is a self-adjoint operator on $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$. Assume further that $\mathcal{H}(\Gamma, \Gamma') < \infty$. Then the inequality*

$$\mathcal{H}(\Gamma,\Gamma') \geq Tr\left[(\Gamma-\Gamma')\frac{H}{\tanh(H/2)}(\Gamma-\Gamma')\right] + \frac{4}{3}\,Tr\left[\Gamma(1-\Gamma)-\Gamma'(1-\Gamma')\right]^2 \quad (3.20)$$

holds.

Proof. In [28, Lemma 1] it has been shown that for any pair of real numbers $0 < x$, $y < 1$, one has

$$x\ln\left(\frac{x}{y}\right)+(1-x)\ln\left(\frac{1-x}{1-y}\right) \geq \frac{\ln\left(\frac{1-y}{y}\right)}{1-2y}(x-y)^2+\frac{4}{3}\left(x(1-x)-y(1-y)\right)^2. \quad (3.21)$$

The strategy for our proof is to approximate Γ and Γ' by finite rank operators in order to be able to apply Klein's inequality. Afterwards, the inequality has to be controlled in the limit $n \to \infty$, where n denotes the dimension of the approximation. Let $\{P_n\}_{n=1}^{\infty}$ be an increasing sequence of orthogonal finite dimensional projections with $P_n \to 1$ in the strong operator topology of $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$. Using Klein's Lemma, see e.g. [35], Eq. (3.21) implies

$$\begin{aligned}\mathcal{H}(\Gamma_n,\Gamma_n') \geq \operatorname{Tr}&\left[(\Gamma_n-\Gamma_n')\frac{\ln\left(\frac{1-\Gamma_n'}{\Gamma_n'}\right)}{1-2\Gamma_n'}(\Gamma_n-\Gamma_n')\right] \quad (3.22)\\ &+\frac{4}{3}\operatorname{Tr}\left[\Gamma_n(1-\Gamma_n)-\Gamma_n'(1-\Gamma_n')\right]^2,\end{aligned}$$

where $\Gamma_n = P_n\Gamma P_n$ and $\Gamma_n' = P_n\Gamma' P_n$. The left-hand side of Eq. (3.22) converges to $\mathcal{H}(\Gamma,\Gamma')$ as n tends to infinity, see [37, Theorem 3]. To treat the terms on the right-hand side, we apply three different versions of Fatou's Lemma.

Since $P_n \xrightarrow{s} 1$ as n tends to infinity, we know that Γ_n converges strongly to Γ and the same is true for Γ_n' and Γ'. But this implies $\left[\Gamma_n(1-\Gamma_n)-\Gamma_n'(1-\Gamma_n')\right]^2 \xrightarrow{s} \left[\Gamma(1-\Gamma)-\Gamma'(1-\Gamma')\right]^2$. Fatou's Lemma for traces, see e.g. [52, Theorem 2.7], is applicable and yields

$$\liminf_{n\to\infty}\operatorname{Tr}\left[\Gamma_n(1-\Gamma_n)-\Gamma_n'(1-\Gamma_n')\right]^2 \geq \operatorname{Tr}\left[\Gamma(1-\Gamma)-\Gamma'(1-\Gamma')\right]^2. \quad (3.23)$$

The same strategy cannot be applied for the first term on the right-hand side of Eq. (3.22) because the operator under the trace in the middle may become unbounded. To obtain a similar result for this term, we choose an ONB $\{e_\alpha\}_{\alpha=1}^{\infty}$ of $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ and use Fatou's Lemma for functions on measure spaces.

$$\liminf_{n\to\infty} \sum_{\alpha=1}^{\infty} \left(e_\alpha, \left(\Gamma_n - \Gamma'_n\right) \frac{\ln\left(\frac{1-\Gamma'_n}{\Gamma'_n}\right)}{1-2\Gamma'_n} \left(\Gamma_n - \Gamma'_n\right) e_\alpha \right) \tag{3.24}$$
$$\geq \sum_{\alpha=1}^{\infty} \liminf_{n\to\infty} \left(e_\alpha, \left(\Gamma_n - \Gamma'_n\right) \frac{\ln\left(\frac{1-\Gamma'_n}{\Gamma'_n}\right)}{1-2\Gamma'_n} \left(\Gamma_n - \Gamma'_n\right) e_\alpha \right).$$

To obtain the result, we think of α as being the integration variable. The application of Fatou's Lemma is justified because $\lambda \mapsto \frac{\ln\left(\frac{1-\lambda}{\lambda}\right)}{1-2\lambda}$ is a positive function and hence the expectation values in Eq. (3.24) are also positive. Let $\mu_{n\alpha}$ be the spectral measure of the operator Γ'_n with respect to the vector $\left(\Gamma_n - \Gamma'_n\right) e_\alpha$. The expectation values in Eq. (3.24) can be written as

$$\left(\left(\Gamma_n - \Gamma'_n\right) e_\alpha, \frac{\ln\left(\frac{1-\Gamma'_n}{\Gamma'_n}\right)}{1-2\Gamma'_n} \left(\Gamma_n - \Gamma'_n\right) e_\alpha \right) = \int_0^1 \frac{\ln\left(\frac{1-\lambda}{\lambda}\right)}{1-2\lambda} \mathrm{d}\mu_{n\alpha}(\lambda). \tag{3.25}$$

Correspondingly, we denote by μ_α the spectral measure of the operator Γ' with respect to the vector $\left(\Gamma - \Gamma'\right) e_\alpha$. Let $\left(\mathcal{C}([0,1]), \|\cdot\|_{\sup}\right)$ be the Banach space of continuous functions from $[0,1]$ to $\mathbb{C}$ and denote by $\mathcal{M}([0,1])$ its dual, the space of Borel measures on $[0,1]$. It can easily be seen that $\mu_{n\alpha} \xrightarrow{n\to\infty} \mu_\alpha$ for all $\alpha \in \mathbb{N}$ in the weak-$*$ topology of $\mathcal{M}([0,1])$. This is because the map $\Phi : X \times \mathcal{L}(X) \to \mathcal{M}$, $\Phi(\psi, A) = \nu_\psi(A)$ (X a separable complex Hilbert space, $\nu_\psi(A)$ the spectral measure of a bounded operator A with respect to the vector ψ) is jointly continuous when X is equipped with the norm topology, $\mathcal{L}(X)$ with the strong operator topology and $\mathcal{M}$ with the weak-$*$ topology.

To obtain a lower bound for the integral on the right-hand side of Eq. (3.25), we use another version of Fatou's Lemma for measures, which gives

$$\liminf_{n\to\infty} \int_0^1 \frac{\ln\left(\frac{1-\lambda}{\lambda}\right)}{1-2\lambda} \mathrm{d}\mu_{n\alpha}(\lambda) \geq \int_0^1 \frac{\ln\left(\frac{1-\lambda}{\lambda}\right)}{1-2\lambda} \mathrm{d}\mu_\alpha(\lambda). \tag{3.26}$$

Let us argue why Eq. (3.26) holds. Abbreviate $f(\lambda) = \frac{\ln\left(\frac{1-\lambda}{\lambda}\right)}{1-2\lambda}$ and define for $0 < \delta < 1/4$ the function $f^\delta(\lambda)$ by

$$f^\delta(\lambda) = \begin{cases} f(\delta) & \text{for } \lambda \in [0, \delta) \\ f(\lambda) & \text{for } \lambda \in [\delta, 1-\delta], \\ f(1-\delta) & \text{for } \lambda \in (1-\delta, 1]. \end{cases} \tag{3.27}$$

Since f is a positive function and each $\mu_{n\alpha}$ is a positive measure we have

$$\begin{aligned} \int_0^1 f^\delta(\lambda)\mathrm{d}\mu_{n\alpha}(\lambda) &\le \int_0^1 f(\lambda)\mathrm{d}\mu_{n\alpha}(\lambda) \\ \Rightarrow \int_0^1 f^\delta(\lambda)\mathrm{d}\mu_{\alpha}(\lambda) &\le \liminf_{n\to\infty} \int_0^1 f(\lambda)\mathrm{d}\mu_{n\alpha}(\lambda) \\ \Rightarrow \int_0^1 f(\lambda)\mathrm{d}\mu_{\alpha}(\lambda) &\le \liminf_{n\to\infty} \int_0^1 f(\lambda)\mathrm{d}\mu_{n\alpha}(\lambda). \end{aligned} \tag{3.28}$$

To come from the first to the second line, we used the weak-$*$ convergence of the measures $\mu_{n\alpha}$. In the second step, we applied monotone convergence. Translated back to the language of operators, Eq. (3.26) implies

$$\begin{aligned} \sum_{\alpha=1}^{\infty} \liminf_{n\to\infty} \left((\Gamma_n - \Gamma'_n)\, e_\alpha, \frac{\ln\left(\frac{1-\Gamma'_n}{\Gamma'_n}\right)}{1-2\Gamma'_n} (\Gamma_n - \Gamma'_n)\, e_\alpha \right) \\ \ge \mathrm{Tr}\left[(\Gamma-\Gamma') \frac{\ln\left(\frac{1-\Gamma'}{\Gamma'}\right)}{1-2\Gamma'} (\Gamma-\Gamma') \right]. \end{aligned} \tag{3.29}$$

A few algebraic manipulations show $\frac{\ln\left(\frac{1-\Gamma'}{\Gamma'}\right)}{1-2\Gamma'} = \frac{H}{\tanh(H/2)}$, which concludes our proof. □

When we apply the inequality for the relative entropy to $\mathcal{H}(\Gamma, \Gamma_0^w)$, we obtain two terms. One of them reads

$$\mathrm{Tr}\left[(\Gamma - \Gamma_0^w) \frac{H_0^w}{\tanh\left(\frac{\beta H_0^w}{2}\right)} (\Gamma - \Gamma_0^w) \right]. \tag{3.30}$$

In order to write this expression as a sum of one positive term plus corrections that are proportional to $h^2 W$, we use the identity [49, (4.3.91)]

$$\frac{x}{\tanh(x/2)} = 2 + \frac{1}{2}\sum_{n=1}^{\infty} \frac{x^2}{x^2/4 + n^2\pi^2} = 2 + \sum_{n=1}^{\infty}\left(2 - \frac{2n^2\pi^2}{x^2/4+n^2\pi^2}\right). \tag{3.31}$$

Insertion of H_0^w yields

$$\frac{H_0^w}{\tanh\left(\frac{H_0^w}{2T}\right)} = 2T + 2T\sum_{n=1}^{\infty}\left(1 - \frac{c^2n^2}{(H_0^w)^2 + c^2n^2}\right), \tag{3.32}$$

where we have introduced the constant $c = 2\pi T$. Next, we expand the resolvent in

Eq. (3.32) as

$$\frac{1}{(H_0^w)^2+c^2n^2} = \frac{1}{H_0^2+c^2n^2} - \frac{1}{H_0^2+c^2n^2}\left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{H_0^2+c^2n^2} + \frac{1}{H_0^2+c^2n^2}\left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{(H_0^w)^2+c^2n^2} \times \left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{H_0^2+c^2n^2}. \tag{3.33}$$

In the above equation, we denote $\omega = \begin{pmatrix} W & 0 \\ 0 & -W \end{pmatrix}$. The resolvent expansion of Eq. (3.33) clearly yields a decomposition of $H_0^w/\tanh(\beta H_0^w/2)$ of the form

$$\frac{H_0^w}{\tanh\left(\frac{\beta H_0^w}{2}\right)} = \mathbb{1}_{\mathbb{C}^2} K_T^{\Delta_0} + A + B, \tag{3.34}$$

where A and B are the operators given by

$$A = \sum_{n=1}^{\infty} \frac{1}{E^2+c^2n^2}\left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{E^2+c^2n^2}, \tag{3.35}$$

$$B = -\sum_{n=1}^{\infty} \frac{1}{E^2+c^2n^2}\left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{(H_0^w)^2+c^2n^2} \times \left(H_0h^2\omega + h^2\omega H_0 + h^4\omega^2\right)\frac{1}{E^2+c^2n^2}.$$

To obtain the results Eq. (3.34) and Eq. (3.35), we have used on the one hand that $H_0^2 = \mathbb{1}_{\mathbb{C}^2} E^2$ holds and on the other hand that $x \mapsto x/\tanh(x/2T)$ is an even function of x. We recall that H_0 and E are multiplication operators in Fourier space and that ω is a multiplication operator in position space.

Since we will have to deal with $\Gamma - \Gamma_0^w$ frequently, we introduce the following notation:

$$Q = \Gamma - \Gamma_0^w, \qquad \Lambda = \alpha - \alpha_0^w \quad \text{and} \quad q = \gamma - \gamma_0^w. \tag{3.36}$$

When we explicitly evaluate the trace over the $\mathbb{C}^2$-matrix structure in the term $\operatorname{Tr}\left[QK_T^{\Delta_0}Q\right]$ and use $K_T^{\Delta_0} \geq 2T$ as well as $K_T^{\Delta_0} \geq C(1+p^2)$, which holds for an appropriately chosen constant $C > 0$, we arrive at the following lower bound for the relative entropy:

$$\frac{1}{2\beta}\mathcal{H}(\Gamma, \Gamma_0^w) \geq \int_{\mathbb{R}^3} \left(\Lambda, K_{T,x}^{\Delta_0}\Lambda\right)_{L^2(\mathbb{R}^3,\mathrm{d}x)} \mathrm{d}y + T\left\|q\right\|^2_{L^2(\mathbb{R}^6)} + C\left\|q\right\|^2_{H^1(\mathbb{R}^6)} + \frac{1}{2}\operatorname{Tr}\left[Q(A+B)Q\right] + \frac{2}{3\beta}\operatorname{Tr}\left[\Gamma\left(1-\Gamma\right) - \Gamma_0^w\left(1-\Gamma_0^w\right)\right]^2. \tag{3.37}$$

The subscript x in the operator $K_{T,x}^{\Delta_0}$ tells us that it is acting on the x-component of the function $\Lambda(x,y)$. The first term on the right-hand side of Eq. (3.37) has to be understood so that one first evaluates the $L^2(\mathbb{R}^3)$-inner product in the x-coordinate and then integrates over y. We recall that the $H^1(\mathbb{R}^6)$-norm in the above equation has, according to our choice of coordinates, a factor of h in front each derivative.

Following the strategy of [28, Chapter 5], we now derive a lower bound for the last term in Eq. (3.37). When we write the traces of the operator-valued matrices explicitly in terms of their components, it can easily be seen that

$$\mathrm{Tr}\,[\Gamma\,(1-\Gamma)-\Gamma_0^w\,(1-\Gamma_0^w)]^2 \geq\ 2\mathrm{Tr}\left[\gamma(1-\gamma)-\gamma_0^w(1-\gamma_0^w)-\alpha\overline{\alpha}+\alpha_0^w\overline{\alpha_0^w}\right]^2 \tag{3.38}$$

holds. We also claim that

$$\begin{aligned}2\mathrm{Tr}\,(\gamma-\gamma_0^w)^2+\frac{4}{3}\ \mathrm{tr}\left[\gamma(1-\gamma)-\gamma_0^w\,(1-\gamma_0^w)-\alpha\overline{\alpha}+\alpha_0^w\overline{\alpha_0^w}\right]^2 \\ \geq \frac{4}{5}\ \mathrm{Tr}\left(\alpha\overline{\alpha}-\alpha_0^w\overline{\alpha_0^w}\right)^2,\end{aligned} \tag{3.39}$$

which follows from

$$\begin{aligned}\|\alpha\overline{\alpha}-\alpha_0^w\overline{\alpha}_0^w\|_2 \leq \|\gamma(1-\gamma)-\gamma_0^w(1-\gamma_0^w)-(\alpha\overline{\alpha}-\alpha_0^w\overline{\alpha}_0^w)\|_2 \\ +\|\gamma(1-\gamma)-\gamma_0^w\,(1-\gamma_0^w)\|_2\end{aligned} \tag{3.40}$$

together with

$$\|\gamma(1-\gamma)-\gamma_0^w\,(1-\gamma_0^w)\|_2 \leq \|\gamma-\gamma_0^w\|_2\,. \tag{3.41}$$

In the above formulas, $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm. Eq. (3.41) can be seen to hold as follows. Choose an increasing sequence of orthogonal finite rank projections $\{P_n\}_{n=1}^{\infty}$ acting on $L^2(\mathbb{R}^3)$ that converges to the identity in the strong operator topology. Using Klein's inequality, one easily shows that

$$\mathrm{Tr}\left(\gamma_n\,(1-\gamma_n)-\gamma_{0,n}^w\left(1-\gamma_{0,n}^w\right)\right)^2 \leq \mathrm{Tr}\left(\gamma_n-\gamma_{0,n}^w\right)^2, \tag{3.42}$$

where we have introduced the notation $\gamma_n = P_n\gamma P_n$ and $\gamma_{0,n}^w = P_n\gamma_0^w P_n$. Let us first have a closer look at the right-hand side of Eq. (3.42). Using $P_n \leq 1$, we find

$$\left(\gamma_n-\gamma_{0,n}^w\right)^2 \leq P_n\left(\gamma-\gamma_{0,n}^w\right)^2 P_n. \tag{3.43}$$

But this implies

$$\liminf_{n\to\infty}\mathrm{Tr}\left(\gamma_n-\gamma_{0,n}^w\right)^2 \leq \liminf_{n\to\infty}\mathrm{Tr}\left(P_n\,(\gamma-\gamma_0^w)^2\,P_n\right)=\mathrm{Tr}\,(\gamma-\gamma_0^w)^2\,. \tag{3.44}$$

On the other hand, we know that

$$\left(\gamma_n\,(1-\gamma_n)-\gamma_{0,n}^w\left(1-\gamma_{0,n}^w\right)\right)^2 \to \left(\gamma\,(1-\gamma)-\gamma_0^w\,(1-\gamma_0^w)\right)^2 \tag{3.45}$$

strongly which allows for an application of the non-commutative Fatou Lemma [52, Theorem 2.7]. We find

$$\begin{aligned}\mathrm{Tr}\left(\gamma\,(1-\gamma)-\gamma_0^w\,(1-\gamma_0^w)\right)^2 &\leq \liminf_{n\to\infty}\mathrm{Tr}\left(\gamma_n\,(1-\gamma_n)-\gamma_{0,n}^w\left(1-\gamma_{0,n}^w\right)\right)^2 \\ &\leq \mathrm{Tr}\,(\gamma-\gamma_0^w)^2\end{aligned} \tag{3.46}$$

which is the inequality we intended to show.

Using these results and Eq. (3.44), we find the following lower bound for the BCS functional

$$\begin{aligned}\mathcal{F}_\beta(\Gamma,\Gamma_0^w) \geq \int_{\mathbb{R}^3} \left(\Lambda, \left(K_{T,x}^{\Delta_0} + V_y(x)\right)\Lambda\right)_{L^2(\mathbb{R}^3,\mathrm{d}x)} \mathrm{d}y + C\left\|q\right\|_{H^1(\mathbb{R}^6)} \\ + \frac{1}{2}\mathrm{Tr}\left[Q(A+B)Q\right] + \frac{2}{5\beta}\mathrm{Tr}\left(\alpha\overline{\alpha} - \alpha_0^w\overline{\alpha_0^w}\right)^2 \\ + 2\mathrm{Re}\int_{\mathbb{R}^6} V\left(\frac{x-y}{h}\right)\Lambda(x,y)\overline{\tilde{\alpha}_0^w}(x,y)\mathrm{d}(x,y).\end{aligned} \tag{3.47}$$

In Eq. (3.47) we write $V_y(x) = V(x-y)$ and $\tilde{\alpha}_0^w(x,y) = \alpha_0^w(x,y) - h^{-3}\alpha_0\left(\frac{x-y}{h}\right)$. By assumption we know that $K_T^{\Delta_0} + V(x)$ is a positive operator and that the only elements in its kernel are functions of the form $\alpha_0\left(\frac{x-y}{h}\right)\psi(y)$. But our Cooper-pair wave functions α are symmetric under an exchange of the variables x and y, which means we can obtain more information than merely that the first term on the right-hand side of Eq. (3.47) is nonnegative.

To distinguish between the part of α lying in the kernel of $K_{T,x}^{\Delta_0} + V_y(x)$ and the part orthogonal to it, we define the function ψ by

$$\psi(y) = \frac{\int_{\mathbb{R}^3}\alpha_0\left(\frac{x-y}{h}\right)\alpha(x,y)\mathrm{d}x}{\int_{\mathbb{R}^3}|\alpha_0(x)|^2\mathrm{d}x}. \tag{3.48}$$

The full Cooper-pair wave function $\alpha(x,y)$ can be written as

$$\alpha(x,y) = h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\psi(y) + \xi_0(x,y), \tag{3.49}$$

where by definition $\int_{\mathbb{R}^3}\alpha_0\left(\frac{x-y}{h}\right)\xi_0(x,y)\mathrm{d}x = 0$ holds true. As already mentioned in Chapter 3.1.2, the above scaling assumes $\alpha_0 \sim h$ for small h and has to be adjusted if α_0 approaches zero at a faster rate. Note that the two parts of our decomposition of α, unlike α itself, are not symmetric under an exchange of the coordinates x and y. It turns out to be useful to introduce the following notation:

$$\varphi(x) = \psi(x) - 1 \quad \text{and} \quad \tilde{\alpha}_0^w(x,y) = \alpha_0^w(x,y) - h^{-3}\alpha_0\left(\frac{x-y}{h}\right). \tag{3.50}$$

For the sake of convenience, we will often write $\alpha_0\psi$ to denote the operator $\alpha_0(-ih\nabla)\psi(x)$ whose integral kernel is given by $h^{-3}\alpha_0\left(\frac{x-y}{h}\right)\psi(y)$ and the same with ψ replaced by φ. By $\kappa > 0$, we denote the second eigenvalue of the operator $K_T^{\Delta_0} + V$, which is independent of h.

Let us continue with our analysis and insert Eq. (3.49) into the first term on the right-hand side of Eq. (3.47). We obtain

$$\begin{aligned}
&\int_{\mathbb{R}^3} \left(\Lambda, \left(K_{T,x}^{\Delta_0} + V_y\right) \Lambda\right)_{L^2(\mathrm{d}x)} \mathrm{d}y \qquad (3.51)\\
&\quad = \int_{\mathbb{R}^3} \left(\alpha_0\psi + \xi_0 - \alpha_0^w, \left(K_{T,x}^{\Delta_0} + V_y\right) \left(\alpha_0\psi + \xi_0 - \alpha_0^w\right)\right)_{L^2(\mathrm{d}x)} \mathrm{d}y\\
&\quad \geq \frac{\kappa}{2} \left\|\xi_0\right\|^2_{L^2(\mathbb{R}^6)} + C_1 \left\|h\nabla_x\xi_0\right\|^2_{L^2(\mathbb{R}^6)}\\
&\qquad - 2 \left\|\left(K_{T,x}^{\Delta_0} + V_y\right)^{1/2} \xi_0\right\|_{L^2(\mathbb{R}^6)} \left\|\left(K_{T,x}^{\Delta_0} + V_y\right)^{1/2} \tilde{\alpha}_0^w\right\|_{L^2(\mathbb{R}^6)}\\
&\qquad\quad - C_2 \left\|\xi_0\right\|_{H^1_x(\mathbb{R}^6)} \left\|\tilde{\alpha}_0^w\right\|_{H^1_x(\mathbb{R}^6)}\\
&\quad \geq \frac{\kappa}{2} \left\|\xi_0\right\|^2_{L^2(\mathbb{R}^6)} + C_1 \left\|h\nabla_x\xi_0\right\|^2_{L^2(\mathbb{R}^6)} - h^{3/2}C_3 \left[\left\|\xi_0\right\|^2_{H^1_x(\mathbb{R}^6)} + 1\right]\\
&\quad \geq C_4 \left\|\xi_0\right\|^2_{H^1_x(\mathbb{R}^6)} - C_3h^{3/2}
\end{aligned}$$

for appropriately chosen constants $C_1, ..., C_4$. To come from the first to the second line, we exploit that since ξ_0 is orthogonal to the kernel of $K_{T,x}^{\Delta_0} + V_y(x)$, we can use a small part of $K_{T,x}^{\Delta_0}$ to generate the $L^2(\mathbb{R}^6)$-norm of $h\nabla_x\xi_0(x, y)$. In this step, we use $K_T^{\Delta_0} \geq C(1 + (hp)^2)$, while in the step that brings us from the second to the third line we use $K_T^{\Delta_0} \leq C\,(1 + (hp)^2)$ and Lemma 3.8. By $\|\cdot\|_{H^1_x(\mathbb{R}^6)}$ we denote that part of the $H^1(\mathbb{R}^6)$-norm where the derivatives with respect to the y-coordinate are dropped. The last inequality holds for h small enough.

On the other hand, $\int_{\mathbb{R}^3} \left(\Lambda, \left(K_{T,x}^{\Delta_0} + V_y\right) \Lambda\right)_{L^2(\mathrm{d}x)} \mathrm{d}y$ can be bounded from below in terms of
$h^2 \left\|(\nabla_x + \nabla_y)\Lambda\right\|^2_{L^2(\mathbb{R}^6)}$ as the following Lemma shows:

Lemma 3.2 *Let $\Lambda \in H^1(\mathbb{R}^6)$ be a symmetric function [$\Lambda(x, y) = \Lambda(y, x)$]. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^3} \left(\Lambda, \left(K_T^{\Delta_0} + V_y\right) \Lambda\right)_{L^2(dx)} dy \geq Ch^2 \int_{\mathbb{R}^6} \left|\left(\nabla_x + \nabla_y\right) \Lambda(x, y)\right|^2 d(x, y). \qquad (3.52)$$

Proof. The proof is nearly exactly the same as the one given in [28, Lemma 3] for a similar version of the above Lemma. The only difference is that we do not have a Fourier sum but a Fourier integral and that our α_0 is the unique ground state of $K_T^{\Delta_0} + V$ whereas the α_0 in [28] is the unique ground state of $K_{T_c} + V$. After these replacements, all arguments in the proof are the same. □

To make a connection between the decomposition $\alpha = \alpha_0\psi + \xi_0$ and Lemma 3.2, we consult Eq. (3.48) and compute

$$\nabla\psi(y) = \frac{\int_{\mathbb{R}^3} \alpha_0\left(\frac{x-y}{h}\right) \left[(\nabla_x + \nabla_y)\alpha(x, y)\right] \mathrm{d}x}{\left\|\alpha_0\right\|^2_{L^2(\mathbb{R}^3)}}. \qquad (3.53)$$

To obtain the result, we used integration by parts and the fact that $\alpha_0(-x) = \alpha_0(x)$, which is assured by Lemma 3.6. Next, we integrate Eq. (3.53) over y and apply Schwarz's inequality once which yields

$$\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} \leq h^3 \frac{\|(\nabla_x + \nabla_y)\,\alpha(x,y)\|^2_{L^2(\mathbb{R}^6)}}{\|\alpha_0\|^2_{L^2(\mathbb{R}^3)}}. \tag{3.54}$$

To replace Λ by α in Eq. (3.52), we use the triangle inequality, the fact that $(\nabla_x + \nabla_y)\alpha_0\left(\frac{x-y}{h}\right) = 0$ and Lemma 3.8:

$$\begin{aligned} \left\|\left(\nabla_x + \nabla_y\right)\Lambda(x,y)\right\|^2_{L^2(\mathbb{R}^6)} \geq & \frac{1}{2}\|(\nabla_x + \nabla_y)\,\alpha(x,y)\|^2_{L^2(\mathbb{R}^6)} \\ & - \|(\nabla_x + \nabla_y)\,\alpha_0^w(x,y)\|^2_{L^2(\mathbb{R}^6)} \\ \geq & \frac{1}{2}\|(\nabla_x + \nabla_y)\,\alpha(x,y)\|^2_{L^2(\mathbb{R}^6)} - Ch^3. \end{aligned} \tag{3.55}$$

Together with Lemma 3.2, Eq. (3.54) and Eq. (3.55) imply

$$\int_{\mathbb{R}^3} \left(\Lambda, \left(K_T^{\Delta_0} + V_y\right)\Lambda\right)_{L^2(\mathrm{d}x)} \mathrm{d}y \geq \frac{C_1\,\|\alpha_0\|^2_{L^2(\mathbb{R}^3)}}{h}\,\|\nabla\psi\|^2 - C_2 h^5 \tag{3.56}$$

with two constants $C_1, C_2 > 0$. We know that $\int_{\mathbb{R}^3}\left(\Lambda, \left(K_T^{\Delta_0} + V_y\right)\Lambda\right)_{L^2(\mathrm{d}x)} \mathrm{d}y$ controls the $L^2(\mathbb{R}^3)$-norm of the gradient of ψ as well as the $L^2(\mathbb{R}^6)$-norm of $\xi_0(x,y)$ and of $h\nabla_x\xi_0(x,y)$, see Eq. (3.51). To also control the $L^2(\mathbb{R}^6)$-norm of $h\nabla_y\xi_0(x,y)$, we use the triangle inequality and obtain

$$\begin{aligned} \|(\nabla_x + \nabla_y)\alpha(x,y)\|^2_{L^2(\mathbb{R}^6)} \geq & \frac{1}{2}\|(\nabla_x + \nabla_y)\xi_0(x,y)\|^2_{L^2(\mathbb{R}^6)} \\ & - \|\alpha_0\|^2_{L^2(\mathbb{R}^3)}\,\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)}. \end{aligned} \tag{3.57}$$

When we apply Lemma 3.7 which tells us that $\|\alpha_0\|_{L^2(\mathbb{R}^3)} \lesssim h$ and put the results of this paragraph together, we finally arrive at

$$\int_{\mathbb{R}^3} \left(\Lambda, \left(K_T^{\Delta_0} + V_y\right)\Lambda\right)_{L^2(\mathrm{d}x)} \mathrm{d}y \geq C_1\left(h\,\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + \|\xi_0\|^2_{H^1(\mathbb{R}^6)}\right) - C_2 h^{3/2} \tag{3.58}$$

for some constants $C_1, C_2 > 0$ and h small enough. Note that we use the notation $a \lesssim b$ which is equivalent to saying that there exists a constant $C > 0$ such that $a \leq Cb$.

At the end of step 1, let us summarize what we have achieved. Insertion of Eq. (3.58) into Eq. (3.47) gives

$$\begin{aligned} \mathcal{F}_\beta(\Gamma, \Gamma_0^w) \geq & C_1\left(h\,\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + \|\xi_0\|^2_{H^1(\mathbb{R}^6)} + \|q\|^2_{H^1(\mathbb{R}^6)}\right) + \frac{1}{2}\mathrm{Tr}\left[Q(A+B)Q\right] \\ & + \frac{2}{5\beta}\mathrm{Tr}\left(\alpha\overline{\alpha} - \alpha_0^w\overline{\alpha_0^w}\right)^2 + 2\mathrm{Re}\int_{\mathbb{R}^6} V\left(\frac{x-y}{h}\right)\Lambda(x,y)\overline{\tilde{\alpha}_0^w}(x,y)\mathrm{d}(x,y) - C_2 h^{3/2}, \end{aligned} \tag{3.59}$$

which again holds for appropriately chosen constants $C_1, C_2 > 0$ and h small enough.

3.2.2 Step 2: Bounds on the remaining non-positive terms

In step 2 we derive bounds for the remaining non-positive terms, that is, the fourth term in the first line on the right-hand side of Eq. (3.59) and the second term in the second line of the same equation. We start with the term proportional to A. In order to give a good estimate for this term, we express the trace in terms of operator kernels. This is crucial because it allows us to capture the effect that W is a function that goes to zero at infinity.

Let us define the operator $\tilde{A}$ by

$$\begin{aligned} \tilde{A} &= h^2 \sum_{n=1}^{\infty} 2c^2 n^2 \frac{1}{E^2 + c^2 n^2} \left(\omega H_0 + H_0 \omega\right) \frac{1}{E^2 + c^2 n^2} \\ &= h^2 \sum_{n=1}^{\infty} 2c^2 n^2 \frac{1}{E^2 + c^2 n^2} \begin{pmatrix} \{W, k\} & [W, \Delta_0] \\ [\Delta_0, W] & \{W, k\} \end{pmatrix} \frac{1}{E^2 + c^2 n^2}. \end{aligned} \tag{3.60}$$

By $[T_1, T_2] = T_1 T_2 - T_2 T_1$ and $\{T_1, T_2\} = T_1 T_2 + T_2 T_1$ we denote the commutator and the anti-commutator of the operators T_1 and T_2, respectively. The matrix elements of $\tilde{A}$ will be called $\tilde{a}_{ij}$ for $i, j = 1, 2$. Since ω^2 is a positive operator we have $\mathrm{Tr}\,[QAQ] \geq \mathrm{Tr}\left[Q\tilde{A}Q\right]$. In order to get rid of the matrix structure, we compute

$$\begin{aligned} \mathrm{Tr}\left[Q\tilde{A}Q\right] &= \mathrm{Tr}\big[\, q\tilde{a}_{11}q + \Lambda\tilde{a}_{21}q + q\tilde{a}_{12}\overline{\Lambda} + \Lambda\tilde{a}_{22}\overline{\Lambda} \\ &\qquad + \overline{\Lambda}\tilde{a}_{11}\Lambda - \overline{q}\tilde{a}_{21}\Lambda - \overline{\Lambda}\tilde{a}_{12}\overline{q} + \overline{q}\tilde{a}_{22}\overline{q}\,\big] \\ &= 2\mathrm{Tr}\,[q\tilde{a}_{11}q] + 2\mathrm{Tr}\left[\overline{\Lambda}\tilde{a}_{11}\Lambda\right] + 4\mathrm{ReTr}\,[\Lambda\tilde{a}_{12}q]\,. \end{aligned} \tag{3.61}$$

To come from the first to the second line, we use the symmetry properties of the kernels of q and Λ, namely $\hat{q}(r,s) = \overline{\hat{q}(s,r)}$ and $\hat{\Lambda}(r,s) = \hat{\Lambda}(s,r)$ where the hat denotes the Fourier transform. From the definition of $\tilde{A}$, we can easily read off the kernels of $\tilde{a}_{11}$ and $\tilde{a}_{12}$. In Fourier space they are given by

$$\begin{aligned} \tilde{a}_{11}(p,q) &= h^2 \hat{W}(p-q)\,[k(hp) + k(hq)] \sum_{n=1}^{\infty} \frac{2c^2 n^2}{E(hp)^2 + c^2 n^2} \frac{1}{E(hq)^2 + c^2 n^2}, \\ \tilde{a}_{12}(p,q) &= h^2 \hat{W}(p-q) \left[\hat{\Delta}_0(hp) - \hat{\Delta}_0(hq)\right] \sum_{n=1}^{\infty} \frac{2c^2 n^2}{E(hp)^2 + c^2 n^2} \frac{1}{E(hq)^2 + c^2 n^2}. \end{aligned} \tag{3.62}$$

In what follows, we write $\zeta(p,q) = \sum_{n=1}^{\infty} \frac{2c^2 n^2}{E(p)^2 + c^2 n^2} \frac{1}{E(q)^2 + c^2 n^2}$. Due to its simple structure, the infinite sum in the definition of ζ can be computed explicitly. The result of this computation as well as some properties of the function ζ and of the operators $\tilde{a}_{11}$ and $\tilde{a}_{12}$ are summarized in the following Lemma.

Lemma 3.3 *The function $\zeta(p,q)$ is given by*

$$\zeta(p,q) = \frac{1}{E(p)+E(q)}\bigg\{ -\frac{\pi}{c^2} + \frac{2\pi/c}{E(p)/c - E(q)/c} \times \left[\frac{E(p)/c}{1-e^{-2\pi E(p)/c}} - \frac{E(q)/c}{1-e^{-2\pi E(q)/c}}\right]\bigg\}. \tag{3.63}$$

It has the following properties: The norms $\|(1+p^2)\zeta(p,q)\|_{L^\infty(\mathbb{R}^6)}$ and $\|\zeta(p,q)\|_{W^{4,\infty}(\mathbb{R}^6)}$ are finite. Additionally, the operators $\tilde{a}_{11}$, $x^m\tilde{a}_{11}$ for $0 \le m \le 2$ and $(1+x^2)\tilde{a}_{11}(1+x^2)$ are bounded from $L^2(\mathbb{R}^3)$ to $L^2(\mathbb{R}^3)$ with operator norm of order h^2. The same holds true if $\tilde{a}_{11}$ is replaced by $\tilde{a}_{12}$ except that the norms are of order h^3 in this case.

Proof. To evaluate the infinite sum in the definition of ζ, we use the Poisson summation formula, see e.g. [51]. Let us define $f(x) = \frac{x^2}{(a^2+x^2)(b^2+x^2)}$. Its Fourier transform can be computed explicitly and reads [50, p. 448]

$$\hat{f}(k) = \frac{\pi}{a^2-b^2}\left(ae^{-a2\pi|k|} - be^{-b2\pi|k|}\right). \tag{3.64}$$

The Poisson summation formula tells us that $\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$. In our case, the sum over the Fourier transform can easily be evaluated and together with the Poisson Summation formula leads to

$$\sum_{n=1}^{\infty} \frac{n^2}{(a^2+n^2)(b^2+n^2)} = -\frac{\pi}{2(a+b)} + \frac{\pi}{a^2-b^2}\left(\frac{a}{1-e^{-2\pi a}} - \frac{b}{1-e^{-2\pi b}}\right). \tag{3.65}$$

Using Eq. (3.65) one easily establishes Eq. (3.63). This concludes the first part of the proof.

To see that $\|(1+p^2)\zeta(p,q)\|_{L^\infty(\mathbb{R}^6)} < \infty$ holds, we write $\zeta(p,q)$ as

$$\zeta(p,q) = \frac{1}{E(p)+E(q)}\bigg\{ -\frac{\pi}{c} + \frac{2\pi}{(1+e^{-\beta E(p)})(1+e^{-\beta E(q)})} \times \left[1 + e^{-\beta E(p)} - e^{-\beta E(p)}E(p)\frac{e^{-\beta(E(q)-E(p))}-1}{E(q)-E(p)}\right]\bigg\}. \tag{3.66}$$

The term in the curly brackets can be bounded by

$$\begin{aligned} &\bigg\| -\frac{\pi}{c} + \frac{2\pi}{(1+e^{-\beta E(p)})(1+e^{-\beta E(q)})} \\ &\quad \times \left[1 + e^{-\beta E(p)} - e^{-\beta E(p)}E(p)\frac{e^{-\beta(E(q)-E(p))}-1}{E(q)-E(p)}\right]\bigg\|_{L^\infty(\mathbb{R}^6)} \\ &\le \frac{\pi}{c} + 4\pi + 2\pi\left\|e^{-\beta E(p)}E(p)\frac{e^{-\beta(E(q)-E(p))}-1}{E(q)-E(p)}\right\|_{L^\infty(\mathbb{R}^6)}, \end{aligned} \tag{3.67}$$

where the norm on the right-hand side of Eq. (3.67) is readily seen to be finite. Let us split $\|(1+p^2)\zeta(p,q)\|_{L^\infty(\mathbb{R}^6)}$ into two parts. If $p^2 \leq 2\mu$ the function $E(p)$ may have zeros if $\hat{\Delta}_0(p) = 0$ and $p^2 = \mu$. On this set, we use $1 + p^2 \leq 1 + 2\mu$, which together with the series representation of $\zeta(p,q)$, yields the desired estimate. On the other hand, for $p^2 > 2\mu$ we can use Eqs. (3.66), (3.67) and the estimate

$$\left\| \chi_{\{x^2>2\mu\}}(p) \frac{1+p^2}{E(p)+E(q)} \right\|_{L^\infty(\mathbb{R}^6)} < \infty, \tag{3.68}$$

where $\chi_{\{x^2>2\mu\}}(p)$ denotes the characteristic function of the set $\{x \in \mathbb{R}^3 \mid x^2 > 2\mu\}$ to prove the claim. The norm in Eq. (3.68) is finite since $E(p) \sim p^2$ for large $|p|$. To show that $\|\zeta(p,q)\|_{W^{4,\infty}(\mathbb{R}^6)} < \infty$, we go back to the definition of $\zeta(p,q)$ as an infinite sum. We have $|\nabla_p (E(p)^2 + c^2n^2)^{-1}| = |2E(p)(\nabla_p E(p))(E(p)^2 + c^2n^2)^{-2}| \lesssim (E(p)^2 + c^2n^2)^{-1}$, which can easily be seen if we keep in mind that $E(p) \sim p^2$ for large $|p|$. Similarly, one can show that $|\partial_{p_i}^m (E(p)^2 + c^2n^2)^{-1}| \lesssim (E(p)^2 + c^2n^2)^{-1}$ for $i = 1,2,3$ and $1 \leq m \leq 4$ which is enough to prove the claim.

To show that the kernel $\tilde{a}_{11}(p,q) = \hat{W}(p-q)\,[k(hp)+k(hq)]\,\zeta(hp,hq)$ defines a bounded operator on $L^2(\mathbb{R}^3)$, we estimate

$$\begin{aligned}
&\left(\int_{\mathbb{R}^3} \left| \int_{\mathbb{R}^3} h^2 \hat{W}(p-q)\,[k(hp)+k(hq)]\,\zeta(hp,hq)\Psi(q)\mathrm{d}q \right|^2 \right)^{1/2} \qquad (3.69)\\
&\leq \|[k(hp)+k(hq)]\,\zeta(hp,hq)\|_{L^\infty(\mathbb{R}^6)} \left(\int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} \left| h^2 \hat{W}(p-q)\Psi(q) \right| \mathrm{d}q \right)^2 \mathrm{d}p \right)^{1/2} \\
&\lesssim h^2 \, \|[k(hp)+k(hq)]\,\zeta(hp,hq)\|_{L^\infty(\mathbb{R}^6)} \left\| \hat{W} \right\|_{L^1(\mathbb{R}^3)} \|\Psi\|_{L^2(\mathbb{R}^3)}\,.
\end{aligned}$$

To come from the second to the third line, we used Young's inequality. The boundedness of $\|[k(hp)+k(hq)]\,\zeta(hp,hq)\|_{L^\infty(\mathbb{R}^6)}$ is assured by the identity $\zeta(p,q) = \zeta(q,p)$ and the boundedness of $\|(1+p^2)\zeta(p,q)\|_{L^\infty(\mathbb{R}^6)}$.

To give a bound on the operator norm of $x\tilde{a}_{11}$, we compute its integral kernel which is given by (in Fourier space x acts as $i\nabla$)

$$\begin{aligned}
(x\tilde{a}_{11})(p,q) = {} & ih^2 \left(\nabla \hat{W}\right)(p-q)\,[k(hp)+k(hq)]\,\zeta(hp,hq) \qquad (3.70)\\
& + ih^3 \hat{W}(p-q)\,(\nabla k)(hp)\zeta(hp,hq) \\
& + ih^2 \hat{W}(p-q)\,[k(hp)+k(hq)]\,\nabla_p \zeta(hp,hq).
\end{aligned}$$

When we use an estimate similar to the one in Eq. (3.69), it can easily be seen that $\|x\tilde{a}_{11}\|_\infty \lesssim h^2$, where by $\|\cdot\|_\infty$ we denote the operator norm on $\mathcal{L}(L^2(\mathbb{R}^3))$. Since a second derivative does not make things worse the same strategy applies to the operator $x^2\tilde{a}_{11}$ and yields $\|x^2\tilde{a}_{11}\|_\infty \lesssim h^2$. The argument showing that

$$\left\|(1+x^2)\tilde{a}_{11}(1+x^2)\right\|_\infty \lesssim h^2 \tag{3.71}$$

is true follows the same lines. The only difference is that x^2 may stand on the right-hand side of $\tilde{a}_{11}$. If we write the norm of the operator $\tilde{a}_{11}x^2$ for example in terms of the integral kernel of $\tilde{a}_{11}$ like in Eq. (3.69) the operator x^2 acts in Fourier space as $-\Delta_q$ on the function $\Psi(q)$. In order to proceed as before, one has to integrate by parts to let the derivatives act on $\tilde{a}_{11}(p,q)$ instead of on the function $\Psi(q)$. The rest of the argument is the same as before. If one considers $\tilde{a}_{12}$ instead of $\tilde{a}_{11}$ the above proofs go through as before. To obtain the additional factor of h in the estimate one has to use that $\left\|\hat{\Delta}_0\right\|_{L^\infty(\mathbb{R}^3)} \lesssim h$ which is assured by Lemma 3.7. This ends the proof of Lemma 3.3. □

Using Lemma 3.3, we can evaluate the traces in Eq. (3.61) in terms of kernels. Let us start with the first term on the right-hand side, which can be estimated by:

$$|\mathrm{Tr}\,[q\tilde{a}_{11}q]| \leq \|\tilde{a}_{11}\|_\infty \|q\|_2^2 \lesssim h^2 \|q\|_2^2 . \tag{3.72}$$

Here and in the following, we denote by $\|T\|_p = \left(\mathrm{Tr}\,(T^*T)^{p/2}\right)^{1/p}$, $p \geq 1$, the p-th Schatten class-norm of the operator T. In order to estimate the second term on the right-hand side of Eq. (3.61), we use the decomposition of α introduced in Eq. (3.49) and write

$$\begin{aligned}\mathrm{Tr}\left[\overline{\Lambda}\tilde{a}_{11}\Lambda\right] = \mathrm{Tr}\,[\overline{\varphi}\alpha_0\tilde{a}_{11}\alpha_0\varphi] &+ 2\mathrm{ReTr}\,[\overline{\varphi}\alpha_0\tilde{a}_{11}\,(\xi_0 - \tilde{\alpha}_0^w)] \\ &+ \mathrm{Tr}\,[(\xi_0 - \tilde{\alpha}_0^w)\,\tilde{a}_{11}\,(\xi_0 - \tilde{\alpha}_0^w)]\,,\end{aligned} \tag{3.73}$$

where $\varphi(x) = \psi(x) - 1$. The last term on the right-hand side of Eq. (3.73) can be estimated like the one in Eq. (3.72), which gives

$$\begin{aligned}\left|\mathrm{Tr}\left[\left(\xi_0^* + \overline{\tilde{\alpha}}_0^w\right)\tilde{a}_{11}\,(\xi_0 + \tilde{\alpha}_0^w)\right]\right| &\lesssim h^2\left(\|\xi_0\|_{L^2(\mathbb{R}^6)} + \|\tilde{\alpha}_0^w\|_{L^2(\mathbb{R}^6)}\right)^2 \\ &\lesssim h^2\left(\|\xi_0\|_{L^2(\mathbb{R}^6)}^2 + h^{3/2}\right).\end{aligned} \tag{3.74}$$

To come from the first to the second line, we used Lemma 3.8. The term proportional to $\alpha_0\varphi$, that is, the second term on the right-hand side of Eq. (3.73), has to be estimated differently because we will not be able to control the $L^2(\mathbb{R}^3)$-norm of φ. Using the positive terms in Eq. (3.59), we will be able to dominate a term of the form $\int_{\mathbb{R}^3}\left(|\psi(x)|^2 - 1\right)g(x)\mathrm{d}x$ where $g(x)$ is a reasonably localized function. We will encounter expressions like this frequently in the following and so we introduce the notation $\Phi(x) = |\psi(x)|^2 - 1$. To bring the second term on the right-hand side of

Eq. (3.73) in this form, we compute

$$
\begin{aligned}
\left|2\mathrm{ReTr}\left[\overline{\varphi}\alpha_0\tilde{\alpha}_{11}\left(\xi_0-\tilde{\alpha}_0^w\right)\right]\right| \leq{}& 2\left|\mathrm{Tr}\left[\left(\frac{\overline{\varphi(x)}}{1+x^2}\right)\hat{\alpha}_0(-ih\nabla)\left(1+x^2\right)\tilde{\alpha}_{11}\left(\xi_0-\tilde{\alpha}_0^w\right)\right]\right| \\
&+2\left|\mathrm{Tr}\left[\left(\frac{\overline{\varphi(x)}}{1+x^2}\right)\left[x^2,\hat{\alpha}_0(-ih\nabla)\right]\tilde{\alpha}_{11}\left(\xi_0-\tilde{\alpha}_0^w\right)\right]\right| \qquad (3.75)\\
\leq{}& 2\left\|\frac{\overline{\varphi(x)}}{1+x^2}\hat{\alpha}_0(-ih\nabla)\right\|_2\left\|(1+x^2)\tilde{\alpha}_{11}\right\|_\infty\left(\|\xi_0\|_2+\|\tilde{\alpha}_0^w\|_2\right)\\
&+2\left\|\frac{\overline{\varphi}(x)}{1+x^2}h(\nabla\hat{\alpha}_0)(-ih\nabla)\right\|_2\left\|x\tilde{\alpha}_{11}\right\|_\infty\left(\|\xi_0\|_2+\|\tilde{\alpha}_0^w\|_2\right)\\
&+2\left\|\frac{\overline{\varphi}(x)}{1+x^2}h^2(\Delta\hat{\alpha}_0)(-ih\nabla)\right\|_2\left\|\tilde{\alpha}_{11}\right\|_\infty\left(\|\xi_0\|_2+\|\tilde{\alpha}_0^w\|_2\right)
\end{aligned}
$$

Because of Lemma 3.3, we already know that all norms containing $\tilde{a}_{11}$ in Eq. (3.75) are of order h^2. To estimate the terms proportional to φ, we use the Seiler-Simon inequality, see e.g. [52, Theorem 4.1], which in the special case we need tells us that $\|a(x)b(-i\nabla)\|_2 \leq (2\pi)^{-3/2}\|a\|_{L^2(\mathbb{R}^3)}\|b\|_{L^2(\mathbb{R}^3)}$ for two functions $a,b\in L^2(\mathbb{R}^3)$. Together with the estimate $\|\hat{\alpha}_0\|_{H^2(\mathbb{R}^3)} \lesssim h$, which is assured by Lemma 3.7, this allows us to obtain

$$
\left|2\mathrm{ReTr}\left[\overline{\varphi}\alpha_0\tilde{a}_{11}\left(\xi_0-\tilde{\alpha}_0^w\right)\right]\right| \lesssim h^{3/2}\left(\int_{\mathbb{R}^3}\frac{|\varphi(x)|^2}{1+x^4}\mathrm{d}x\right)^{1/2}\left(\|\xi_0\|_2+\|\tilde{\alpha}_0^w\|_2\right). \qquad (3.76)
$$

This is not yet what we wanted to show, but it is close as the following computation shows:

$$
\begin{aligned}
\int_{\mathbb{R}^3}\frac{|\varphi(x)|^2}{1+x^4}\mathrm{d}x &= \int_{\mathbb{R}^3}\frac{\Phi(x)-2\mathrm{Re}\left[\psi(x)-1\right]}{1+x^4}\mathrm{d}x \qquad (3.77)\\
&\leq \int_{\mathbb{R}^3}\frac{\Phi(x)}{1+x^4}\mathrm{d}x+2\left(\int_{\mathbb{R}^3}\frac{|\psi(x)|^2}{1+x^4}\mathrm{d}x\right)^{1/2}\left(\int_{\mathbb{R}^3}\frac{1}{1+x^4}\mathrm{d}x\right)^{1/2}\\
&\qquad +2\left(\int_{\mathbb{R}^3}\frac{1}{1+x^4}\mathrm{d}x\right)^{1/2}\\
&\leq 2\int_{\mathbb{R}^3}\frac{\Phi(x)}{1+x^4}\mathrm{d}x+2\left(\int_{\mathbb{R}^3}\frac{1}{1+x^4}\mathrm{d}x\right)+2\left(\int_{\mathbb{R}^3}\frac{1}{1+x^4}\mathrm{d}x\right)^{1/2}.
\end{aligned}
$$

Insertion of the result from Eq. (3.77) into Eq. (3.76) together with an application of Lemma 3.8 yields

$$
\begin{aligned}
\left|2\mathrm{ReTr}\overline{\varphi}\alpha_0\tilde{a}_{11}\left(\xi_0-\tilde{\alpha}_0^w\right)\right| &\lesssim h^{3/2}\left[\left(\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p\right)^{1/2}+1\right]\left(\|\xi_0\|_2+h^{3/2}\right)\\
&\lesssim h\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p+h^2\|\xi_0\|_2^2+h, \qquad (3.78)
\end{aligned}
$$

where by $\widehat{\frac{1}{1+(\cdot)^4}}(p)$ we denote the Fourier transform of the function $x \mapsto (1+x^4)^{-1}$. To come to the last line, we used the inequality $ab \leq \frac{1}{2}(a^2+b^2)$ for two numbers $a, b \in \mathbb{R}$.

It remains to give a similar bound for the first term on the right-hand side of Eq. (3.73), which can be done in the same spirit as the previous estimate. We compute

$$\begin{aligned} |\mathrm{Tr}\,[\overline{\varphi}\alpha_0\tilde{a}_{11}\alpha_0\varphi]| = & \left|\mathrm{Tr}\left[\frac{\overline{\varphi}(x)}{1+x^2}\hat{\alpha}_0(-ih\nabla)(1+x^2)\tilde{a}_{11}(1+x^2)\hat{\alpha}_0(-ih\nabla)\frac{\varphi(x)}{1+x^2}\right]\right| \\ & + \left|\mathrm{Tr}\left[\frac{\overline{\varphi}(x)}{1+x^2}\left[x^2, \hat{\alpha}_0(-ih\nabla)\right]\tilde{a}_{11}(1+x^2)\hat{\alpha}_0(-ih\nabla)\frac{\varphi(x)}{1+x^2}\right]\right| \\ & + \left|\mathrm{Tr}\left[\frac{\overline{\varphi}(x)}{1+x^2}\hat{\alpha}_0(-ih\nabla)(1+x^2)\tilde{a}_{11}\left[x^2, \hat{\alpha}_0(-ih\nabla)\right]\frac{\varphi(x)}{1+x^2}\right]\right| \\ & + \left|\mathrm{Tr}\left[\frac{\overline{\varphi}(x)}{1+x^2}\left[x^2, \hat{\alpha}_0(-ih\nabla)\right]\tilde{a}_{11}\left[x^2, \hat{\alpha}_0(-ih\nabla)\right]\frac{\varphi(x)}{1+x^2}\right]\right|. \end{aligned} \tag{3.79}$$

Let us have a closer look at the first term on the right-hand side of Eq. (3.79). It can be estimated by

$$\begin{aligned} & \left|\mathrm{Tr}\left[\frac{\overline{\varphi}(x)}{1+x^2}\hat{\alpha}_0(-ih\nabla)(1+x^2)\tilde{a}_{11}(1+x^2)\hat{\alpha}_0(-ih\nabla)\frac{\varphi(x)}{1+x^2}\right]\right| \\ & \qquad \leq \left\|\hat{\alpha}_0(-ih\nabla)\frac{\varphi(x)}{1+x^2}\right\|_2^2 \left\|(1+x^2)\tilde{a}_{11}(1+x^2)\right\|_\infty \\ & \qquad \lesssim h\int_{\mathbb{R}^3}\frac{|\varphi(x)|^2}{1+x^4}\mathrm{d}x \\ & \qquad \lesssim h\left(\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p + 1\right). \end{aligned} \tag{3.80}$$

To come to the third line, we used Lemma 3.3, the Seiler-Simon inequality, Lemma 3.7 and the computation carried out in Eq. (3.77).

Let us come to the last term on the right-hand side of Eq. (3.61). Using the same techniques as in the estimate of the other two terms, we find

$$|\mathrm{Tr}\,[\Lambda\tilde{a}_{12}q]| \lesssim h\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p + h^2\left(\|q\|_2^2 + \|\xi_0\|_2^2\right) + h. \tag{3.81}$$

Putting our results together, we obtain an upper bound for $|\mathrm{Tr}\,[QAQ]|$. It reads

$$|\mathrm{Tr}\,[QAQ]| \lesssim h\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p + h^2\left(\|\xi_0\|_2^2 + \|q\|_2^2\right) + h. \tag{3.82}$$

We note that the first term on the right-hand side of Eq. (3.82) is proportional to $|\psi|^2$ and of the same order in h as the contribution proportional to $\|\nabla\psi\|_{L^2(\mathbb{R}^3)}^2$ in Eq. (3.59).

Let us continue our analysis and derive a bound for the term $\mathrm{Tr}\,[QBQ]$ that is similar to the one given in Eq. (3.82) for $\mathrm{Tr}\,[QAQ]$. We abbreviate $\eta = H_0 h^2\omega + h^2\omega H_0 + h^4\omega^2$ and compute

$$
\begin{aligned}
|\mathrm{Tr}\,[QBQ]| &= \left|\mathrm{Tr}\left[Q\left(\sum_{n=1}^{\infty}\frac{1}{E^2+c^2n^2}\eta\frac{2c^2n^2}{(H_0^w)^2+c^2n^2}\eta\frac{1}{E^2+c^2n^2}\right)Q\right]\right| \qquad (3.83)\\
&\leq \sum_{n=1}^{\infty}\left\|Q\frac{1}{E^2+c^2n^2}\eta\frac{2c^2n^2}{(H_0^w)^2+c^2n^2}\eta\frac{1}{E^2+c^2n^2}Q\right\|_1\\
&\leq \sum_{n=1}^{\infty}\left\|Q\frac{1}{E^2+c^2n^2}\eta\right\|_2\left\|\frac{2c^2n^2}{(H_0^w)^2+c^2n^2}\right\|_\infty\left\|\eta\frac{1}{E^2+c^2n^2}Q\right\|_2\\
&\lesssim \sum_{n=1}^{\infty}\left\|Q\frac{1}{E^2+c^2n^2}\left(H_0h^2\omega+h^2\omega H_0+h^4\omega^2\right)\right\|_2^2
\end{aligned}
$$

We first give a bound on the term proportional to $h^4\omega^2$:

$$
\begin{aligned}
\sum_{n=1}^{\infty}\left\|Q\frac{1}{E^2+c^2n^2}h^4\omega^2\right\|_2^2 &\lesssim h^8\,\|Q\|_\infty^2\sum_{n=1}^{\infty}\left\|\frac{1}{E^2+c^2n^2}W\right\|_2^2 \qquad (3.84)\\
&\lesssim h^8\,\|W\|_{L^2(\mathbb{R}^3)}^2\sum_{n=1}^{\infty}\left\|\frac{1}{E^2+c^2n^2}\right\|_{L^2(\mathbb{R}^3)}^2\\
&\lesssim h^5\sum_{n=1}^{\infty}\frac{1}{n^{5/2}}.
\end{aligned}
$$

The result in the second line follows from the Seiler-Simon inequality and $|Q|\leq 1$. To come to the last line, we need the estimate $\|(E^2+c^2n^2)^{-1}\|_{L^2(\mathbb{R}^3)}\lesssim h^{-3/2}n^{-5/4}$, which can be derived in a straight-forward way.

Next, we treat the terms in Eq. (3.83) that are proportional to $H_0h^2\omega$. To that end, we write the Hilbert-Schmidt norm in terms of the matrix elements of our $\mathbb{C}^2$-matrix structure and use the symmetries of the kernels of q and Λ which gives

$$
\begin{aligned}
\left\|Q\frac{1}{E^2+c^2n^2}h^2\{H_0,\omega\}\right\|_2^2 &= 2\Bigg(\left\|q\frac{1}{E^2+c^2n^2}h^2\{k,W\}\right\|_2^2+\left\|\Lambda\frac{1}{E^2+c^2n^2}[W,\Delta_0]\right\|_2^2\\
&\quad+\left\|q\frac{1}{E^2+c^2n^2}[W,\Delta_0]\right\|_2^2+\left\|\Lambda\frac{1}{E^2+c^2n^2}h^2\{k,W\}\right\|_2^2\Bigg). \qquad (3.85)
\end{aligned}
$$

A bound for the terms proportional to q is readily obtained:

$$
\begin{aligned}
\left\|q\frac{1}{E^2+c^2n^2}h^2\{k,W\}\right\|_2 &\leq \|q\|_2\left\|\frac{1}{E^2+c^2n^2}h^2\{k,W\}\right\|_\infty \qquad (3.86)\\
&\lesssim \frac{h^2}{n}\,\|q\|_2\,.
\end{aligned}
$$

To come to the last line, we used that $\left\|\frac{k}{E^2+c^2n^2}\right\|_{L^\infty(\mathbb{R}^3)} \leq \left\|\frac{E}{E^2+c^2n^2}\right\|_{L^\infty(\mathbb{R}^3)} \lesssim \frac{1}{n}$. Note that this is sufficient to guarantee the convergence of the infinite sum in Eq. (3.83). A similar bound with n^{-1} replaced by n^{-2} and h^2 replaced by h^3 holds for the third term on the right-hand side of Eq. (3.85). To estimate the fourth term on the right-hand side of Eq. (3.85) we decompose $\Lambda = \alpha_0\varphi + \xi_0 - \tilde{\alpha}_0^w$ and consider only the term proportional to $\alpha_0\psi$. The other contributions are treated like the ones proportional to q. We obtain

$$\begin{aligned} \left\|\alpha_0\varphi\frac{1}{E^2+c^2n^2}h^2\left\{k,W\right\}\right\|_2 &\leq \left\|\alpha_0\frac{\varphi}{1+x^2}\right\|_2\left\|(1+x^2)\frac{1}{E^2+c^2n^2}h^2\left\{k,W\right\}\right\|_\infty \\ &\lesssim \frac{h^{3/2}}{n}\left[\left(\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+x^4}}(p)\right|\mathrm{d}p\right)^{1/2}+1\right]. \end{aligned} \tag{3.87}$$

To come from the first to the second line, we use the Seiler-Simon inequality and commute the operator $(1+x^2)$ to the right until it stands next to $W(x)$. As before, we need $\left\|\frac{E}{E^2+c^2n^2}\right\|_{L^\infty(\mathbb{R}^3)} \lesssim \frac{1}{n}$. The same bound with n^{-1} replaced by n^{-2} and with h^2 replaced by h^3 is obtained for the second term on the right-hand side of Eq. (3.85). Putting our estimates together and again using Lemma 3.8, we finally find

$$|\mathrm{Tr}QBQ| \lesssim h^3\left[\left(\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p\right)+1\right]+Ch^4\left(\|q\|_2^2+\|\xi_0\|_2^2\right)+h^3. \tag{3.88}$$

This ends the construction of a bound for the term $\mathrm{Tr}\left[Q(A+B)Q\right]$.

The remaining non-positive term to estimate is the last term on the right-hand side of Eq. (3.59). Along the lines of the previous estimates, we compute

$$\begin{aligned} &\left|\int_{\mathbb{R}^6}V\left(\frac{x-y}{h}\right)(\alpha-\alpha_0^w)(x,y)\overline{\tilde{\alpha}}_0^w(x,y)\mathrm{d}(x,y)\right| \\ &\quad\lesssim \left|\int_{\mathbb{R}^6}h^{-3}\Delta_0\left(\frac{x-y}{h}\right)\varphi(y)\overline{\tilde{\alpha}}_0^w(x,y)\mathrm{d}(x,y)\right| \\ &\qquad\qquad + \|V\|_{L^\infty(\mathbb{R}^3)}\left(\|\xi_0\|_2+\|\tilde{\alpha}_0^w\|_2\right)\|\tilde{\alpha}_0^w\|_2 \\ &\quad\lesssim \left|\mathrm{Tr}\Delta_0\frac{\varphi}{1+x^2}(1+x^2)\overline{\tilde{\alpha}}_0^w\right|+h^{3/2}\|\xi_0\|_2^2+h^3 \\ &\quad\lesssim h^{-1/2}\left[\left(\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p\right)^{1/2}+1\right]\left\|(1+x^2)\tilde{\alpha}_0^w\right\|_2+h^{3/2}\|\xi_0\|_2^2+h^3 \\ &\quad\lesssim h\int_{\mathbb{R}^3}\left|\hat{\Phi}(p)\widehat{\frac{1}{1+(\cdot)^4}}(p)\right|\mathrm{d}p+h^{3/2}\|\xi_0\|_2+h \end{aligned} \tag{3.89}$$

To come to the third line, we used Lemma 3.9. The rest of the computation is carried out like the ones we did to estimate $\mathrm{Tr}\left[QAQ\right]$.

When we put our estimates for the non-positive terms in Eq. (3.59) together, that is, Eq. (3.82), Eq. (3.88) and Eq. (3.89), we arrive at the following lower bound for the BCS functional:

$$\begin{aligned}\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \geq C_1 h \Big(\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + \|\xi_0\|^2_{H_1(\mathbb{R}^6)} + \|q\|^2_{H^1(\mathbb{R}^6)} \Big) \\ + \frac{2}{5\beta} \mathrm{Tr}\,(\alpha\overline{\alpha} - \alpha_0^w \overline{\alpha}_0^w)^2 - hC_2 \int_{\mathbb{R}^3} \left| \hat{\Phi}(p) \widehat{\frac{1}{1+(\cdot)^4}}(p) \right| \mathrm{d}p - C_3 h.\end{aligned} \tag{3.90}$$

Eq. (3.90) holds for appropriately chosen constants $C_1, C_2, C_3 > 0$ and h small enough. This ends the second step.

3.2.3 Step 3: Construction of the lower bound and a-priori estimates

The construction of the lower bound starting from Eq. (3.86) needs one crucial ingredient - estimates of the size of $\|\nabla\psi\|_{L^2(\mathbb{R}^3)}$ and $\|\xi_0\|_{L^2(\mathbb{R}^6)}$. We will see that the condition $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \leq 0$ is strong enough to guarantee that $\|\nabla\psi\|_{L^2(\mathbb{R}^3)}$ is of order 1 with respect to $\|\hat{\Phi}\|_{L^2(B_r)}$ and that $\|\xi_0\|_{L^2(\mathbb{R}^6)}$ is of order $h^{1/2}$ with respect to the same norm. Here, B_r denotes the ball of radius r centered at zero with $r > 0$ chosen large enough but finite. This guarantees a separation of scales with respect to the decomposition $\alpha = \alpha_0\psi + \xi_0$. Our analysis starts with the following lemma

Lemma 3.4 *Let $g : \mathbb{R}^3 \to \mathbb{R}_+$ be a measurable function, $C_1 > 0$ and assume that for some $\beta \in \mathbb{R}$ with $\beta > 2$ one has $\left\|(1+|p|^\beta)g\right\|_{L^1(\mathbb{R}^3)} + \left\|(1+|p|^\beta)g\right\|_{L^\infty(\mathbb{R}^3)} < \infty$. Then there exists $R > 0$ such that for all $r \geq R$ one has*

$$\int_{\mathbb{R}^3} |\nabla\psi(x)|^2\, dx - C_1 \int_{B_r^c} \left|\hat{\Phi}(p)\right| g(p)\, dp \geq -C(r). \tag{3.91}$$

In the above equation $B_r \subset \mathbb{R}^3$ denotes the ball with radius r centered at zero and the superscript c stands for complement. The constant $C(r) > 0$ satisfies $\lim_{r\to\infty} C(r) = 0$.

Proof. We start by expressing the above quantities in terms of $\widehat{|\psi| - 1}$, the Fourier transform of $|\psi| - 1$. This can be done because $(|\psi|^2 - 1) = (|\psi| - 1)^2 + 2(|\psi| - 1)$. An application of the triangle inequality yields

$$\begin{aligned}\int_{B_r^c} \left|\hat{\Phi}(p)\right| g(p)\mathrm{d}p &\leq \int_{B_r^c} \left|\widehat{|\psi| - 1} * \widehat{|\psi| - 1}(p)\right| g(p)\mathrm{d}p + 2\int_{B_r^c} \left|\widehat{|\psi| - 1}(p)\right| g(p)\mathrm{d}p \\ &\leq \frac{1}{1+r^\beta} \int_{B_r^c} \left| (\widehat{|\psi| - 1}) * (\widehat{|\psi| - 1})(p) \left[\left(1 + |p|^\beta\right) \right] g(p) \right| \mathrm{d}p \\ &\quad + 2 \left\| (\widehat{|\psi| - 1}) \right\|_{L^2(B_r^c)} \|g\|_{L^2(B_r^c)}.\end{aligned} \tag{3.92}$$

To obtain a bound for the first two terms on the right-hand side of Eq. (3.92), we write all functions as a sum of one part living in B_r and another one living in its complement. In other words, we insert $1 = \chi_{B_r}(p) + \chi_{B_r^c}(p)$ in front of each function where χ_Ω denotes the characteristic function of the set Ω. We note that $\left\| \widehat{(|\psi|-1)} \right\|_{L^1(B_r)} \leq \left\| \frac{1}{|p|} \right\|_{L^2(B_r)} \left\| p\ \widehat{(|\psi|-1)} \right\|_{L^2(B_r)}$, and therefore an application of Young's inequality tells us that

$$\begin{aligned} &\int_{B_r^c} \left| \widehat{(|\psi|-1)} * \widehat{(|\psi|-1)}(p) \left[\left(1+|p|^\beta\right) \right] \hat{g}(p) \right| \mathrm{d}p \\ &\quad \lesssim \left\| \widehat{|\psi|-1} \right\|^2_{L^2(B_r^c)} \left\| \left(1+|p|^\beta\right) \hat{g} \right\|_{L^1(B_r^c)} \\ &\quad + \left\| p\widehat{(|\psi|-1)} \right\|^2_{L^2(B_r)} \left\| \frac{1}{|p|} \right\|^2_{L^2(B_r)} \left\| (1+|p|^\beta)\hat{g} \right\|_{L^\infty(B_r^c)} \\ &\quad + \left\| \widehat{(|\psi|-1)} \right\|_{L^2(B_r^c)} \left\| p\ \widehat{(|\psi|-1)} \right\|_{L^2(B_r)} \left\| \frac{1}{|p|} \right\|_{L^2(B_r)} \left\| (1+|p|^\beta)\hat{g} \right\|_{L^2(B_r^c)} . \end{aligned} \tag{3.93}$$

The gradient term on the other hand is bounded from below by

$$\begin{aligned} \|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} &\geq \|\nabla(|\psi(x)|-1)\|^2_{L^2(\mathbb{R}^3)} \\ &\geq r^2 \left\| \widehat{(|\psi|-1)} \right\|^2_{L^2(B_r(0)^c))} + \left\| p\widehat{(|\psi|-1)} \right\|^2_{L^2(B_r(0))} . \end{aligned} \tag{3.94}$$

Since $\int_{B(r)} \frac{1}{|p|}\ \mathrm{d}p = 2\pi r^2$, we finally obtain

$$\begin{aligned} &\int_{\mathbb{R}^3} |\nabla\psi(x)|^2 \,\mathrm{d}x - C_1 \int_{B_r(0)^c} \left| \hat{\Phi}(p) \right| g(p) \mathrm{d}p \\ &\quad \geq r^2 \left\| \widehat{(|\psi|-1)} \right\|^2_{L^2(B_r^c))} + \left\| p\ \widehat{(|\psi|-1)} \right\|^2_{L^2(B_r))} \\ &\qquad - 2C_1 \left\| \widehat{(|\psi|-1)} \right\|_{L^2(B_r^c)} \|g\|_{L^2(B_r^c)} \\ &\qquad - \frac{C_1 C_2}{1+|r|^\beta} \left[\left\| \widehat{(|\psi|-1)} \right\|^2_{L^2(B_r^c)} \left\| (1+|p|^\beta) g \right\|_{L^1(B_r^c)} \right. \\ &\qquad + \left\| p\widehat{(|\psi|-1)} \right\|^2_{L^2(B_r)} r^2 \left\| (1+|p|^\beta) g \right\|_{L^\infty(B_r^c)} \\ &\qquad \left. + \left\| \widehat{(|\psi|-1)} \right\|_{L^2(B_r^c)} \left\| p\widehat{(|\psi|-1)} \right\|_{L^2(B_r)} r \left\| (1+|p|^\beta) g \right\|_{L^2(B_r^c)} \right] \end{aligned} \tag{3.95}$$

for an appropriately chosen constant $C_2 > 0$. Choosing $\beta > 2$ and r large enough, we easily see that the expression on the right-hand side behaves as claimed. □

Let us apply Lemma 3.4 to derive a-priori estimates for $\|\nabla\psi\|_{L^2(\mathbb{R}^3)}$ and $\|\xi_0\|_{L^2(\mathbb{R}^6)}$. We choose Γ such that $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \leq 0$. This is always possible because $\mathcal{F}_\beta(\Gamma_0^w, \Gamma_0^w) = 0$.

Using Eq. (3.90), one easily finds

$$0 \geq C_1 \left(h \|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + \|\xi_0\|^2_{L^2(\mathbb{R}^6)} \right) - hC_2 \int_{\mathbb{R}^3} \left| \hat{\Phi}(p) \widehat{\frac{1}{1+(\cdot)^4}}(p) \right| \mathrm{d}p - C_3 h \tag{3.96}$$

$$\geq C_1 \|\xi_0\|^2_{L^2(\mathbb{R}^6)} - hC_4(r) \int_{B_r} \left| \hat{\Phi}(p) \widehat{\frac{1}{1+(\cdot)^4}}(p) \right| \mathrm{d}p - C_5(r) h$$

for appropriately chosen constants $C_1, ..., C_4(r), C_5(r)$ and r large enough. To come from the first to the second line, we applied Lemma 3.4. Note that since $(1+(\cdot)^4)^{-1} \in H^k(\mathbb{R}^3)$ for all $k \in \mathbb{N}$, its Fourier transform fulfils all requirements needed to apply Lemma 3.4 (Choose for example $\beta = 4$.). Eq. (3.96) yields the following bound for $\|\xi_0\|^2_{L^2(\mathbb{R}^6)}$:

$$\|\xi_0\|^2_{L^2(\mathbb{R}^6)} \lesssim h + h \left\| \widehat{\frac{1}{1+(\cdot)^4}} \right\|_{L^\infty(B_r)} \left\| \hat{\Phi} \right\|_{L^2(B_r)} \tag{3.97}$$

$$\lesssim h \left\| \hat{\Phi} \right\|_{L^2(B_r)} + h.$$

Hence, the $L^2(\mathbb{R}^6)$-norm of ξ_0 is always of order $h^{1/2}$ compared with the $L^2(B_r)$-norm of $\hat{\Phi}$. Arguing the same way, we obtain

$$\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} \lesssim \left\| \hat{\Phi} \right\|_{L^2(B_r)} + 1. \tag{3.98}$$

Having the a-priori estimates Eqs. (3.96) and (3.97) at hand, we can construct the lower bound for the BCS functional.

We start with Eq. (3.90) and assume that $\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \leq 0$ holds. An application of Lemma 3.4 gives

$$\mathcal{F}_\beta(\Gamma, \Gamma_0^w) \geq C_1 h \left(\|\nabla\psi\|^2_{L^2(\mathbb{R}^3)} + \|\xi_0\|^2_{H_1(\mathbb{R}^6)} + \|q\|^2_{H^1(\mathbb{R}^6)} \right) + \frac{2}{5\beta} \mathrm{Tr} \left(\alpha\overline{\alpha} - \alpha_0^w \overline{\alpha}_0^w \right)^2$$

$$- C_2 h \int_{B_r} \left| \hat{\Phi}(p) \right| \mathrm{d}p - C_3 h, \tag{3.99}$$

for appropriately chosen constants $C_1, C_2, C_3 > 0$ depending on r, r large enough and h small enough. We define the functions $\Phi_\leq$ and $\Phi_>$ to be the inverse Fourier transforms of $\hat{\Phi}(p)\chi_{\{|p|\leq r\}}(p)$ and $\hat{\Phi}(p)\chi_{\{|p|>r\}}(p)$, respectively. When we insert $\alpha = \alpha_0\psi + \xi_0$ into the fourth term on the right-hand side of Eq. (3.99), it reads

$$\begin{aligned} \mathrm{Tr} \left[(\alpha\overline{\alpha} - \alpha_0^w \overline{\alpha}_0^w)^2 \right] &= \mathrm{Tr} \left[(\alpha_0 \Phi_\leq \alpha_0)^2 \right] \\ &\quad + 2\mathrm{Re}\mathrm{Tr} \left[\alpha_0 \Phi_\leq \alpha_0 \left(\alpha_0 \Phi_> \alpha_0 + \xi_0 \xi_0^* + \alpha_0 \psi \xi_0^* + \xi_0 \overline{\psi} \alpha_0 - \alpha_0^w \overline{\alpha_0^w} + \alpha_0^2 \right) \right] \\ &\quad + \mathrm{Tr} \left[\left(\alpha_0 \Phi_> \alpha_0 + \xi_0 \xi_0^* + \alpha_0 \psi \xi_0^* + \xi_0 \overline{\psi} \alpha_0 - \alpha_0^w \overline{\alpha_0^w} + \alpha_0^2 \right)^2 \right] \\ &\geq \mathrm{Tr} \left[(\alpha_0 \Phi_\leq \alpha_0)^2 \right] + 2\mathrm{Re}\mathrm{Tr} \left[\alpha_0 \Phi_\leq \alpha_0 \left(\xi_0 \xi_0^* + \alpha_0 \psi \xi_0^* + \xi_0 \overline{\psi} \alpha_0 - \alpha_0^w \overline{\alpha_0^w} + \alpha_0^2 \right) \right]. \end{aligned} \tag{3.100}$$

To come to the last line, we used that $\mathrm{Tr}\left[\alpha_0\Phi_{\leq}\alpha_0^2\Phi_{>}\alpha_0\right]=0$, which results from the fact that $\hat{\Phi}_{\leq}(p)$ and $\hat{\Phi}_{>}(p)$ have disjoint support. The first term on the right-hand side of Eq. (3.100) can be rewritten as

$$\begin{aligned}\mathrm{Tr}\left[(\alpha_0\Phi_{\leq}\alpha_0)^2\right] &= h^{-3}\int_{\mathbb{R}^3}\left|\hat{\Phi}_{\leq}(p)\right|^2\hat{\alpha}_0^2*\hat{\alpha}_0^2(hp)\mathrm{d}p \\ &\geq \inf_{p\in B_r}\left(\hat{\alpha}_0^2*\hat{\alpha}_0^2(hp)\right)h^{-3}\int_{\mathbb{R}^3}\left|\hat{\Phi}_{\leq}(p)\right|^2\mathrm{d}p,\end{aligned} \tag{3.101}$$

where, like above, B_r denotes the ball of radius r centered at zero. Let us note that since α_0 is the unique ground state of the real operator $K_T^{\Delta_0}+V$ its Fourier transform is a real function. To come to the second line, we used that $\hat{\Phi}_{\leq}(p)$ equals zero outside B_r. Using Lemma 3.6, it can easily be seen that $\inf_{p\in B_r}\left(\hat{\alpha}_0^2*\hat{\alpha}_0^2(hp)\right)=m(r)>0$. Obviously, $m(r)$ tends to zero as r tends to infinity. The remaining terms on the right-hand side can be bounded with the help of the a-priori estimates for $\|\nabla\psi\|_{L^2(\mathbb{R}^3)}$ and $\|\xi_0\|_{L^2(\mathbb{R}^6)}$ derived above. We find

$$\begin{aligned}\left|\mathrm{Tr}\left[\alpha_0\Phi_{\leq}\alpha_0\xi_0\xi_0^*\right]\right| &\leq \|\alpha_0\Phi_{\leq}\alpha_0\|_2\|\xi_0\|_4^2 \\ &\leq h^{3/2}\left\|\hat{\alpha}_0^2*\hat{\alpha}_0^2(p)\right\|_{L^\infty(\mathbb{R}^3)}^{1/2}\left\|\hat{\Phi}_{\leq}\right\|_{L^2(\mathbb{R}^3)}\|\xi_0\|_2^2 \\ &\lesssim h^{3/2}\left\|\hat{\Phi}_{\leq}\right\|_{L^2(\mathbb{R}^3)}^2+h^{3/2}.\end{aligned} \tag{3.102}$$

To come from the first to the second line, we used that lower Schatten class-norms dominate higher ones, in particular $\|A\|_4\leq\|A\|_2$ for all compact operators A. Using Young's inequality and the a-priori estimates, one similarly finds

$$\begin{aligned}\left|\mathrm{Tr}\left[\alpha_0\Phi_{\leq}\alpha_0\left(\alpha_0\psi\xi_0^*+\xi_0\overline{\psi}\alpha_0\right)\right]\right| &\lesssim \|\alpha_0\Phi_{\leq}\alpha_0\|_2\|\alpha_0\psi\|_\infty\|\xi_0\|_2 \\ &\lesssim h^{1/2}\left\|\hat{\Phi}_{\leq}\right\|_{L^2(\mathbb{R}^3)}\left\|h^{-3}\alpha_0(x/h)\right\|_{L^{6/5}(\mathbb{R}^3)}\|\nabla\psi\|_{L^2(\mathbb{R}^3)}\|\xi_0\|_2 \\ &\leq h^{3/2}\left(\left\|\hat{\Phi}_{\leq}\right\|_{L^2(\mathbb{R}^3)}^2+1\right).\end{aligned} \tag{3.103}$$

On the other hand, an application of Lemma 3.8 yields

$$\begin{aligned}\left|\mathrm{Tr}\left[\alpha_0\Phi_{\leq}\alpha_0\left(-\alpha_0^w\overline{\alpha_0^w}+\alpha_0^2\right)\right]\right| &\leq \|\alpha_0\Phi_{\leq}\alpha_0\|_2\left(\left\|-\alpha_0^w\overline{\alpha_0^w}+\alpha_0\overline{\alpha}_0^w\right\|_2+\left\|-\alpha_0\overline{\alpha}_0^w+\alpha_0^2\right\|_2\right) \\ &\lesssim h^2\left\|\hat{\Phi}_{\leq}\right\|_{L^2(\mathbb{R}^3)}.\end{aligned} \tag{3.104}$$

When we insert these into Eq. (3.100) and afterwards insert the resulting expression into Eq. (3.99), we finally obtain

$$\begin{aligned}\mathcal{F}_\beta(\Gamma,\Gamma_0^w) \geq C_1\Bigg(&h\|\nabla\psi\|_{L^2(\mathbb{R}^3)}^2+h\left\|\widehat{|\psi|^2-1}\right\|_{L^2(B_r)}^2+\|\xi_0\|_{H_1(\mathbb{R}^6)}^2 \\ &+\|q\|_{H^1(\mathbb{R}^6)}^2\Bigg)-C_2h\end{aligned} \tag{3.105}$$

for two constants $C_1, C_2 > 0$ which both depend on r and for r large enough as well as h small enough. This concludes the proof of the lower bound for the BCS functional. All a-priori bounds except the one for $\||\psi|^2 - 1\|_{L^2(\mathbb{R}^3)}$ can easily be read off from Eq. (3.105).

In order to prove the remaining a-priori estimate, we first derive a uniform bound for $\||\psi| - 1\|_{L^2(\mathbb{R}^3)}$. The uniform bound for $\left\||\psi|^2 - 1\right\|_{L^2(\mathbb{R}^3)}$ will then follow easily. For the sake of convenience, we introduce the abbreviation $\eta(x) = |\psi(x)| - 1$ and always assume $\eta \in L^2(\mathbb{R}^3)$. As a straight forward computation shows, we have

$$\Phi(x) = |\psi(x)|^2 - 1 = \eta(x)^2 - 2\eta(x). \tag{3.106}$$

Since $\|\nabla\psi\|_{L^2(\mathbb{R}^3)} \geq \|\nabla\left(|\psi| - 1\right)\|_{L^2(\mathbb{R}^3)} = \|p\hat{\eta}(p)\|_{L^2(\mathbb{R}^3)}$ we have a uniform bound on the $L^2(\mathbb{R}^3)$-norm of $p\hat{\eta}(p)$. Hence, we can write $\hat{\eta}(p) = \hat{\eta}_1(p) + \hat{\eta}_2(p)$ with $\hat{\eta}_1(p) = \hat{\eta}(p)\chi_{B_1}(p)$ and $\hat{\eta}_2(p) = \hat{\eta}(p)\chi_{B_1^c}(p)$, where the $L^1(\mathbb{R}^3)$-norm of $\hat{\eta}_1$ and the $L^2(\mathbb{R}^3)$-norm of $\hat{\eta}_2$ are both uniformly bounded. This is because

$$\begin{aligned} &\|\hat{\eta}_1(p)\|_{L^1(\mathbb{R}^3)} \leq \left\|\frac{1}{|p|}\right\|_{L^2(B_r)} \|p\hat{\eta}_1(p)\|_{L^2(\mathbb{R}^3)} \lesssim \|\nabla\psi\|_{L^2(\mathbb{R}^3)} \quad \text{and} \\ &\|\hat{\eta}_2(p)\|_{L^2(\mathbb{R}^3)} \leq \|p\hat{\eta}_2(p)\|_{L^2(\mathbb{R}^3)} \lesssim \|\nabla\psi\|_{L^2(\mathbb{R}^3)}. \end{aligned} \tag{3.107}$$

We have chosen the symbols $\hat{\eta}_1$ and $\hat{\eta}_2$ for the functions in our decomposition of η to highlight that their inverse Fourier transform exists. To bound the whole $L^2(\mathbb{R}^3)$-norm of η, we need a uniform bound on the $L^2(\mathbb{R}^3)$-norm of $\hat{\eta}_1$. This can be obtained with the help of the uniform bound on $\left\|\hat{\Phi}\right\|_{L^2(B_r)}$ as we will see in the next paragraph.

Written in terms of $\hat{\eta}_1$ and $\hat{\eta}_2$, this bound reads

$$\begin{aligned} \left\|\hat{\Phi}(p)\right\|_{L^2(B_r)} &= \|\hat{\eta}_1 * \hat{\eta}_1(p) + \hat{\eta}_2 * \hat{\eta}_2(p) + 2\hat{\eta}_1 * \hat{\eta}_2(p) - 2\hat{\eta}_1(p) - \hat{\eta}_2(p)\|_{L^2(B_r)} \\ &\leq C. \end{aligned} \tag{3.108}$$

Since

$$\begin{aligned} \|\hat{\eta}_2 * \hat{\eta}_2(p)\|_{L^2(B_r)} &\leq \mu\left(B_r\right)^{1/2} \|\hat{\eta}_2\|^2_{L^2(\mathbb{R}^3)}, \\ \|\hat{\eta}_1 * \hat{\eta}_2(p)\|_{L^2(B_r)} &\leq \|\hat{\eta}_1\|_{L^1(\mathbb{R}^3)} \|\hat{\eta}_2\|_{L^2(\mathbb{R}^3)} \end{aligned} \tag{3.109}$$

we know that

$$\|\hat{\eta}_1 * \hat{\eta}_1 - 2\hat{\eta}_1\|_{L^2(B_r)} \leq C. \tag{3.110}$$

The function $\hat{\eta}_1$ and therefore also the function $\hat{\eta}_1 * \hat{\eta}_1$ have compact support which implies that if r is large enough we can replace the $L^2(B_r)$-norm in Eq. (3.110) by the $L^2(\mathbb{R}^3)$-norm. And since $\hat{\eta}_1$ has an inverse Fourier transform η_1, we can go back to coordinate space to find

$$\left\|\eta_1^2 - 2\eta_1\right\|_{L^2(\mathbb{R}^3)} \leq C. \tag{3.111}$$

In order to decompose η_1 another time, we define the sets $\Omega_{\leq} = \{x \in \mathbb{R}^3 | \, |\eta_1(x)| \leq 1\}$ and $\Omega_{>} = \{x \in \mathbb{R}^3 | \, |\eta_1(x)| > 1\}$ which have the obvious property $\Omega_{\leq} \cup \Omega_{>} = \mathbb{R}^3$. Note that

$$\mu\left(\Omega_{>}\right) \leq \int_{\mathbb{R}^3} |\eta_1(x)|^6 \,\mathrm{d}x \lesssim \|p\hat{\eta}_1(p)\|^6_{L^2(\mathbb{R}^3)} \lesssim \|\nabla\psi\|^6_{L^2(\mathbb{R}^3)} \tag{3.112}$$

where the first step follows from Sobolev's inequality and the second step from Chebyshev's inequality. Of course μ denotes Lebesgue measure on $\mathbb{R}^3$. Hence, $\mu\left(\Omega_{>}\right)$ is uniformly bounded. On $\Omega_{>}$, the function η_1 is square integrable with uniformly bounded $L^2(\Omega_{>})$-norm because

$$\|\eta_1\|_{L^2(\Omega_{>})} \leq \mu\left(\Omega_{>}\right)^{1/3} \|\eta_1\|_{L^6(\Omega_{>})} \lesssim \|\nabla\psi\|^3_{L^2(\mathbb{R}^3)}. \tag{3.113}$$

Hence, we only need to construct a uniform bound for the $L^2(\Omega_{\leq})$-norm of η_1 to obtain a uniform bound on $\|\eta\|_{L^2(\mathbb{R}^3)}$. This is possible because we also know

$$\begin{aligned} \left\|\eta_1^2 - 2\eta_1\right\|_{L^2(\mathbb{R}^3)} \geq \left\|\eta_1^2 - 2\eta_1\right\|_{L^2(\Omega_{\leq})} &\geq 2\left\|\eta_1\right\|_{L^2(\Omega_{\leq})} - \left\|\eta_1^2\right\|_{L^2(\Omega_{\leq})} \\ &\geq \|\eta_1\|_{L^2(\Omega_{\leq})}. \end{aligned} \tag{3.114}$$

To obtain the result, we used that $|\eta_1(x)^2| \leq |\eta_1(x)|$ holds on $\Omega_{\leq}$.

It remains to construct a uniform bound on $\|\Phi\|_{L^2(\mathbb{R}^3)}$. To that end, we again write $\Phi = \eta^2 - 2\eta$ and estimate

$$\|\Phi\|_{L^2(\mathbb{R}^3)} \leq \|\eta\|^2_{L^4(\mathbb{R}^3)} + 2\|\eta\|_{L^2(\mathbb{R}^3)} \leq \|\eta\|^{1/2}_{L^2(\mathbb{R}^3)} \|\eta\|^{3/2}_{L^6(\mathbb{R}^3)} + 2\|\eta\|_{L^2(\mathbb{R}^3)}. \tag{3.115}$$

This concludes the proof of Theorem 3.1.

3.3 Properties of α_0 and α_0^w

3.3.1 Properties of α_0

In this section we establish some properties of the minimizer of the translation-invariant BCS functional that are used frequently in the main text. Without mentioning it explicitly, we always assume $V \in L^{3/2}(\mathbb{R}^3)$ and $\mu > 0$. We remind the reader that the translation-invariant BCS functional has a minimizer $\alpha_0 \in H^1(\mathbb{R}^3)$ and that $\Delta_0(x) = 2V(x)\alpha_0(x)$, see [16]. By $\|\cdot\|_{H^k(\mathbb{R}^d)}$ and $\|\cdot\|_{W^{k,p}(\mathbb{R}^d)}$ we denote the usual Sobolev-norms and by $H^k(\mathbb{R}^d)$ and $W^{k,p}(\mathbb{R}^d)$ the corresponding function spaces.

Lemma 3.5 *Let the pair (γ_0, α_0) with $\alpha_0 \neq 0$ be a solution of the Euler-Lagrange equations of the translation-invariant BCS functional and choose $k \in \mathbb{N}_0$. If $V \in W^{k,\infty}(\mathbb{R}^3)$ then $\alpha_0 \in H^{k+2}(\mathbb{R}^3)$. On the other hand, $\hat{V} \in H^k(\mathbb{R}^3)$ implies $\hat{\alpha}_0 \in H^k(\mathbb{R}^3)$.*

Proof. Since the pair (γ_0, α_0) minimizes the translation-invariant BCS functional, α_0 is a solution of the corresponding Euler-Lagrange equation

$$K_T^{\Delta_0}(p)\hat{\alpha}_0(p) + (2\pi)^{-3/2}\hat{V} * \hat{\alpha}_0(p) = 0, \tag{3.116}$$

which holds pointwise almost everywhere [16]. Using this equation, we can estimate the $L^2(\mathbb{R}^3)$-norm of derivatives of α_0 as follows:

$$\begin{aligned}
\|\partial_i^m \alpha_0\|_{L^2(\mathbb{R}^3)} &= (2\pi)^{-3/2} \left\| p_i^m \frac{1}{K_T^{\Delta_0}} \hat{V} * \hat{\alpha}_0 \right\|_{L^2(\mathbb{R}^3)} \leq \left\| \frac{p^2}{K_T^{\Delta_0}} \right\|_{L^\infty(\mathbb{R}^3)} \left\| p_i^{m-2} \hat{V} * \hat{\alpha}_0 \right\|_{L^2(\mathbb{R}^3)} \\
&\lesssim \sum_{j=0}^{m-2} \left\| \partial_i^j V \right\|_\infty \left\| \partial_i^{m-2-j} \alpha_0 \right\|_{L^2(\mathbb{R}^3)} .
\end{aligned} \tag{3.117}$$

To come to the last line, we used that $K_T^{\Delta_0} \lesssim (1 + E(p)) \lesssim (1 + p^2)$ which holds because $\left\| \hat{\Delta}_0 \right\|_{L^\infty(\mathbb{R}^3)} \lesssim \left\| \hat{V} \right\|_{L^2(\mathbb{R}^3)} \|\hat{\alpha}_0\|_{L^2(\mathbb{R}^3)}$. Eq. (3.117) shows that $V \in W^{k,\infty}(\mathbb{R}^3)$ implies $\alpha_0 \in H^{k+2}(\mathbb{R}^3)$.

In order to obtain the second property, we again use Eq. (3.116) together with the fact that in the second term, $\hat{\alpha}_0$ appears only in the convolution with $\hat{V}$. Hence, derivatives act only on $\hat{V}$ and not on $\hat{\alpha}_0$. We compute

$$\begin{aligned}
\|\partial_i \hat{\alpha}_0(p)\|_{L^2(\mathbb{R}^3)} &\lesssim \left\| \frac{2k(p)p_i + \hat{\Delta}_0(p)\left(\partial_i \hat{\Delta}_0(p)\right)}{4TE(p)\left(K_T^{\Delta_0}(p)\right)^2} \left[\frac{\frac{E(p)}{2T} - \frac{1}{2}\sinh\left(\frac{E(p)}{T}\right)}{E(p)^2 \cosh\left(\frac{E(p)}{2T}\right)^2} \right] \hat{\Delta}_0(p) \right\|_{L^2(\mathbb{R}^3)} \\
&\quad + \left\| \frac{1}{K_T^{\Delta_0}(p)} \left(\partial_i \hat{V}\right) * \hat{\alpha}_0(p) \right\|_{L^2(\mathbb{R}^3)} \\
&\lesssim \left\| \frac{4k(p)p_i + 2\hat{\Delta}_0(p)\left(\partial_i \hat{\Delta}_0(p)\right)}{4TE(p)K_T^{\Delta_0}(p)} \right\|_{L^\infty(\mathbb{R}^3)} \left\| \frac{\frac{E(p)}{2T} - \frac{1}{2}\sinh\left(\frac{E(p)}{T}\right)}{E(p)^2 \cosh\left(\frac{E(p)}{2T}\right)^2} \right\|_{L^\infty(\mathbb{R}^3)} \\
&\qquad \times \left\| \hat{V} \right\|_{L^2(\mathbb{R}^3)} \|\hat{\alpha}_0\|_{L^2(\mathbb{R}^3)} \\
&\quad + \left\| \frac{1}{K_T^{\Delta_0}} \right\|_{L^2(\mathbb{R}^3)} \left\| \partial_i \hat{V} \right\|_{L^2(\mathbb{R}^3)} \|\hat{\alpha}_0\|_{L^2(\mathbb{R}^3)} \\
&\leq C\left(\left\| \hat{V} \right\|_{H^1(\mathbb{R}^3)} \right) \|\hat{\alpha}_0\|_{L^2(\mathbb{R}^3)} .
\end{aligned} \tag{3.118}$$

To obtain the result in the last line, we used the identity $\hat{\Delta}_0(p) = -2(2\pi)^{-3/2}\hat{V} * \hat{\alpha}_0(p)$ and Young's inequality. Looking at the above terms, one can easily see that another differentiation does not change this structure, except that second derivatives of $\hat{V}$ appear. The extension to k derivatives is a simple exercise in differentiation together with applications of estimates of the above kind. □

The results contained in the next Lemma have been shown in [19] in the case $T = 0$. Note that at $T = 0$ the BCS functional can be written as a functional that depends only on α. In this situation the uniqueness part is easier to show than for $T > 0$.

Lemma 3.6 *Let $V \in L^2(\mathbb{R}^3)$ with $\hat{V} \leq 0$, $\hat{V}$ not identically zero and assume $T_c > 0$. Then for $0 < T < T_c$ the translation-invariant BCS functional $\mathcal{F}_{TI}(\gamma, \alpha)$ admits a unique (up to a phase in front of α_0) minimizer (γ_0, α_0). We can choose the phase of α_0 such that $\hat{\alpha}_0 \geq 0$. Additionally, $\mu(\{p \in \mathbb{R}^3 | \ \hat{\alpha}_0(p) = 0\}) = 0$ and α_0 is the unique ground state of the linear operator $K_T^{\Delta_0} + V$ where $\Delta_0(x) = 2V(x)\alpha_0(x)$. If $V(-x) = V(x)$ then $\alpha_0(-x) = \alpha_0(x)$.*

Proof. We recall the definition of the translation-invariant BCS functional [16]

$$\mathcal{F}_{TI}(\Gamma) = \int_{\mathbb{R}^3} (p^2 - \mu)\gamma(p)\mathrm{d}p + \int_{\mathbb{R}^3} V(x)|\alpha(x)|^2 \mathrm{d}x - \frac{1}{\beta} S(\Gamma), \tag{3.119}$$

$$S(\Gamma) = -\frac{1}{2}\int_{\mathbb{R}^3} \mathrm{Tr}\left[\Gamma(p)\ln(\Gamma(p)) + (1-\Gamma(p))\ln(1-\Gamma(p))\right]\mathrm{d}p$$

$$= -\int_{\mathbb{R}^3} \left[s(p)\ln s(p) + (1-s(p))\ln(1-s(p))\right]\mathrm{d}p.$$

The function s is determined by the equation $s(p)(1-s(p)) = \gamma(p)(1-\gamma(p)) - |\hat{\alpha}(p)|^2$ and appears in the computation of the eigenvalues of the matrix in the second line of Eq. (3.119). Hence, the entropy $S(\Gamma)$ depends only on the modulus of $\hat{\alpha}(p)$. On the other hand, the interaction energy can be reduced when $\hat{\alpha}$ is replaced by its modulus:

$$\int_{\mathbb{R}^3} \overline{\hat{\alpha}}(p)(\hat{V} * \hat{\alpha})(p)\mathrm{d}p \geq -\int_{\mathbb{R}^3} |\hat{\alpha}|(p)(|\hat{V}| * |\hat{\alpha}|)(p)\mathrm{d}p. \tag{3.120}$$

The kinetic energy is also not affected by this transformation. This is so, because the only connection between γ and α is given by the inequality $|\hat{\alpha}(p)|^2 \leq \hat{\gamma}(p)(1-\hat{\gamma}(p))$. Hence, we have $\mathcal{F}_{TI}(\Gamma) \geq \mathcal{F}_{TI}(\Gamma')$ where Γ' equals Γ except that $\hat{\alpha}'(p) = |\hat{\alpha}(p)|$ which implies that $\mathcal{F}_{TI}$ admits a minimizer Γ_0 with $\hat{\alpha}_0 \geq 0$.

As a minimizer of the translation-invariant BCS functional, α_0 solves the Euler-Lagrange equation Eq. (3.116) which can be written as

$$K_T^{\Delta_0}(p)\hat{\alpha}_0(p) = (2\pi)^{-3/2}|\hat{V}| * \hat{\alpha}_0(p). \tag{3.121}$$

It holds for almost every $p \in \mathbb{R}^3$. Now let us assume that $\mu(\{p \in \mathbb{R}^3 | \ \hat{\alpha}_0(p) = 0\}) \neq 0$. The support of $K_T^{\Delta_0}(p)\hat{\alpha}_0(p)$ equals the support of $\hat{\alpha}_0(p)$. On the other hand, our assumption implies that the support of $|\hat{V}| * \hat{\alpha}_0(p)$ has strictly larger Lebesgue measure than the support of $\hat{\alpha}_0(p)$. Since Eq. (3.121) holds this is a contradiction and we conclude that $\mu(\{p \in \mathbb{R}^3 | \ \hat{\alpha}_0(p) = 0\}) = 0$.

To prove that α_0 is the unique ground state of $K_T^{\Delta_0} + V$, we consider $Q(\psi) = \left(\psi, \left(K_T^{\Delta_0} + V\right)\psi\right)$ for $\psi \in H^1(\mathbb{R}^3)$ with $\|\psi\|_{L^2(\mathbb{R}^3)} = 1$. The quadratic form Q is bounded from below because $K_T^{\Delta_0}(p) \geq C\left(1+p^2\right)$ for an appropriately chosen constant $C > 0$ depending on T. Since $K_T^{\Delta_0}(p) \geq 2T$ and $\sigma_{ess}\left(K_T^{\Delta_0} + V\right) = [2T, \infty)$ (V is relatively compact with respect to $K_T^{\Delta_0}$), we know that $K_T^{\Delta_0} + V$ has only discrete eigenvalues below $2T$ and since $\left(K_T^{\Delta_0} + V\right)\alpha_0 = 0$ it follows that Q attains its infimum. Under our assumptions on $\hat{V}$, we have the obvious inequality

$$\left(\hat{\psi}, K_T^{\Delta_0}\hat{\psi}\right) + (2\pi)^{-3/2}\left(\hat{\psi}, \hat{V} * \hat{\psi}\right) \geq \left(|\hat{\psi}|, K_T^{\Delta_0}|\hat{\psi}|\right) - (2\pi)^{-3/2}\left(|\hat{\psi}|, |\hat{V}| * |\hat{\psi}|\right). \tag{3.122}$$

This implies that Q admits a minimizer ψ_0 with $\hat{\psi}_0 \geq 0$. Since ψ_0 also solves the corresponding Euler-Lagrange equation we argue like before to show that

$$\mu\left(\left\{p \in \mathbb{R}^3 \middle|\ \hat{\psi}_0(p) = 0\right\}\right) = 0. \tag{3.123}$$

Let φ_0 be another eigenvector of $K_T^{\Delta_0} + V$, corresponding to its lowest eigenvalue λ. It has to be orthogonal to ψ_0, and hence its Fourier transform cannot have a definite sign. Since $K_T^{\Delta_0}$ and $\hat{V}$ are real functions the minimizer $\hat{\varphi}_0$ can be chosen real. We write $\hat{\varphi}_0 = \hat{\varphi}_0^+ - \hat{\varphi}_0^-$, where $\hat{\varphi}_0^{+/-}$ denotes the positive/negative part of $\hat{\varphi}_0$ with the obvious property $\hat{\varphi}_0^+(p)\hat{\varphi}_0^-(p) = 0$ for almost every $p \in \mathbb{R}^3$. Insertion into Q gives

$$\begin{aligned} Q(\varphi_0) = \left(\hat{\varphi}_0^+, K_T^{\Delta_0}\hat{\varphi}_0^+\right) - \int_{\mathbb{R}^3} \hat{\varphi}_0^+(p)|\hat{V}| * \hat{\varphi}_0^+(p)\mathrm{d}p \\ + \left(\hat{\varphi}_0^-, K_T^{\Delta_0}\hat{\varphi}_0^-\right) - \int_{\mathbb{R}^3} \hat{\varphi}_0^-(p)|\hat{V}| * \hat{\varphi}_0^-(p)\mathrm{d}p \\ + 2\int_{\mathbb{R}^3} \hat{\varphi}_0^+(p)|\hat{V}| * \hat{\varphi}_0^-(p)\mathrm{d}p \end{aligned} \tag{3.124}$$

To minimize the above expression, we have to choose both, $\hat{\varphi}_0^+$ and $\hat{\varphi}_0^-$, to be minimizers of Q. Since they are both positive functions this implies $\mu\left(\left\{p \in \mathbb{R}^3 \middle|\ \hat{\varphi}_0^{+/-}(p) = 0\right\}\right) = 0$. It remains to minimize the last (positive) term on the right-hand side of Eq. (3.124). This can only be achieved if either $\hat{\varphi}_0^+$ or $\hat{\varphi}_0^-$ is identically zero since otherwise their support is $\mathbb{R}^3$. As already argued above, the remainder cannot be orthogonal to $\hat{\psi}_0$ which implies the uniqueness of ψ_0. Since $\hat{\alpha}_0$ is positive as well we necessarily have $\psi_0 = \alpha_0$.

It remains to show the uniqueness (up to a constant of modulus 1 in front of α_0) of Γ_0 and the property that if $V(-x) = V(x)$ we also have $\alpha_0(-x) = \alpha_0(x)$. We note that if a pair (γ, α) solves the Euler-Lagrange equation of the BCS functional, then γ is uniquely determined by α, see [16, Eqs. (3.3),(3.4)]. A straight forward computation yields

$$\begin{aligned} \mathcal{F}_{TI}(\Gamma) - \mathcal{F}_{TI}(\Gamma_0) &= \frac{1}{2\beta}\mathcal{H}_{TI}(\Gamma, \Gamma_0) + \int_{\mathbb{R}^3} V(x)\left|\alpha(x) - \alpha_0(x)\right|^2 \mathrm{d}x, \\ \mathcal{H}_{TI}(\Gamma, \Gamma_0) &= \int_{\mathbb{R}^3} \mathrm{Tr}_{\mathbb{C}^2}\left[\varphi\left(\Gamma(p)\right) - \varphi\left(\Gamma_0(p)\right) - \varphi'\left(\Gamma_0(p)\right)\left(\Gamma(p) - \Gamma_0(p)\right)\right]\mathrm{d}p. \end{aligned} \tag{3.125}$$

As in the main part of the text, $\varphi(x) = x\ln(x) + (1-x)\ln(1-x)$ as well as

$$\Gamma(p) = \begin{pmatrix} \gamma(p) & \hat{\alpha}(p) \\ \overline{\hat{\alpha}(p)} & 1-\overline{\gamma(p)} \end{pmatrix} \tag{3.126}$$

and the same for $\Gamma_0(p)$. Using Eq. (3.21) and Klein's inequality, one easily shows

$$\mathcal{H}_{TI}\left(\Gamma, \Gamma_0\right) \geq \int_{\mathbb{R}^3} \mathrm{Tr}_{\mathbb{C}^2}\left[\left(\Gamma(p) - \Gamma_0(p)\right) \frac{\beta H_0(p)}{\tanh\left(\frac{\beta H_0(p)}{2}\right)} \left(\Gamma(p) - \Gamma_0(p)\right)\right] \mathrm{d}p. \tag{3.127}$$

When we estimate the right-hand side of Eq. (3.125) with the help of Eq. (3.127) and use on the one hand that $H_0(p)^2 = \mathbb{1}_{\mathbb{C}^2} E(p)^2$ and on the other hand that $x \mapsto x/\tanh(x/(2T))$ is an even function, we obtain

$$\mathcal{F}_{TI}\left(\Gamma\right) - \mathcal{F}_{TI}\left(\Gamma_0\right) \geq \left(\alpha, \left(K_T^{\Delta_0} + V\right)\alpha\right)_{L^2(\mathbb{R}^3)} + 2T\int_{\mathbb{R}^3} \left(\gamma(p) - \gamma_0(p)\right)\mathrm{d}p. \tag{3.128}$$

To come to the right-hand side of the above equation, we used that α_0 lies in the kernel of $K_T^{\Delta_0} + V$ and that $K_T^{\Delta_0} \geq 2T$. From our previous considerations we know that $K_T^{\Delta_0} + V \geq 0$ and that the only element in its kernel is the function α_0. This together with Eq. (3.128) shows that the generalized one-particle density matrix Γ of any pair (γ, α) with either $\gamma \neq \gamma_0$ and/or $\alpha \neq \alpha_0$ (except for a phase in front of α) has energy strictly larger than $\mathcal{F}_{TI}(\gamma_0, \alpha_0)$. This proves the uniqueness of Γ_0. Assume now that $V(-x) = V(x)$. Then the pair $(\gamma_0, \alpha_0^-))$ where $\alpha_0^-(x) = \alpha_0(-x)$ has the same energy as the pair (γ_0, α_0). If $\alpha_0 \neq \alpha_0^-$ the two pairs would minimize the translation-invariant BCS functional which contradicts the uniqueness of Γ_0, hence $\alpha_0 = \alpha_0^-$. This concludes the proof of Lemma 3.6. □

The next Lemma shows how α_0 behaves close to T_c.

Lemma 3.7 *Let $V \in L^2(\mathbb{R}^3)$ with $\hat{V} \in L^p(\mathbb{R}^3)$ for a number $p \in \left[1, \frac{12}{7}\right)$ be such that $T_c > 0$ and chose $T = T_c(1 - h^2 D)$ with $D > 0$ and $h \ll 1$. Additionally, let $\{\alpha_*^i \mid i = 1, ..., m\}$ be an orthonormal basis of the kernel of $K_{T_c} + V$. Then the Cooper-pair wave function of any minimizer Γ_0 of the translation-invariant BCS functional $\mathcal{F}_{TI}$ is of the form $\alpha_0(x) = \sum_{i=1}^m c_i(h)\alpha_*^i(x) + \eta(x)$ where $0 < \sum_{i=1}^m |c_i(h)| \lesssim h$. The function η lies in the complement of the kernel of $K_{T_c} + V$ and is such that $\|\eta\|_{H^1(\mathbb{R}^3)} \lesssim h^2$ as well as $\|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)} \lesssim h^2$. If in addition $V \in W^{k,\infty}(\mathbb{R}^3)$ for $k \in \mathbb{N}$, then $\|\eta\|_{H^{k+2}(\mathbb{R}^3)} \lesssim h^2$ and if $\hat{V} \in H^k(\mathbb{R}^3)$ then $\|\hat{\eta}\|_{H^k(\mathbb{R}^3)} \lesssim h^2$.*

Proof. Since Γ_0 minimizes $\mathcal{F}_{TI}$ its Cooper-pair wave function α_0 solves the corresponding Euler-Lagrange equation which can be written as

$$\left(K_T^{\Delta_0} + V\right)\alpha_0 = 0. \tag{3.129}$$

Here $\Delta_0(x) = 2V(x)\alpha_0(x)$. During our analysis, we will have to compare the operator $K_T^{\Delta_0}$ with the operator K_{T_c}. To that end, we write

$$K_T^{\Delta_0} = K_{T_c} + \underbrace{(K_T - K_{T_c})}_{A\leq 0} + \underbrace{\left(K_T^{\Delta_0} - K_T\right)}_{B\geq 0}. \tag{3.130}$$

The operators A and B have a sign because for $T, \Delta \in \mathbb{R}_+$ the expression

$$\frac{\sqrt{k(p)^2 + |\Delta|^2}}{\tanh\left(\frac{\sqrt{k(p)^2+|\Delta|^2}}{2T}\right)} \tag{3.131}$$

is monotone increasing in T for fixed Δ and $p \in \mathbb{R}^3$ and the same when the roles of T and Δ are exchanged. Both, A and B are pseudo-differential operators. By a slight abuse of notation we denote by $A(p)$ the symbol of A and similarly for B. A simple computation yields

$$A(p) = \int_0^1 \frac{\frac{k(p)^2}{2}\int_0^1 \frac{h^2 T_c D}{(T_c+h^2 DT_c s)}\mathrm{d}s}{\sinh^2\left(\frac{k(p)}{2T_c} - \frac{tk(p)}{2}\int_0^1 \frac{h^2 DT_c}{(T_c+h^2DT_cs)^2}\mathrm{d}s\right)}\mathrm{d}t. \tag{3.132}$$

For large $|p|$ the smooth function $A(p)$ and all its derivatives have exponential decay. It is easy to show that $\left\|p_i^n \partial_{p_j}^m A(p)\right\|_{L^\infty(\mathbb{R}^3)} \leq C(n,m)h^2$ for all $n, m \in \mathbb{N}_0$ and $i, j = 1, 2, 3$. On the other hand, the function $B(p)$ is given by

$$B(p) = \int_0^1 \frac{T\sinh\left(\frac{|k(p)|+t\delta E(p)}{T}\right) - |k(p)| - t\delta E(p)}{2T\sinh^2\left(\frac{|k(p)|+t\delta E(p)}{2T}\right)}\delta E(p)\mathrm{d}t, \tag{3.133}$$

$$\delta E(p) = \int_0^1 \frac{\left|\hat{\Delta}_0(p)\right|^2}{2\sqrt{k(p)^2 + s\left|\hat{\Delta}_0(p)\right|^2}}\mathrm{d}s.$$

To come to Eq. (3.131), we have used that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x}{\tanh\left(\frac{x}{2T}\right)} = \frac{T\sinh\left(\frac{x}{T}\right) - x}{2T\sinh^2\left(\frac{x}{2T}\right)}. \tag{3.134}$$

Since the function on the right-hand side of Eq. (3.134) is bounded, we know that $|B(p)| \lesssim |\delta E(p)|$ holds. Together with $|\delta E(p)| \lesssim \left|\hat{\Delta}_0(p)\right|$, which directly follows from Eq. (3.133), this implies $|B(p)| \lesssim \left|\hat{\Delta}_0(p)\right|$ and in particular $\|B(p)\|_{L^\infty(\mathbb{R}^3)} < \infty$.

In the first step, we will show that $\|\alpha_0\|_{L^2(\mathbb{R}^3)} \lesssim h$ holds. To that end, let us start with Eq. (3.129), which implies

$$(\alpha_0, B\alpha_0) \leq -(\alpha_0, A\alpha_0) \lesssim h^2 \|\alpha_0\|^2_{L^2(\mathbb{R}^3)} \tag{3.135}$$

because $K_{T_c} + V \geq 0$ and $\|A\|_\infty \lesssim h^2$. To derive the estimate for α_0, we need a lower bound for the term on the left-hand side of Eq. (3.135). When we use the properties of the function from Eq. (3.134), the gap equation in the form $\hat{\alpha}_0(p) = -\frac{1}{2} K_T^{\Delta_0}(p)^{-1} \hat{\Delta}_0(p)$ and the bound $K_T^{\Delta_0}(p) \geq C(1+p^2)$, the estimate

$$B(p) \geq C(1+p^2) \left|\hat{\alpha}_0(p)\right|^2 \tag{3.136}$$

can be justified. Together with Eq. (3.135) this implies $\|(1+p^2)\hat{\alpha}_0\|_{L^4(\mathbb{R}^3)}^4 \lesssim h^2 \|\alpha_0\|_{L^2(\mathbb{R}^3)}^2$. Next, we realize that as long as $r \in (2,4]$, one has $\|\hat{\alpha}_0\|_{L^r(\mathbb{R}^3)} \lesssim \left\|(1+p^2)^{1/4}\hat{\alpha}_0(p)\right\|_{L^4(\mathbb{R}^3)}$. The relation $\hat{\Delta}_0(p) = -2/(2\pi)^{3/2} \hat{V} * \hat{\alpha}_0(p)$, which follows from the Euler-Lagrange equations of the BCS functional, can be used to bound the $L^\infty(\mathbb{R}^3)$-norm of $\hat{\Delta}_0$ in terms of the $L^r(\mathbb{R}^3)$-norm of $\hat{\alpha}_0$. On the other hand, the gap equation together with the above bounds implies

$$\|\alpha_0\|_{L^2(\mathbb{R}^3)} \lesssim \left\|K_T^{\Delta_0}(p)^{-1}\right\|_{L^2(\mathbb{R}^3)} \left\|\hat{\Delta}_0\right\|_{L^\infty(\mathbb{R}^3)} \lesssim \|\hat{\alpha}_0\|_{L^r(\mathbb{R}^3)}. \tag{3.137}$$

Putting our estimates together, we finally obtain

$$\|\alpha_0\|_{L^2(\mathbb{R}^3)}^4 \lesssim h^2 \|\alpha_0\|_{L^2(\mathbb{R}^3)}^2 \tag{3.138}$$

which is what we wished to show.

The operator $V K_{T_c}^{-1}$ is compact, and hence the kernel of $K_{T_c} + V$ is finite dimensional. Accordingly, we can make the ansatz $\alpha_0(x) = \sum_{i=1}^m c_i(h) \alpha_*^i + \eta(x)$, where from the previous paragraph we know that $\sum_{i=1}^m |c_i(h)| + \|\eta\|_{L^2(\mathbb{R}^2)} \lesssim h$. We need to show that the $H^1(\mathbb{R}^3)$-norm of η is of order h^2. To that end, we insert the above ansatz of α_0 into Eq. (3.129) and take the inner product of the resulting expression with η, which yields

$$(\eta, (K_{T_c} + V + B)\,\eta) = -\sum_{i=1}^m c_i(h) \left(\eta, (A+B)\alpha_*^i\right) - (\eta, A\eta). \tag{3.139}$$

Let κ be the smallest eigenvalue of $K_{T_c} + V$ above zero. We then have $(\eta, (K_{T_c} + V)\,\eta) \geq \kappa\,(\eta,\eta)$. In fact, one can do better and take a small part of K_{T_c} and the inequality $K_{T_c} \geq C(1+p^2)$ to generate the $H^1(\mathbb{R}^3)$-norm of η. Hence, for some appropriately chosen $0 < \kappa' < \kappa$ we have

$$(\eta, (K_{T_c} + V + B)\,\eta) \geq \kappa' \|\eta\|_{H^1(\mathbb{R}^3)}^2. \tag{3.140}$$

In combination with Eq. (3.139), this implies

$$\begin{aligned} \|\eta\|_{H^1(\mathbb{R}^3)}^2 \leq{}& \frac{1}{\kappa'} \left(\|A(p)\|_{L^\infty(\mathbb{R}^3)} + \|B(p)\|_{L^\infty(\mathbb{R}^3)}\right) \left\|\sum_{i=1}^m c_i(h)\alpha_*^i\right\|_{L^2(\mathbb{R}^3)} \|\eta\|_{L^2(\mathbb{R}^3)} \\ &+ \frac{1}{\kappa'} \|A(p)\|_{L^\infty} \|\eta\|_{L^2(\mathbb{R}^3)}^2 \\ \lesssim{}& h \|\eta\|_{L^2(\mathbb{R}^3)} \left(h^2 + \left\|\hat{\Delta}_0\right\|_{L^\infty(\mathbb{R}^3)}\right) + h^2 \|\eta\|_{L^2(\mathbb{R}^3)}^2. \end{aligned} \tag{3.141}$$

To come to the expression in the last line of Eq. (3.141), we used the bounds on the $L^\infty(\mathbb{R}^3)$-norms of $A(p)$ and $B(p)$. Together with $\left\|\hat{\Delta}_0\right\|_{L^\infty} \lesssim \|\alpha_0\|_{L^2(\mathbb{R}^3)} \|V\|_{L^2(\mathbb{R}^3)} \lesssim h$, Eq. (3.141) gives

$$\|\eta\|^2_{H^1(\mathbb{R}^3)} \lesssim h^2 \|\eta\|_{L^2(\mathbb{R}^3)} + h^2 \|\eta\|^2_{L^2(\mathbb{R}^3)} . \tag{3.142}$$

This can easily be seen to imply $\|\eta\|_{H^1(\mathbb{R}^3)} \lesssim h^2$.

Next, we treat the bound for $\|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)}$. Eq. (3.129) can be written in the form

$$\begin{aligned}\hat{\eta}(p) = -K_{T_c}(p)^{-1}(A(p)+B(p)) \sum_{i=1}^m c_i(h)\hat{\alpha}^i_*(p) \\ - K_{T_c}(p)^{-1}\hat{V} * \hat{\eta}(p) - K_{T_c}(p)^{-1}(A(p)+B(p))\hat{\eta}(p)\end{aligned} \tag{3.143}$$

where we denote the symbol of the pseudo-differential operator K_{T_c} by $K_{T_c}(p)$. Using Eq. (3.143), we estimate

$$\begin{aligned}\|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)} \lesssim & \left\|K_{T_c}(p)^{-1}\right\|_{L^\infty(\mathbb{R}^3)} \Big(\|A(p)\|_{L^\infty(\mathbb{R}^3)} + \|B(p)\|_{L^\infty(\mathbb{R}^3)}\Big) \left\|\sum_{i=1}^m c_i(h)\hat{\alpha}^i_*\right\|_{L^\infty(\mathbb{R}^3)} \\ & + \left\|K_{T_c}(p)^{-1}\right\|_{L^\infty(\mathbb{R}^3)} \|V\|_{L^2(\mathbb{R}^3)} \|\eta\|_{L^2(\mathbb{R}^3)} \\ & + \left\|K_{T_c}(p)^{-1}\right\|_{L^\infty(\mathbb{R}^3)} \Big(\|A(p)\|_{L^\infty(\mathbb{R}^3)} + \|B(p)\|_{L^\infty(\mathbb{R}^3)}\Big) \|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)} \\ \lesssim & \, h^2 + h\|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)} .\end{aligned} \tag{3.144}$$

To come to the last line, we used $\left\|\hat{\Delta}_0(p)\right\|_{L^\infty(\mathbb{R}^3)} \lesssim h$ and $\|\hat{\alpha}^i_*(p)\|_{L^\infty(\mathbb{R}^3)} < \infty$ for $i = 1, ..., m$. Obviously, Eq. (3.144) implies $\|\hat{\eta}\|_{L^\infty(\mathbb{R}^3)} \lesssim h^2$. To obtain the estimate for $\hat{\alpha}^i_*(p)$, we use the equation $(K_{T_c} + V)\alpha^i_* = 0$ and obtain

$$\left\|\hat{\alpha}^i_*(p)\right\|_{L^\infty(\mathbb{R}^3)} = \left\|K_{T_c}(p)^{-1}\hat{V} * \hat{\alpha}^i_*(p)\right\|_{L^\infty(\mathbb{R}^3)} \lesssim \|V\|_{L^2(\mathbb{R}^3)} \left\|\alpha^i_*\right\|_{L^2(\mathbb{R}^3)} . \tag{3.145}$$

Except for the fact that not all $c_i(h), i = 1, ..., m$ vanish, which we show below, this proves the first part of the claim. In order to show the estimates for the Sobolev-norms of η, one considers Eq. (3.143) and uses ideas from above as well as ideas used in the proof of Lemma 3.5. Since no additional difficulties occur we leave this part of the proof to the reader.

To show that not all $c_i(h), i = 1, ..., m$ vanish, we assume the contrary. In this case, the gap equation implies

$$\kappa \|\eta\|_{L^2(\mathbb{R}^3)} \leq (\eta, (K_{T_c} + V)\,\eta) \leq |(\eta, (A+B)\eta)| . \tag{3.146}$$

If one uses the above bounds on the operator norms of A and B, it is not hard to see that the right-hand side of Eq. (3.146) is bounded by a constant times $h\|\eta\|_{L^2(\mathbb{R}^3)}$. This implies a contradiction and ends our proof of Lemma 3.7. □

3.3.2 Properties of α_0^w

In this section we prove two statements concerning $\alpha_0^w(x,y)$ that are used frequently in Section 3.2.

Lemma 3.8 *Let $V \in H^1(\mathbb{R}^3) \cap W^{1,\infty}(\mathbb{R}^3)$, $\hat{V} \in L^1(\mathbb{R}^3)$ and $W \in H^1(\mathbb{R}^3) \cap W^{1,\infty}(\mathbb{R}^3)$. Then $\left\| \alpha_0^w(x,y) - h^{-3}\alpha_0\left(\frac{x-y}{h}\right) \right\|_{H^1(\mathbb{R}^6)} \lesssim h^{3/2}$.*

Proof. To prove the claim, we make use of a representation of the operator

$$\alpha_0^w = \left[\left(1 + e^{\beta H_0^w}\right)^{-1} \right]_{12} \tag{3.147}$$

in terms of a Cauchy integral. This is possible because the function $g(z) = (1 + e^z)^{-1}$ is analytic in the strip $\{z \in \mathbb{C} | \, |\mathrm{Im}(z)| < \pi\}$. Let

$$\mathcal{C}_R = \{r - i\pi/(2\beta), r \in [-R, R]\} \cup \{-r + i\pi/(2\beta), r \in [-R, R]\}. \tag{3.148}$$

Then

$$\alpha_0^w = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\mathcal{C}_R} g(\beta z) \left[\frac{1}{z - H_0^w} \right]_{12} \mathrm{d}z, \tag{3.149}$$

where the limit $R \to \infty$ is to be taken in the weak operator topology. For further details on the construction of the above integral, see [28, p. 696, p. 704]. In the following, we will often write

$$\alpha_0^w = \frac{1}{2\pi i} \int_{\mathcal{C}} g(\beta z) \left[\frac{1}{z - H_0^w} \right]_{12} \mathrm{d}z, \tag{3.150}$$

where $\mathcal{C} = \cup_{R \geq 0} \mathcal{C}_R$ to denote the above limit. Since $\alpha_0^w = \alpha_0 + \tilde{\alpha}_0^w$, we have

$$\tilde{\alpha}_0^w = \frac{1}{2\pi i} \int_{\mathcal{C}} g(\beta z) \left[\frac{1}{z - H_0^w} - \frac{1}{z - H_0} \right]_{12} \mathrm{d}z. \tag{3.151}$$

To derive the bound we are interested in, we expand the resolvent of H_0^w in the following way

$$\frac{1}{z - H_0^w} = \frac{1}{z - H_0} + \frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0} + \frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0^w} \tag{3.152}$$

and treat the different contributions term by term. Since $\tilde{\alpha}_0^w(x,y) = \tilde{\alpha}_0^w(y,x)$, it is sufficient to consider derivatives acting on the x-component.

In order to show that the Hilbert-Schmidt norm

$$\left\| (-ih\nabla_j) \frac{1}{2\pi} \int_{\mathcal{C}} g(\beta z) \left[\frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0} \right]_{12} \mathrm{d}z \right\|_2 \tag{3.153}$$

for $j = 1, 2, 3$ is finite, we compute the kernel of the operator defined by the integral explicitly. Note that H_0 is a multiplication operator in Fourier space, which allows us to compute its resolvent:

$$\frac{1}{z - H_0} = \frac{1}{(z-E)(z+E)} \begin{pmatrix} z+k & \Delta_0 \\ \Delta_0 & z-k \end{pmatrix}. \tag{3.154}$$

This gives

$$\begin{aligned} \left[\frac{1}{z-H_0} h^2 \omega \frac{1}{z-H_0} \right]_{12} (p,q) = & \frac{(z + k(hp)) h^2 \hat{W}(p-q) \hat{\Delta}_0(hq)}{\left(z^2 - E(hp)\right)^2 \left(z^2 - E(hq)^2\right)} \\ & - \frac{\hat{\Delta}_0(hp) h^2 \hat{W}(p-q)(z - k(hq))}{\left(z^2 - E(hp)\right)^2 \left(z^2 - E(hq)^2\right)}. \end{aligned} \tag{3.155}$$

Having the explicit expression for this kernel at hand, we can compute the kernel of the corresponding Cauchy integral:

$$\begin{aligned} \left(\frac{1}{2\pi i} \int_{\mathcal{C}} g(\beta z) \left[\frac{1}{z-H_0} h^2 \omega \frac{1}{z-H_0} \right]_{12} \mathrm{d}z \right) (p,q) \\ = h^2 \hat{W}(p-q) \frac{\beta}{2} \frac{g_0(\beta E(hq)) - g_0(\beta E(hp))}{E(hp) - E(hq)} \\ \times \frac{\hat{\Delta}(hq) k(hp) + \hat{\Delta}(hp) k(hq)}{E(hp) + E(hq)}. \end{aligned} \tag{3.156}$$

Here $g_0(z) = \tanh(z/2)/z$. We must now bound the Hilbert-Schmidt norm of the operator $(-ih\nabla_j)$ acting on the operator defined by this kernel, which can be done as follows:

$$\begin{aligned} & \left\| hp_j h^2 \hat{W}(p-q) \frac{\beta}{2} \frac{g_0(\beta E(hq)) - g_0(\beta E(hp))}{E(hp) - E(hq)} \frac{\hat{\Delta}(hq) k(hp) + \hat{\Delta}(hp) k(hq)}{E(hp) + E(hq)} \right\|_{L^2(\mathbb{R}^6)} \\ & \leq \frac{\beta}{2} \left\| \frac{g_0(\beta E(hq)) - g_0(\beta E(hp))}{E(hp) - E(hq)} \right\|_{L^\infty(\mathbb{R}^6)} \left(\left\| hp_j h^2 \hat{W}(p-q) \frac{\hat{\Delta}(hq) k(hp)}{E(hp) + E(hq)} \right\|_{L^2(\mathbb{R}^6)} \right. \\ & \qquad \left. + \left\| hp_j h^2 \hat{W}(p-q) \frac{\hat{\Delta}(hp) k(hq)}{E(hp) + E(hq)} \right\|_{L^2(\mathbb{R}^6)} \right). \end{aligned} \tag{3.157}$$

Using $[g_0(x) - g_0(y)]/(x-y) = \int_0^1 g_0'(x - t(x-y))\mathrm{d}t$, one easily verifies that the $L^\infty(\mathbb{R}^6)$-norm on the right-hand side of Eq. (3.157) is bounded independently of h. The first $L^2(\mathbb{R}^3)$-norm on the right-hand side of the same equation obeys the desired bound because

$$\begin{aligned} & \left\| hp_j h^2 \hat{W}(p-q) \frac{\hat{\Delta}(hq) k(hp)}{E(hp) + E(hq)} \right\|_{L^2(\mathbb{R}^6)} \leq \left\| \frac{k(hp)}{E(hp) + E(hq)} \right\|_{L^\infty(\mathbb{R}^6)} \\ & \times \left(h^3 \left\| p_j \hat{W}(p) \right\|_{L^2(\mathbb{R}^3)} h^{-3/2} \left\| \hat{\Delta}_0(q) \right\|_{L^2(\mathbb{R}^3)} + h^2 \left\| \hat{W}(p) \right\|_{L^2(\mathbb{R}^3)} h^{-3/2} \left\| q_j \hat{\Delta}_0(q) \right\|_{L^2(\mathbb{R}^3)} \right) \\ & \quad \lesssim h^{3/2}. \end{aligned} \tag{3.158}$$

To obtain the estimate, we performed the coordinate transformation $p \to p + q$ and applied Lemma 3.7. To see that the right-hand side of Eq. (3.158) is finite under our assumptions one has to use the fact that $\hat{\Delta}_0(p) = -2(2\pi)^{-3/2}\hat{V} * \hat{\alpha}_0(p)$ and apply Lemma 3.5. Using the same arguments, the second $L^2(\mathbb{R}^6)$-norm on the right-hand side of Eq. (3.157) is seen to obey the same bound.

To bound the contribution coming from the third term on the right-hand side of Eq. (3.152), we first have to compute the upper right component of the operator-valued matrix under investigation.

$$
\begin{aligned}
&\left[\frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0} h^2 \omega \frac{1}{z - H_0^w}\right]_{12} \\
&\qquad = \left[D(z+k) h^2 W D(z+k) - D\Delta_0 h^2 W D \Delta_0\right] h^2 W \left[\frac{1}{z - H_0^w}\right]_{12} \\
&\qquad\quad - \left[D(z+k) h^2 W D \Delta_0 - D\Delta_0 h^2 W D(z-k)\right] h^2 W \left[\frac{1}{z - H_0^w}\right]_{22}.
\end{aligned} \tag{3.159}
$$

For the sake of convenience, we have introduced the operator-valued function $D(z) = (z^2 - E^2)^{-1}$. Since the argument z is obvious we omit it in the following. We will only show how to bound the Cauchy integral of the first and the last terms on the right-hand side of Eq. (3.159). All remaining terms can be bounded with the same techniques. Let us start with the contribution coming from the first term. We estimate

$$
\begin{aligned}
&\left\| (-ih\nabla_j)\, D(z+k) h^2 W D(z+k) h^2 W \left(\frac{1}{z - H_0^w}\right)_{12} \right\|_2 \\
&\qquad \lesssim h^4 \left\| D(z+k) \right\|_\infty^2 \|W\|_{W^{1,\infty}(\mathbb{R}^3)} \left\| (-ih\nabla_j) W \left[\frac{1}{z - H_0^w}\right]_{12} \right\|_2 .
\end{aligned} \tag{3.160}
$$

To obtain the estimate, we have commuted $(-ih\nabla_j)$ to the right. The first factor on the right-hand side of Eq. (3.160) can be estimated by

$$
\|D(z+k)\|_\infty \lesssim \begin{cases} 1 & \text{for } r \gg 1, \\ \frac{1}{1+|r|} & \text{for } r \ll -1 \end{cases} \tag{3.161}
$$

and has to be understood to hold on the contour $\mathcal{C}$. By r we refer to the natural coordinates on $\mathcal{C}$. Let us also note that $g(\beta z)$ decays exponentially for $r \gg 1$ and that it is bounded by 1 for $r \ll -1$. It remains to give a bound on the Hilbert-Schmidt norm for the right-hand side of Eq. (3.160). We write $\left[\frac{1}{z-H_0^w}\right]_{12} = \left[\frac{1}{z-H_0} + \frac{1}{z-H_0} h^2 \omega \frac{1}{z-H_0^w}\right]_{12}$

and estimate

$$
\begin{aligned}
&\left\| (-ih\nabla_j)\, W \left(D\Delta_0 + D(z+k)h^2 W \left[\frac{1}{z-H_0^w}\right]_{12} - D\Delta_0 h^2 W \left[\frac{1}{z-H_0^w}\right]_{22} \right) \right\|_2 \\
&\qquad \lesssim \|W\|_{H^1(\mathbb{R}^3)} \Bigg\{ h^{-3/2} \left\| p_j \hat{\Delta}_0(p) \right\|_{L^2(\mathbb{R}^3)} \frac{1}{1+|r|} \\
&\qquad\qquad + h^2 \left\| p_j \frac{z+k(p)}{(z^2-E(p)^2)} \right\|_{L^\infty(\mathbb{R}^3)} \left\| W \left[\frac{1}{z-H_0^w}\right]_{12} \right\|_2 \\
&\qquad\qquad + h^{1/2} \|W\|_{H^1(\mathbb{R}^3)} \left\| p_j \hat{\Delta}_0(p) \right\|_{L^2(\mathbb{R}^3)} \left\| \left[\frac{1}{z-H_0^w}\right]_{22} \right\|_\infty \Bigg\}.
\end{aligned} \tag{3.162}
$$

To come to the right-hand side of the above equation, we used the Seiler-Simon inequality which is this special case reads $\|a(-i\nabla)b(x)\|_2 \leq (2\pi)^{-3/2} \|a\|_{L^2(\mathbb{R}^3)} \|b\|_{L^2(\mathbb{R}^3)}$ for two functions $a, b \in L^2(\mathbb{R}^3)$. To estimate the second term on the right-hand side of Eq. (3.162), we note that

$$
\left\| \frac{p_j}{z-E(p)} \right\|_{L^\infty(\mathbb{R}^3)} \lesssim \begin{cases} |r|^{1/2} & \text{for } r \gg 1, \\ \left(\frac{1}{1+|r|}\right)^{1/2} & \text{for } r \ll -1. \end{cases} \tag{3.163}
$$

On the other hand, when we expand $(z-H_0^w)^{-1}$ another time and use the bound

$$
\left\| \frac{1}{z-E(hp)} \right\|_{L^2(\mathbb{R}^3)} \lesssim h^{-3/2} \begin{cases} |r|^{1/4} & \text{for } r \gg 1, \\ \left(\frac{1}{1+|r|}\right)^{1/4} & \text{for } r \ll -1, \end{cases} \tag{3.164}
$$

it can easily be checked that

$$
\left\| W \left[\frac{1}{z-H_0^w}\right]_{12} \right\|_2 \lesssim h^{-3/2} \begin{cases} |r|^{3/4} & \text{for } r \gg 1, \\ \left(\frac{1}{1+|r|}\right)^{3/4} & \text{for } r \ll -1. \end{cases} \tag{3.165}
$$

Combining the estimates from Eq. (3.160) - Eq. (3.165), we obtain that the Cauchy integral of the first term on the right-hand side of Eq. (3.159) is bounded by a constant times $h^{5/2}$.

It remains to estimate the contribution coming from the last term on the right-hand side of Eq. (3.159). We compute

$$
\begin{aligned}
&\left\| (-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z-H_0^w}\right]_{22} \right\|_2 \\
&\lesssim h^{5/2} \left\| p_j \hat{\Delta}_0(p) \right\|_{L^2(\mathbb{R}^3)} \|W\|_{L^2(\mathbb{R}^3)} \|D\|_\infty \|D(z-k)\|_\infty \|W\|_{L^\infty(\mathbb{R}^3)} \left\| \left[\frac{1}{z-H_0^w}\right]_{22} \right\|_\infty \\
&\lesssim h^{5/2} \begin{cases} \left(\frac{1}{1+|r|}\right)^3 & \text{for } r \gg 1, \\ \frac{1}{1+|r|} & \text{for } r \ll -1, \end{cases}
\end{aligned} \tag{3.166}
$$

where, as above, we used the Seiler-Simon inequality. To come to the last line, we used the estimate

$$\left\| \left[\frac{1}{z - H_0^w} \right]_{22} \right\|_\infty \lesssim \begin{cases} \frac{1}{1+|r|} & \text{for } r \gg 1, \\ 1 & \text{for } r \lll -1, \end{cases} \tag{3.167}$$

which can easily be justified by expanding $(z - H_0^w)^{-1}$ another time. Since the function g is exponentially decaying for $r \gg 1$ but is only bounded by 1 for $r \leq 0$, we cannot control the Cauchy integral over the part of $\mathcal{C}$ on which $r \ll -1$ holds because we only have a decay of $(1 + |r|)^{-1}$ in that direction.

To circumvent this problem, we use an algebraic identity for the function g, namely $g(\beta z) = -g(-\beta z) + 1$. Hence, we have

$$\begin{aligned} &\left\| \frac{1}{2\pi i} \int_{\mathcal{C}} g(\beta z)(-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z - H_0^w} \right]_{22} \mathrm{d}z \right\|_2 \\ &\qquad \leq \left\| \frac{1}{2\pi i} \int_{\mathcal{C}} g(-\beta z)(-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z - H_0^w} \right]_{22} \mathrm{d}z \right\|_2 \\ &\qquad\quad + \left\| \frac{1}{2\pi i} \int_{\mathcal{C}} (-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z - H_0^w} \right]_{22} \mathrm{d}z \right\|_2 . \end{aligned} \tag{3.168}$$

The first term on the right-hand side of Eq. (3.168) can be estimated using Eq. (3.166) and yields

$$\left\| \frac{1}{2\pi i} \int_{\mathcal{C}} g(-\beta z)(-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z - H_0^w} \right]_{22} \mathrm{d}z \right\|_2 \lesssim h^{5/2}. \tag{3.169}$$

The second term on the right-hand side of Eq. (3.160) equals zero which can be seen with the help of the subsequent argument. Let $\psi, \phi \in L^2(\mathbb{R}^3)$ be such that $\hat{\psi} \in \mathcal{C}_c^\infty(\mathbb{R}^3)$. Then

$$\begin{aligned} &\int_{\mathcal{C}} \left(\psi, (-ih\nabla_j) D\Delta_0 h^2 W D(z-k) h^2 W \left[\frac{1}{z - H_0^w} \right]_{22} \phi \right) \mathrm{d}z \\ &= \int_{\mathcal{C}} \left(\int_{\mathbb{R}^9} \overline{\hat{\psi}(p)} \frac{hp_j \hat{\Delta}(hp)}{z^2 - E(hp)^2} h^2 \hat{W}(p-r) \frac{z - k(hr)}{z^2 - E(hr)^2} h^2 \hat{W}(r-q) \hat{\phi}_z(q) \mathrm{d}(p,r,q) \right) \mathrm{d}z \end{aligned} \tag{3.170}$$

where $\phi_z = \left[\frac{1}{z - H_0^w} \right]_{22} \phi \in L^2(\mathbb{R}^3)$. The map $z \mapsto \phi_z$ is a vector-valued analytic function on the set $\mathbb{C} \setminus \mathbb{R}$. A short introduction to vector-valued and operator-valued analytic functions can be found in [53, Chapter 3].

We now check that the expression on the right-hand side of Eq. (3.170) is finite. To that end, we compute

$$\begin{aligned}
&\int_{\mathbb{R}^9} \left| \overline{\hat{\psi}(p)} \frac{hp_j \hat{\Delta}(hp)}{z^2 - E(hp)^2} h^2 \hat{W}(p-r) \frac{z - k(hr)}{z^2 - E(hr)^2} h^2 \hat{W}(r-q) \hat{\phi}_z(q) \right| \mathrm{d}(p,r,q) \\
&\quad \lesssim h^4 \left\| \overline{\hat{\psi}(p)} \frac{hp_j \hat{\Delta}(hp)}{z^2 - E(hp)^2} \right\|_{L^2(\mathbb{R}^3)} \left\| \hat{W} \right\|^2_{L^1(\mathbb{R}^3)} \left\| \frac{z - k(hr)}{z^2 - E(hr)^2} \right\|_{L^\infty(\mathbb{R}^3)} \left\| \hat{\phi}_z(q) \right\|_{L^2(\mathbb{R}^3)} \\
&\quad \lesssim h^4 \begin{cases} \left(\frac{1}{1+|r|} \right)^3 & \text{for } r \gg 1, \\ \left(\frac{1}{1+|r|} \right)^2 & \text{for } r \ll -1. \end{cases}
\end{aligned} \tag{3.171}$$

To come to the last line, we explicitly used that the support of $\hat{\psi}$ is finite. Accordingly, the constant in the last line of Eq. (3.171) depends on the size of the support of $\hat{\psi}$. The bound is strong enough to guarantee that the Cauchy integral on the right-hand side of Eq. (3.170) yields a finite value as long as ψ is fixed. Additionally, the bound in Eq. (3.171) shows that this term is continuous in W in the sense that if W_n is a sequence of functions in $L^2(\mathbb{R}^3)$ such that $\hat{W}_n \to \hat{W}$ strongly in $L^1(\mathbb{R}^3)$ then the corresponding Cauchy integrals converge. Hence, we can approximate W with a sequence W_n such that each $\hat{W}_n$ has compact support.

Next, we go back to Eq. (3.170) and examine the inner product under the Cauchy integral, which defines an analytic function

$$f(z) = \big(\psi, (-ih\nabla_j) D(z) \Delta_0 h^2 W_n D(z)(z-k) h^2 W_n \phi_z\big) \tag{3.172}$$

on $\mathbb{C} \setminus \mathbb{R}$. This function has to be integrated over the contour $\mathcal{C}$, which has two parts, one in the upper half-plane and another one in the lower half-plane. Using the fact that $\hat{\psi}$ and $\hat{W}_n$ have compact support, one can easily justify the estimate

$$|f(z)| \lesssim \frac{1}{1+|y|} \frac{1}{1+|z|^3}, \tag{3.173}$$

where $z = x + iy$ and $y \neq 0$. Let us consider the integral in the upper half-plane. Standard arguments from complex analysis together with the estimate from Eq. (3.173) show that

$$\int_{-\infty}^{\infty} f(x+iy)\mathrm{d}x = \int_{-\infty}^{\infty} f(x+iy')\mathrm{d}x \tag{3.174}$$

holds for all $y, y' > 0$. On the other hand,

$$\left| \int_{-\infty}^{\infty} f(x+iy)\mathrm{d}x \right| \lesssim \frac{1}{1+|y|}, \tag{3.175}$$

which, together with Eq. (3.174), implies that the absolute value of the integral $\int_{-\infty}^{\infty} f(x+iy)\mathrm{d}x$ is smaller than any given positive number, and therefore equals zero. The same argument can be done for the integral in the lower half-plane. Since ψ has been chosen from a dense subset of $L^2(\mathbb{R}^3)$, this shows that the second term on the right-hand side of Eq. (3.168) equals zero and ends the proof of Lemma 3.8. □

Next, we investigate the decay properties of the function $\alpha_0^w(x,y) - h^{-3}\alpha_0\left(\frac{x-y}{h}\right)$.

Lemma 3.9 *Let* $V \in L^2(\mathbb{R}^3)$, $\hat{V} \in L^1(\mathbb{R}^3) \cap H^2(\mathbb{R}^3) \cap W^{2,\infty}(\mathbb{R}^3)$ *as well as* $(1+x^2)W \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ *and* $\hat{W} \in L^1(\mathbb{R}^3)$. *Then*

$$\|\hat{\alpha}_0^w(p,q) - \hat{\alpha}_0\left(h(p-q)\right)\|_{H^2(\mathbb{R}^6)} \lesssim h^{3/2} \tag{3.176}$$

holds.

Proof. The proof goes along the same lines as the one of Lemma 3.8. The only difference is that we now have to commute x^2 until it stands next to W, while in the previous proof we commuted $(-ih\nabla_j)$ until it stood next to Δ_0. Since no additional difficulties arise, we leave the proof to the reader. □

CHAPTER 4

No translational symmetry breaking in the BCS model with radial pair interaction

We consider the two-dimensional BCS functional with a radial pair interaction. We show that the translational symmetry is not broken in a certain temperature interval below the critical temperature. Our result carries over to the three-dimensional case if the Cooper-pairs have vanishing angular momentum.

4.1 Introduction

In 1957 Bardeen, Cooper and Schrieffer published their famous paper with the title "Theory of Superconductivity", which contained the first, generally accepted, microscopic theory of superconductivity. In recognition of this work, they were awarded the Nobel prize in 1972. Originally introduced to describe the phase transition from the normal to the superconducting state in metals and alloys, BCS theory can also be applied to describe the phase transition to the superfluid state in cold fermionic gases. In this situation, one has to replace the usual non-local phonon-induced interaction in the gap equation by a local pair potential. Apart from being a paradigmatic model in solid state physics and in the field of cold quantum gases, the BCS theory of superconductivity, that is, the gap equation and the BCS functional show a rich mathematical structure, which has been well recognized. See [9, 10, 11, 12, 13, 14] for works on the gap equation with interaction kernels suitable to describe the physics of conduction electrons in solids and [16, 17, 18, 19, 21, 25, 32] for works that treat the translation-invariant BCS functional with a local pair interaction. One main question in the study of BCS theory is whether the gap equation

$$\Delta(p) = -\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{V}(p-q) \frac{\tanh\left(\frac{E(q)}{2T}\right)}{E(q)} \Delta(q) \mathrm{d}q, \tag{4.1}$$

with $E(q) = \sqrt{(q^2-\mu)^2 + |\Delta(q)|^2}$ has a non-trivial solution, that is, one with $\Delta \neq 0$. In [16] it has been demonstrated that, although the gap equation is highly non-linear, this can be decided with the help of a linear criterion. To be more precise, it was shown that the existence of a non-trivial solution of the gap equation is equivalent to the fact that a certain linear operator has a negative eigenvalue. Based on a characterization of the critical temperature in terms of this linear operator, its behavior has been investigated in the limit of small couplings and in the low-density limit, see [17, 19] and [20], respectively. Recently, there has also been considerable interest in the BCS functional with external fields, and in particular, in its connection to the Ginzburg-Landau theory of superconductivity, see [22, 23, 24, 28, 29, 33, 34].

In this work we consider the two-dimensional BCS functional with a radial pair interaction and show that there exists a certain temperature interval below the critical temperature, in which the translational symmetry of the system is not broken. Our analysis carries over to the three-dimensional case if the Cooper-pairs are in an s-wave state. Prior to this work, such a result was known only in the case of $\hat{V} \leq 0$ and not identically zero, see [31]. Apart from these considerations, we show for the two-dimensional model, that Cooper-pairs are in an angular momentum eigenstate if the temperature lies in the temperature interval mentioned above. A similar result in three spatial dimensions allows us to determine a situation, in which the Cooper-pairs are in an s-wave state.

4.2 Main results

We consider a two-dimensional periodic sample of fermionic atoms filling all of $\mathbb{R}^d$, $d \geq 1$, within the framework of BCS theory. BCS states are most conveniently described by their generalized one-particle density matrix, that is, by an operator $\Gamma \in \mathcal{L}\left(L^2(\mathbb{R}^d) \oplus L^2(\mathbb{R}^d)\right)$ of the form

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \overline{\alpha} & 1-\overline{\gamma} \end{pmatrix}, \tag{4.2}$$

for which $0 \leq \Gamma \leq 1$ holds. In the above equation, the complex conjugate operators are defined by $\overline{\gamma} = C\gamma C$, where C denotes complex conjugation. In terms of integral kernels, this translates to $\overline{\gamma}(x,y) = \overline{\gamma(x,y)}$. The condition $0 \leq \Gamma \leq 1$ implies $0 \leq \gamma \leq 1$, $\alpha\overline{\alpha} \leq \gamma(1-\gamma)$ as well as $\alpha^* = \overline{\alpha}$, where again in terms of integral kernels the last statement reads $\alpha(x,y) = \alpha(y,x)$. Our BCS states are assumed to be periodic with period one, which means, that γ and α commute with translations by vectors $v \in \mathbb{Z}^d$. For the integral kernels, this implies $\gamma(x+v, y+v) = \gamma(x,y)$ for all $v \in \mathbb{Z}^d$ and the same for α. Let $\Omega = [0,1]^d$ and let χ_Ω be the characteristic function of Ω. For a periodic operator A, we introduce the trace per unit volume by $\mathrm{Tr}_\Omega\,[A] = \mathrm{Tr}\,[\chi_\Omega(x) A \chi_\Omega(x)]$. Obviously, the location of the cube Ω does not play any role. A periodic BCS state is called *admissible* if $\mathrm{Tr}_\Omega\,[(-\nabla^2+1)\gamma] < \infty$ holds. The set of all such BCS states will be denoted by $\mathcal{D}^{per}$. For $T \geq 0$, the periodic BCS functional with chemical potential $\mu \in \mathbb{R}$, interaction potential V and entropy

$$S(\Gamma) = -\frac{1}{2}\,\mathrm{Tr}_\Omega[\Gamma \log \Gamma + (1-\Gamma)\log(1-\Gamma)], \tag{4.3}$$

is given by

$$\mathcal{F}^{\mathrm{per}}(\Gamma) = \mathrm{Tr}_\Omega[(-\nabla^2-\mu)\gamma] + \int_{\Omega\times\mathbb{R}^d} V(x-y)|\alpha(x,y)|^2 \,\mathrm{d}(x,y) - TS(\Gamma). \tag{4.4}$$

In case of additional external electric and magnetic fields, this BCS functional and its connection to the Ginzburg-Landau theory of superconductivity has been studied in [28, 29].

The generalized one-particle density matrix of a translation-invariant BCS state, that is, a state commuting with all translations in $\mathbb{R}^d$, is a matrix-valued pseudo-differential operator of the form

$$\Gamma(-i\nabla) = \begin{pmatrix} \gamma(-i\nabla) & \hat{\alpha}(-i\nabla) \\ \overline{\hat{\alpha}(-i\nabla)} & 1-\overline{\gamma(-i\nabla)} \end{pmatrix}, \tag{4.5}$$

For those states, the condition $0 \leq \Gamma \leq 1$ translates to $|\hat{\alpha}(p)|^2 \leq \gamma(p)(1-\gamma(p))$ for almost all $p \in \mathbb{R}^d$. The one-particle density matrix γ and the Cooper-pair wave function α of a translation-invariant *admissible* BCS state satisfy $\gamma \in L^1\left(\mathbb{R}^d, (1+p^2)\mathrm{d}p\right)$ and $\alpha \in H^1_{sym}(\mathbb{R}^d) = \left\{\alpha \in H^1(\mathbb{R}^d) | \alpha(x) = \alpha(-x)\right\}$. In particular, γ is a real function. By $\mathcal{D}$ we denote the set of all translation-invariant *admissible* BCS states. We will, with a slight abuse of notation, write $\Gamma \in \mathcal{D}$ as well as $(\gamma, \alpha) \in \mathcal{D}$ if $\gamma = \Gamma_{11}$ and $\alpha = \Gamma_{12}$. In the case of translation-invariant states, the entropy takes the form

$$S(\Gamma) = -\frac{1}{2}\int_{\mathbb{R}^d} \mathrm{Tr}_{\mathbb{C}^2}\left[\Gamma(p)\log\Gamma(p) + (1-\Gamma(p))\log(1-\Gamma(p))\right]\mathrm{d}p, \tag{4.6}$$

and the BCS functional can be written as

$$\mathcal{F}(\Gamma) = \int_{\mathbb{R}^d}(p^2-\mu)\gamma(p)\,\mathrm{d}p + \int_{\mathbb{R}^d} V(x)|\alpha(x)|^2\,\mathrm{d}x - TS(\Gamma). \tag{4.7}$$

We call the BCS functional in the form of Eq. (4.7) the translation-invariant BCS functional. In [16] it was shown, that the three-dimensional version of $\mathcal{F}$ is bounded from below and attains its infimum in $\mathcal{D}$. The same results hold in two spatial dimensions and with symmetric Cooper-pair wave functions by analogous arguments.

Let K_T denote denote the pseudo-differential operator with symbol

$$K_T(p) = \frac{p^2-\mu}{\tanh\left(\frac{p^2-\mu}{2T}\right)}. \tag{4.8}$$

It was shown in [16, Theorem 1] that the linear operator $K_T + V$ has at least one negative eigenvalue if and only if the normal state is unstable, that is, the energy can be lowered by the formation of Cooper-pairs. Note that K_T is monotone increasing in T, which allows us to define the critical temperature for the translation-invariant BCS functional by

$$T_c = \inf\{T \geq 0 \mid K_T + V|_{sym} \geq 0\}. \tag{4.9}$$

Note that we view $K_T + V$ as an operator on $L^2_{sym}(\mathbb{R}^d)$. This is because our Cooper-pair wave functions are assumed to be symmetric, that is $\alpha(x) = \alpha(-x)$. From now on, we consider the case $d = 2$. For a radial interaction potential $V \in L^2(\mathbb{R}^2)$, the linear operator $K_T + V$ is rotation invariant. In particular, all eigenstates of $K_T + V$ are of the form $\hat{\alpha}_\ell(p) = e^{i\ell\theta}\sigma_\ell(|p|)$, for some $\ell \in 2\mathbb{Z}$, where θ denotes the angle of p in polar coordinates. We define the sector of states $\Gamma \in \mathcal{D}$, whose Cooper-pair wave function lies in the sector of angular momentum ℓ via the sets

$$\mathcal{D}_\ell = \left\{(\gamma, \alpha) \in \mathcal{D} \mid \gamma(p) = \tilde{\gamma}(|p|), \hat{\alpha}(p) = e^{i\ell\theta}\sigma_\ell(|p|) \text{ for all } p \in \mathbb{R}^2\right\}, \tag{4.10}$$

for $\ell \in 2\mathbb{Z}$. The fact that γ is chosen radial can be motivated with the Euler-Lagrange equation of the BCS functional. We call

$$\mathcal{F}_\ell = \mathcal{F}|_{\mathcal{D}_\ell} \tag{4.11}$$

the BCS functional on the sector of Cooper-pair wave functions of angular momentum ℓ. For each functional $\mathcal{F}_\ell$, we obtain a critical temperature $T_c(\ell)$. Let $\mathcal{H}_\ell = \{\alpha_\ell \in H^1_{sym}(\mathbb{R}^2) \mid \alpha_\ell(p) = e^{i\ell\theta}\sigma_\ell(|p|)\}$. Then the critical temperature $T_c(\ell)$ of the functional $\mathcal{F}_\ell$ is given by

$$T_c(\ell) = \inf\left\{T \geq 0 \mid (K_T + V)|_{\mathcal{H}_\ell} \geq 0\right\}. \tag{4.12}$$

Note that $T_c(-\ell) = T_c(\ell)$ for all $\ell \in 2\mathbb{Z}$. Moreover, T_c is given by

$$T_c = \max_{\ell \in 2\mathbb{Z}} T_c(\ell), \tag{4.13}$$

since $T_c \geq T_c(\ell)$ for all $\ell \in 2\mathbb{Z}$. Having these definitions at hand, we state our main result:

Theorem 4.1 *Let $V \in L^2(\mathbb{R}^2)$ with $\hat{V} \in L^r(\mathbb{R}^2)$, $r \in [1, 2)$, be radial and assume that $T_c > 0$. Suppose that there exist $\ell_0, \ell_1 \in 2\mathbb{Z}$ such that*

$$T_c(\ell_0) > T_c(\ell_1) \geq T_c(\ell) \tag{4.14}$$

for all $\ell \in 2\mathbb{Z} \setminus \{\pm\ell_0\}$. If $(\gamma_{\ell_0}, \alpha_{\ell_0}) \in \mathcal{D}_{\ell_0}$ is a minimizer of $\mathcal{F}_{\ell_0}$ (which always exists), then $(\gamma_{\ell_0}, \alpha_{\ell_0})$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0})$ minimize the full BCS functional $\mathcal{F}^{\text{per}}$ for $T \in [T_c(\ell_1), T_c)$. Moreover, $\sigma_{\ell_0} = \sigma_{-\ell_0}$ up to phases. For $T \in (T_c(\ell_1), T_c)$ these are the only minimizers of $\mathcal{F}^{\text{per}}$.

Remark 4.1 *If $\ell_0 = 0$ and $T \in (T_c(\ell_1), T_c)$ the BCS functional has, up to a phase in front of α_0, a unique minimizer (γ_0, α_0) with radial Cooper-pair wave function α_0. If $\ell_0 \neq 0$ it has two minimizers, namely $(\gamma_{\ell_0}, \alpha_{\ell_0}) \in \mathcal{D}_\ell$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0}) \in \mathcal{D}_{-\ell_0}$. We can choose α_{ℓ_0} and $\alpha_{-\ell_0}$ such that their radial parts coincide.*

Remark 4.2 *Note that the Fourier transform is a bijection from $\mathcal{H}_\ell$ to $\mathcal{H}_\ell$. To be more precise, the Fourier transform of a function $\alpha(x) = e^{i\ell\varphi}\sigma_\ell(|x|)$ is given by*

$$\hat{\alpha}(p) = e^{i\ell\theta}e^{i\ell\pi/2}\int_0^\infty |x|\,\sigma_\ell(|x|)J_\ell(|p||x|)\,\mathrm{d}|x|, \tag{4.15}$$

where φ and θ denote the angles of x and p in polar coordinates, respectively. By J_ℓ we denote the ℓ-th Bessel function.

Remark 4.3 *Having in mind that the linear operator $K_T + V$, which characterizes T_c, is related to the second variation of $\mathcal{F}$ at the normal state $\Gamma_n = (\gamma_n, 0)$ in direction of α by*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{F}(\gamma_n, t\alpha)\bigg|_{t=0} = 2\,(\alpha, (K_T + V)\alpha)\,, \tag{4.16}$$

one can understand Theorem 4.1 as follows. We find $T < T_c$ such that $K_T + V$ has exactly one negative eigenvalue λ_0. Hence the second variation is smallest (and, in particular, negative) if α is an element of the eigenspace of λ_0, and therefore one could hope to find a minimizer of $\mathcal{F}$, whose Cooper-pair wave function lies approximately in this eigenspace. In fact, Theorem 4.1 states that the minimizers of $\mathcal{F}$ for temperatures T in a certain temperature interval below $T_c = T_c(\ell_0)$ lie in exactly one/two specific angular momentum sectors, namely the ones with angular momentum ℓ_0 and $-\ell_0$. For $T = T_c(\ell_1)$ the second eigenvalue λ_1 and its eigenspace become important, since now elements in the eigenspace of λ_1 are also candidates for lowering the energy.

In the special situation where $\ell_0 = 0$, Theorem 4.1 also holds in three spatial dimensions. Let $\mathcal{H}_r = \left\{\alpha \in H^1_{sym}(\mathbb{R}^3) \mid \alpha(p) = \sigma_0(|p|)\right\}$ as well as

$$T' = \inf\left\{T \geq 0 \;\middle|\; (K_T + V)|_{\mathcal{H}_r^\perp} \geq 0\right\}. \tag{4.17}$$

Note that the orthogonal complement in the definition of T' is taken in $L^2_{sym}(\mathbb{R}^3)$. It should be compared to $T_c(\ell_1)$ in the two-dimensional case. For the three-dimensional model the following statement holds:

Theorem 4.2 *Let $V \in L^2(\mathbb{R}^3)$ with $\hat{V} \in L^r(\mathbb{R}^3)$ for some $r \in [1, 12/7)$ be radial and assume that $T_c > 0$. Assume further that zero is a non-degenerate eigenvalue of $K_{T_c} + V$. Then, for $T \in [T', T_c)$, there exists a pair (γ_0, α_0) with γ_0 and α_0 being radial functions, that minimizes the BCS functional $\mathcal{F}^{per}$. Moreover, up to a phase in front of α_0, (γ_0, α_0) is the only minimizer of $\mathcal{F}^{per}$ for $T \in (T'_c, T_c)$.*

Remark 4.4 *In case of $\hat{V} \leq 0$ and not identically zero, the kernel of $K_{T_c} + V$ is one-dimensional for all $d \geq 1$. Hence, the assumptions of Theorem 4.2 are satisfied for radial interaction potentials with this property.*

Remark 4.5 *For of a small interaction potential, the methods of [17, 19] can be used to decide in which angular momentum sector the ground state of $K_{T_c} + V$ lies. In this case, it is sufficient to consider a certain operator acting on functions on the Fermi sphere, whose eigenvalues can be computed explicitly. This method works in two and three spatial dimensions.*

4.3 Preparations

To prove Theorem 4.1, we will show that the minimizers of $\mathcal{F}_{\ell_0}$ and $\mathcal{F}_{-\ell_0}$ also minimize $\mathcal{F}^{per}$. The following considerations lay the basis for this approach. Let us start by introducing some notation. To each Cooper-pair wave function α, one usually associates a gap function Δ in the following way:

$$\Delta(p) = \frac{2}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{V}(p-q)\hat{\alpha}(q)\,\mathrm{d}q. \tag{4.18}$$

If α belongs to a minimizer of the translation-invariant BCS functional both functions contain the same amount of information. For $T > 0$, we define the operator K_T^Δ with the help of the function $E(p) = \sqrt{(p^2-\mu)^2 + |\Delta(p)|^2}$ by

$$K_T^\Delta = \frac{E(-i\nabla)}{\tanh\left(E(-i\nabla)/(2T)\right)}. \tag{4.19}$$

The symbol of K_T^Δ will be denoted by $K_T^\Delta(p)$. The Euler-Lagrange equations of $\mathcal{F}$ are most conveniently formulated in terms of $K_T^\Delta(p)$ and read

$$\begin{aligned} \gamma(p) &= \frac{1}{2} - \frac{p^2-\mu}{2K_T^\Delta(p)}, \\ \hat{\alpha}(p) &= -\frac{\Delta(p)}{2K_T^\Delta(p)}, \end{aligned} \tag{4.20}$$

see [16, 31]. The equation for α is referred to as the gap equation and is often written in the form $(K_T^\Delta + V)\alpha = 0$. It is equivalent to the formulation of the gap equation in the introduction, see Eq. (4.1). Another useful form of these equations is given by

$$\begin{aligned} \Gamma(p) &= \begin{pmatrix} \gamma(p) & \hat{\alpha}(p) \\ \overline{\hat{\alpha}(p)} & 1-\overline{\gamma(p)} \end{pmatrix} = \frac{1}{1+e^{\beta H_\Delta(p)}}, \\ H_\Delta(p) &= \begin{pmatrix} p^2-\mu & \Delta(p) \\ \overline{\Delta(p)} & -(p^2-\mu) \end{pmatrix}. \end{aligned} \tag{4.21}$$

Note that H_Δ depends only on α and not on γ.

We continue with the first Lemma, which captures some properties of the functional $\mathcal{F}_\ell$.

Lemma 4.1 *The BCS functional $\mathcal{F}_\ell$, $\ell \in 2\mathbb{Z}$, is bounded from below and attains its infimum in $\mathcal{D}_\ell$.*

Proof. Boundedness from below of $\mathcal{F}_\ell$ follows from the fact that $\mathcal{F}$ is bounded from below. As in the proof of [16, Lemma 1] we find a minimizing sequence $\left(\gamma_\ell^{(n)}, \hat{\alpha}_\ell^{(n)}\right)$ in the convex set $\mathcal{D}_\ell$, that converges strongly in $L^p(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ to $(\gamma, \hat{\alpha})$ for any $p \in (1, \infty)$, as n tends to infinity. It is an easy consequence that $(\gamma, \hat{\alpha}) \in \mathcal{D}_\ell$. □

Next, we determine the Euler-Lagrange equations of the functionals $\mathcal{F}_\ell$. The existence of a non-trivial solution of one of these equations in the relevant temperature interval is a key ingredient of our proof of Theorem 4.1.

Lemma 4.2 *The Euler-Lagrange equation of the functional $\mathcal{F}_\ell$, $\ell \in 2\mathbb{Z}$, is the one of the translation-invariant BCS functional $\mathcal{F}$.*

Proof. The proof is analogous to the one of [16, Lemma 1], which states that for $T > 0$ any minimizer $(\gamma, \alpha) \in \mathcal{D}$ of $\mathcal{F}$ satisfies the pair of equations

$$\begin{aligned}\left(\hat{V} * \hat{\alpha}\right)(p) &= \left(p^2 - \mu\right) \frac{\hat{\alpha}(p)}{2\gamma(p) - 1}, \\ (p^2 - \mu) &= -2T(\gamma(p) - 1) f\left(2\sqrt{(\gamma(p) - 1/2)^2 + |\hat{\alpha}(p)|^2}\right),\end{aligned} \tag{4.22}$$

for almost every $p \in \mathbb{R}^2$, where $f(a) = \frac{1}{a}\log\frac{1+a}{1-a}$ for $0 \leq a < 1$. The proof applies also for minimizers $(\gamma_\ell, \alpha_\ell) \in \mathcal{D}_\ell$ of $\mathcal{F}_\ell$, because for $\hat{g}$ of the form $\hat{g}(p) = e^{i\ell\varphi}\tilde{g}(|p|)$ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\gamma_\ell, \alpha_\ell + tg)\bigg|_{t=0} &= 2\,\mathrm{Re}\int_{\mathbb{R}^2} V(x)\overline{g(x)}\alpha_\ell(x)\,\mathrm{d}x \\ &\quad + 2T\,\mathrm{Re}\int_{\mathbb{R}^2} \overline{\hat{g}(p)}\alpha_\ell(p) f\left(2\sqrt{(\gamma_\ell(p) - 1/2)^2 + |\hat{\alpha}(p)|^2}\right)\mathrm{d}p.\end{aligned} \tag{4.23}$$

Since both integrands are radial functions the reasoning from the proof of [16, Lemma 1] goes through. It can easily be seen that Eq. (4.22) is equivalent to the form of the Euler-Lagrange equations we have introduced above, see [31]. □

The following lemma clarifies the connection between minimizers of $\mathcal{F}_\ell$ and $\mathcal{F}_{-\ell}$.

Lemma 4.3 *If $(\gamma_\ell, \alpha_\ell) \in \mathcal{D}_\ell$ is a minimizer of $\mathcal{F}_\ell$, then $(\gamma_\ell, \alpha_{-\ell}) \in \mathcal{D}_{-\ell}$ is a minimizer of $\mathcal{F}_{-\ell}$ and $\sigma_\ell = \sigma_{-\ell}$ up to phases.*

Proof. The first and the last term on the right hand side of equation Eq. (4.7) do not depend on ℓ. We insert $\hat{\alpha}_\ell(p) = e^{i\ell\theta}\sigma_\ell(|p|)$ into the interaction term and find

$$\begin{aligned}&\int_{\mathbb{R}^2} |\alpha_\ell(x)|^2 V(x)\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_0^\infty\int_0^\infty e^{i\ell(\theta_q - \theta_p)}\overline{\sigma_\ell(|p|)}\hat{V}\left(|p-q|\right)\sigma_\ell(|q|)\,\mathrm{d}|p|\mathrm{d}|q|\mathrm{d}\theta_p\mathrm{d}\theta_q.\end{aligned} \tag{4.24}$$

Observe that $|p - q| = \sqrt{p^2 + q^2 - 2|p||q|\cos(\theta_q - \theta_p)}$. Substituting θ_q by $-\theta_q$ and θ_p by $-\theta_p$ finishes the proof. □

4.4 Proof of Theorem 4.1 and 4.2

Let Γ_{ℓ_0} be a minimizer of $\mathcal{F}_{\ell_0}$. Our aim is to show that $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) \geq 0$ holds for all $\Gamma \in \mathcal{D}^{per}$. To do so, we need an inequality for a difference of two free energies.

Assume that Γ_ℓ is a minimizer of the functional $\mathcal{F}_\ell$ with $(\Gamma_\ell)_{11} = \gamma_\ell$, $(\Gamma_\ell)_{12} = \alpha_\ell$ and let $\varphi : (0,1) \to \mathbb{R}$ be given by $\varphi(x) = x\ln(x) + (1-x)\ln(1-x)$. The difference $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell)$ reads

$$\begin{aligned}\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) = \mathrm{Tr}_\Omega\left[\left(-\nabla^2 - \mu\right)(\gamma - \gamma_\ell)\right] \\ + \int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y)|^2 - |\alpha_{\ell_0}(x-y)|^2\right)\,\mathrm{d}(x,y) \\ + T\,\mathrm{Tr}_\Omega\left[\varphi(\Gamma) - \varphi(\Gamma_\ell)\right].\end{aligned} \tag{4.25}$$

First, we complete the square in the difference of the interaction terms, which yields

$$\begin{aligned}&\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y)|^2 - |\alpha_{\ell_0}(x-y)|^2\right)\,\mathrm{d}(x,y) \\ &\quad= \int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y) - \alpha_{\ell_0}(x-y)|^2\right)\,\mathrm{d}(x,y) \\ &\quad\quad - 2\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha_{\ell_0}(x-y)|^2 - \mathrm{Re}\left(\overline{\alpha(x,y)}\alpha_{\ell_0}(x-y)\right)\right)\,\mathrm{d}(x,y).\end{aligned} \tag{4.26}$$

Before continuing, we introduce the following piece of notation. Let

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.27}$$

and define

$$\mathrm{Tr}_{\Omega,0}[A] = \mathrm{Tr}_\Omega[PAP] + \mathrm{Tr}_\Omega[(1-P)A(1-P)] \tag{4.28}$$

for $A \in \mathcal{L}\left(L^2(\mathbb{R}^2) \oplus L^2(\mathbb{R}^2)\right)$. Using this definition, the second term on the right hand side of Eq. (4.26) together with the kinetic energies can be written as

$$\begin{aligned}&\mathrm{Tr}_\Omega\left[\left(-\nabla^2 - \mu\right)(\gamma - \gamma_\ell)\right] \\ &\quad + 2\mathrm{Re}\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(\alpha_\ell(x-y)\overline{\alpha(x,y)} - |\alpha_\ell(x-y)|^2\right)\,\mathrm{d}(x,y) \\ &= \frac{1}{2}\,\mathrm{Tr}_{\Omega,0}\left[H_{\Delta_\ell}(\Gamma - \Gamma_\ell)\right],\end{aligned} \tag{4.29}$$

where H_{Δ_ℓ} is given as before, see Eq. (4.21). The new trace assures that the trace is taken with respect to a certain basis of the $\mathbb{C}^2$-matrix structure of the generalized one-particle density matrices. This is necessary because only the diagonal elements of $H_{\Delta_\ell}(\Gamma - \Gamma_\ell)$ are locally trace-class. At this point, it also turns out to be convenient to introduce the relative entropy $\mathcal{H}_0$, which for two states $\Gamma, \tilde{\Gamma} \in \mathcal{D}^{per}$ is defined by

$$\mathcal{H}_0(\Gamma, \tilde{\Gamma}) = \mathrm{Tr}_{\Omega,0}\left[\varphi(\Gamma) - \varphi(\tilde{\Gamma}) - \varphi'(\tilde{\Gamma})(\Gamma - \tilde{\Gamma})\right]. \tag{4.30}$$

The fact that $\varphi'(\Gamma_\ell) = -\beta H_{\Delta_\ell}$ yields the following statement:

Lemma 4.4 *Let $\Gamma_\ell \in \mathcal{D}_\ell$ with $\alpha_\ell = (\Gamma_\ell)_{12}$ be a minimizer of $\mathcal{F}_\ell$. Then*

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) = \frac{1}{2\beta}\mathcal{H}_0(\Gamma, \Gamma_\ell) + \int_{\Omega\times\mathbb{R}^2} V(x-y)|\alpha(x,y) - \alpha_\ell(x-y)|^2 \,\mathrm{d}(x,y) \quad (4.31)$$

for all $\Gamma \in \mathcal{D}^{per}$ with $\alpha = \Gamma_{12}$.

Based on this identity we estimate $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell)$ from below as stated in the following Proposition.

Proposition 4.1 *Let $\Gamma_\ell \in \mathcal{D}_\ell$ with $\gamma_\ell = (\Gamma_\ell)_{11}$ and $\alpha_\ell = (\Gamma_\ell)_{12}$ be a minimizer of $\mathcal{F}_\ell$ and denote $V_y(x) = V(x-y)$. Then*

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) \geq \int_\Omega \left(\alpha, \left(K_T^{\Delta_\ell} + V_y(x)\right)\alpha\right)_{L^2(\mathbb{R}^2, \mathrm{d}x)} \mathrm{d}y + \mathrm{Tr}_\Omega\left[K_T^{\Delta_\ell}(\gamma - \gamma_\ell)^2\right] \quad (4.32)$$

for all $\Gamma \in \mathcal{D}^{per}$ with $\gamma = \Gamma_{11}$ and $\alpha = \Gamma_{12}$. The first term on the right hand side of Eq. (4.32) has to interpreted in the following way: The operator $K_T^{\Delta_\ell} + V_y(x)$ acts on the x-coordinate of $\alpha(x,y)$. After the $L^2(\mathbb{R}^2)$-inner product in the x-coordinate is evaluated, one integrates the resulting expression in the y-coordinate over Ω.

Proof. The claimed estimate is a direct consequence of an inequality for the relative entropy that has been proven in [28, Lemma 5]. An application of this inequality yields

$$\begin{aligned}\mathcal{F}^{\mathrm{per}}(\Gamma) - \mathcal{F}^{\mathrm{per}}(\Gamma_\ell) \geq{}& \frac{1}{2}\mathrm{Tr}_\Omega\left[(\Gamma - \Gamma_\ell)\frac{H_{\Delta_\ell}}{\tanh\left(H_{\Delta_\ell}/(2T)\right)}(\Gamma - \Gamma_\ell)\right] \\ &+ \int_{\Omega\times\mathbb{R}^2} V(x-y)|\alpha(x,y) - \alpha_\ell(x-y)|^2\,\mathrm{d}(x,y)\end{aligned} \quad (4.33)$$

The facts that $x \mapsto x/(\tanh(x/2))$ is an even function and

$$H^2_{\Delta_\ell}(p) = \mathbb{1}_{\mathbb{C}^2}\left((p^2 - \mu)^2 + |\Delta_\ell(p)|^2\right) \quad (4.34)$$

is diagonal, imply the statement. □

Next, we show that the operator $K_T^{\Delta_{\ell_0}} + V$ is nonnegative for $T \in [T_c(\ell_1), T_c)$ and characterize its kernel.

Proposition 4.2 *Assume $V \in L^2(\mathbb{R}^2)$ with $\hat{V} \in L^r(\mathbb{R}^2)$ for some $r \in [1,2)$. Assume there exist $\ell_1, \ell_0 \in 2\mathbb{Z}$ such that*

$$T_c(\ell_0) > T_c(\ell_1) \geq T_c(\ell) \quad (4.35)$$

for all $\ell \in 2\mathbb{Z} \setminus \{\pm\ell_0\}$. Let Γ_{ℓ_0} with $(\Gamma_{\ell_0})_{12} = \alpha_{\ell_0}$ and $\Gamma_{-\ell_0}$ with $(\Gamma_{-\ell_0})_{12} = \alpha_{-\ell_0}$ be minimizers of $\mathcal{F}_{\ell_0}$ and $\mathcal{F}_{-\ell_0}$, respectively. Then $K_T^{\Delta_{\ell_0}} + V$ is nonnegative as an operator on $L^2(\mathbb{R}^2)$ for all $T \in [T_c(\ell_1), T_c)$ and $\alpha_{\ell_0}, \alpha_{-\ell_0} \in \ker\left(K_T^{\Delta_{\ell_0}} + V\right)$. If on the other hand $T \in (T_c(\ell_1), T_c)$, then α_{ℓ_0} and $\alpha_{-\ell_0}$ span the kernel of $K_T^{\Delta_{\ell_0}} + V$.

The proof of this proposition is based on spectral perturbation theory and relies on the fact that $K_T^{\Delta_{\ell_0}} + V \to K_{T_c} + V$ in norm resolvent sense for $T \to T_c$. We will derive this convergence from Lemma 4.5 and 4.6. For the sake of convenience, we write $a \lesssim b$ if there exists a constant $c > 0$ such that $a \leq cb$.

Lemma 4.5 *Let $T \in [0, T_c)$. The operators $A_T := K_T - K_{T_c}$ and $B_T := K_T^{\Delta} - K_T$ are bounded. More precisely, $\|A_T\| \lesssim (T_c - T)$ and $\|B_T\| \lesssim \|\Delta\|_{L^\infty(\mathbb{R}^2)}$. Moreover, $A_T \leq 0$ and $B_T \geq 0$.*

Proof. Note that

$$K_T^{\Delta}(p) = \frac{\sqrt{k(p)^2 + |\Delta(p)|^2}}{\tanh\left(\sqrt{k(p)^2 + |\Delta(p)|^2}/(2T)\right)} \tag{4.36}$$

is an increasing function in T for fixed Δ and vice versa. Hence $A_T \leq 0$ and $B_T \geq 0$. Both, A_T and B_T are pseudo-differential operators and by a slight abuse of notation we denote by $A_T(p)$ the symbol of A_T and by $B_T(p)$ the symbol of B_T. In the following, we abbreviate $T_c - T = \delta T$ and

$$I_T = \int_0^1 \frac{\delta T}{(T_c - s\delta T)^2} \, \mathrm{d}s. \tag{4.37}$$

A simple calculation yields

$$A_T(p) = -\int_0^1 \frac{I_T k(p)^2}{2\sinh^2\left(k(p)/(2T_c) + tI_T k(p)/2\right)} \, \mathrm{d}t. \tag{4.38}$$

For large $|p|$ the smooth function $p \mapsto A(p)$ and all its derivatives have exponential decay. Moreover, $|I_T| \lesssim T_c - T$ implies $\|A_T(p)\|_{L^\infty(\mathbb{R}^2)} \lesssim T_c - T$. In order to derive an analogous representation for $B_T(p)$ we define

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x} \frac{x}{\tanh(x/(2T))} = \frac{T\sinh(x/T) - x}{2T\sinh^2(x/(2T))} \tag{4.39}$$

as well as

$$\delta E(p) = \sqrt{k(p)^2 + |\Delta(p)|^2} - |k(p)| = \int_0^1 \frac{|\Delta(p)|^2}{2\sqrt{k(p)^2 + s|\Delta(p)|^2}} \, \mathrm{d}s. \tag{4.40}$$

A straightforward calculation shows that

$$B_T(p) = \delta E(p) \int_0^1 f(|k(p)| + t\delta E(p)) \, \mathrm{d}t. \tag{4.41}$$

Since the function f defined in Eq. (4.39) is bounded by 1, we find that $|B_T(p)| \leq |\delta E(p)|$ for almost all $p \in \mathbb{R}^2$. It can be seen directly from the definition of $\delta E(p)$, Eq. (4.40), that $|\delta E(p)| \lesssim |\Delta(p)|$ for almost all $p \in \mathbb{R}^2$, which implies $\|B_T\| \lesssim \|\Delta(p)\|_{L^\infty(\mathbb{R}^2)}$. □

Lemma 4.6 *Let $T \in [0, T_c)$. If α is a solution of the gap equation $(K_T^\Delta + V)\alpha = 0$ with $\Delta(p) = \frac{1}{\pi}\hat{V} * \hat{\alpha}(p)$, then $\|\alpha\|_{L^2(\mathbb{R}^2)} \lesssim (T_c - T)^{1/2}$. Additionally, $\|\Delta\|_{L^\infty(\mathbb{R}^2)} \lesssim (T_c - T)^{1/2}$.*

Proof. The gap equation, Eq. (4.20), can be written as

$$(\alpha, (K_{T_c} + V)\,\alpha) + (\alpha, B\alpha) = -\,(\alpha, A\alpha)\,, \tag{4.42}$$

where we use the notation introduced in Lemma 4.5 but drop the subscript, i.e. $A = A_T$ and $B = B_T$ for brevity. Lemma 4.5 and the definition of T_c imply that

$$(\alpha, B\alpha) \leq -\,(\alpha, A\alpha) \lesssim (T_c - T)\|\alpha\|^2_{L^2(\mathbb{R}^2)}. \tag{4.43}$$

We will show below that $\|(1 + (\cdot)^2)^{1/4}\hat{\alpha}\|^4_{L^4(\mathbb{R}^2)} \lesssim (\alpha, B\alpha)$ holds. This estimate implies the claim by the following arguments. From Eq. (4.43) we conclude that

$$\|\left(1 + (\cdot)^2\right)^{1/4}\hat{\alpha}\|^4_{L^4(\mathbb{R}^2)} \lesssim (T_c - T)\|\alpha\|^2_{L^2(\mathbb{R}^2)}. \tag{4.44}$$

On the other hand, the $L^r(\mathbb{R}^2)$-norm of $\hat{\alpha}$ is bounded from above by

$$\|\hat{\alpha}\|_{L^r(\mathbb{R}^2)} \leq \|\left(1 + (\cdot)^2\right)^{-1/4}\|_{L^s(\mathbb{R}^2)}\|\left(1 + (\cdot)^2\right)^{1/4}\hat{\alpha}\|_{L^4(\mathbb{R}^2)}, \tag{4.45}$$

where $r > 2$, due to the fact that we have to choose $s > 4$. Thus,

$$\|\hat{\alpha}\|^4_{L^r(\mathbb{R}^2)} \lesssim (T_c - T)\|\hat{\alpha}\|^2_{L^2(\mathbb{R}^2)}. \tag{4.46}$$

Furthermore, we conclude from the definition of Δ that

$$\|\Delta\|_{L^\infty(\mathbb{R}^2)} \lesssim \|\hat{V}\|_{L^t(\mathbb{R}^2)}\|\hat{\alpha}\|_{L^r(\mathbb{R}^2)}, \tag{4.47}$$

where we choose $r = 2$ and $t \in [1, 2)$ appropriately. A combination of the gap equation in the form Eq. (4.20) together with Eq. (4.46) and Eq. (4.47) finally shows

$$\|\hat{\alpha}\|_{L^2(\mathbb{R}^2)} \lesssim \|\Delta\|_{L^\infty(\mathbb{R}^2)} \lesssim (T_c - T)^{1/4}\|\hat{\alpha}\|^{1/2}_{L^2(\mathbb{R}^2)}, \tag{4.48}$$

which is what we intended to show.

It is left to prove that $\|(1 + (\cdot)^2)^{1/4}\hat{\alpha}\|^4_{L^4(\mathbb{R}^2)} \lesssim (\alpha, B\alpha)$ holds. We recall that the symbol of the operator B can be written as

$$B(p) = \delta E(p)\int_0^1 f(|k(p)| + t\delta E(p))\,\mathrm{d}t, \tag{4.49}$$

where f is the function defined in Eq. (4.39). Since f is strictly increasing, $B(p)$ is bounded from below by

$$B(p) \geq \int_{1/2}^1 f(|k(p)| + t\delta E(p))\,\mathrm{d}t \geq \frac{\delta E(p)}{2} f(|k(p)| + \delta E(p)/2). \tag{4.50}$$

Now, we choose a constant $c > 0$ and distinguish two cases. First, consider $p \in \mathbb{R}^2$ such that $|k(p)| + \delta E(p)/2 \geq c$ holds. Making use of the fact that we only have to consider $\Gamma \in \mathcal{D}$ such that $\mathcal{F}(\Gamma) \leq \mathcal{F}(\Gamma_n)$ holds, where Γ_n denotes the normal state, one can easily show that $\|\Delta\|_{L^\infty(\mathbb{R}^2)}$ is uniformly bounded. Together with the relation $\Delta(p) = -2K_T^\Delta(p)\hat{\alpha}(p)$, this implies

$$\begin{aligned} B(p) &\geq \frac{f(c)}{2}\delta E(p) \geq \frac{f(c)}{2}\frac{|\Delta(p)|^2}{2|k(p)| + |\Delta(p)|} \\ &\geq K_T^\Delta(p)|\hat{\alpha}(p)|^2 \inf_{p\in\mathbb{R}^2} \frac{K_T^\Delta(p)}{2|k(p)| + \|\Delta\|_{L^\infty(\mathbb{R}^2)}} \\ &\gtrsim \left(1+p^2\right)|\hat{\alpha}(p)|^2. \end{aligned} \tag{4.51}$$

For the second case, we consider all $p \in \mathbb{R}^2$ such that $|k(p)| + \delta E(p)/2 < c$. Since f is a concave function and $f(0) = 0$, we conclude that $f(x) \geq xf(c)/c$ as long as $x \leq c$. Consequently, we see that

$$\begin{aligned} B(p) &\geq \frac{f(c)}{2c}\,\delta E(p)\left(|k(p)| + \frac{\delta E(p)}{2}\right) \\ &\geq \frac{f(c)}{2c}\,|\Delta(p)|^2\left(\frac{1}{2} - \frac{|\Delta(p)|\,|k(p)|}{(2|k(p)| + |\Delta(p)|)^2}\right) \\ &\geq \frac{3f(c)}{16c}\,|\Delta(p)|^2. \end{aligned} \tag{4.52}$$

Combining Eq. (4.51), Eq. (4.52) and Eq. (4.20), we arrive at

$$B(p) \gtrsim \left(1+p^2\right)|\hat{\alpha}(p)|^2. \tag{4.53}$$

We insert this estimate in Eq. (4.49), which yields

$$(\alpha, B\alpha) \gtrsim \int_{\mathbb{R}^2} \left(1+p^2\right)|\hat{\alpha}(p)|^4\,\mathrm{d}p = \|\left(1+(\cdot)^2\right)\hat{\alpha}\|^4_{L^4(\mathbb{R}^2)} \tag{4.54}$$

and hence concludes the proof. □

Let $T \in [0, T_c)$ and $z \in \mathbb{C}\setminus\mathbb{R}$. Lemma 4.5 and Lemma 4.6 together show that

$$\begin{aligned} &\left\|\frac{1}{z-(K_{T_c}+V)} - \frac{1}{z-(K_T^\Delta+V)}\right\| \\ &\leq \left\|\frac{1}{z-(K_{T_c}+V)}\right\|\|A_T + B_T\|\left\|\frac{1}{z-(K_T^\Delta+V)}\right\| \\ &\lesssim |\mathrm{Im}(z)|^{-2}\sqrt{T_c - T}. \end{aligned} \tag{4.55}$$

In other words, $K_T^\Delta + V \to K_{T_c} + V$ for $T \to T_c$ in norm resolvent sense for an arbitrary $z \in \mathbb{C}\setminus\mathbb{R}$ and consequently for all $z \in \rho(K_{T_c}+V)$. We are now prepared for the proof of Proposition 4.2.

Proof of Proposition 4.2. Let us consider the case $\ell_0 \neq 0$. The proof for the case $\ell_0 = 0$ is analogous. The assumption $T_c(\ell_0) > T_c(\ell_1) \geq T_c(\ell)$ for all $\ell \in 2\mathbb{Z} \setminus \{\pm\ell_0\}$ together with the fact that

$$\inf_{\alpha \in \mathcal{H}_{\ell_0}} (\alpha, (K_{T_c} + V)\alpha) = \inf_{\alpha \in \mathcal{H}_{-\ell_0}} (\alpha, (K_{T_c} + V)\alpha) \tag{4.56}$$

ensure that the lowest eigenvalue of $K_{T_c} + V$ is twice degenerate. In the case $\ell_0 = 0$ this eigenvalue is non-degenerate. From the convergence of $K_T^{\Delta_{\ell_0}} + V$ to $K_{T_c} + V$ in norm resolvent sense, which follows from the considerations above, one concludes that the lowest eigenvalue of $K_T^{\Delta_{\ell_0}} + V$ is stable.

In particular, we know that for some $\epsilon > 0$ the operator $K_T^{\Delta_{\ell_0}} + V$ has exactly two eigenvalues $\lambda_1(T), \lambda_2(T) \in \{z \in \mathbb{C} \mid |z| < \epsilon\}$ for all $\tilde{T} < T < T_c$ and $T_c - \tilde{T}$ small enough. The Euler-Lagrange equations of $\mathcal{F}_{\ell_0}$ and $\mathcal{F}_{-\ell_0}$ can be written as

$$\begin{aligned} (K_T^{\Delta_{\ell_0}} + V)\alpha_{\ell_0} &= 0, \\ (K_T^{\Delta_{\ell_0}} + V)\alpha_{-\ell_0} &= 0, \end{aligned} \tag{4.57}$$

respectively, which tells us that $\lambda_1(T) = 0 = \lambda_2(T)$. Thus, we have shown that there exists a $\tilde{T} < T_c$ such that $K_T^{\Delta_{\ell_0}} + V$ is nonnegative if $T \in (\tilde{T}, T_c)$.

In order to prove the full statement, it remains to show that this holds true for all $T \in [T_c(\ell_1), T_c)$. The above argument can be repeatedly applied, as long as the gap between the ground state of $K_T^{\Delta_{\ell_0}} + V$ and its smallest eigenvalue $\lambda_3(T) > 0$ does not close. This is the case if $T > T_c(\ell_1)$, as the following argument shows. Assume $\varphi \perp \mathcal{D}_{\ell_0}$ and estimate

$$\begin{aligned} \left(\varphi, (K_T^{\Delta_{\ell_0}} + V)\varphi\right) &\geq (\varphi, (K_T + V)\varphi) \\ &= \left(\varphi, (K_{T_c(\ell_1)} + V)\varphi\right) + \left(\varphi, \left(K_T - K_{T_c(\ell_1)}\right)\varphi\right). \end{aligned} \tag{4.58}$$

Let us distinguish two cases. If, in the first case, $\varphi \notin \ker(K_{T_c(\ell_1)} + V)$, then

$$\left(\varphi, (K_{T_c(\ell_1)} + V)\varphi\right) \geq \kappa_1 \|\varphi\|^2_{L^2(\mathbb{R}^2)}, \tag{4.59}$$

for some $\kappa_1 > 0$. If, on the other hand, $\varphi \in \ker(K_{T_c(\ell_1)} + V)$, there exists a radius $R > 0$ such that $\int_{B_R(0)} |\hat{\varphi}(p)|^2 \,\mathrm{d}p = \frac{1}{2}\|\varphi\|^2_{L^2(\mathbb{R}^2)}$. From the monotonicity of K_T in T, we deduce that $K_T(p) - K_{T_c(\ell_1)}(p)$ is a strictly positive function for all $T \in (T_c(\ell_1), T_c)$. Hence,

$$\begin{aligned} \left(\varphi, \left(K_T - K_{T_c(\ell_1)}\right)\varphi\right) &\geq \int_{B_R(0)} |\hat{\varphi}(p)|^2 \left(K_T(p) - K_{T_c(\ell_1)}(p)\right) \mathrm{d}p \\ &\geq \frac{\kappa_2(T)}{2} \|\varphi\|^2_{L^2(\mathbb{R}^2)}, \end{aligned} \tag{4.60}$$

for some positive constant $\kappa_2(T)$ depending on T. Since the kernel of $K_{T_c(\ell_1)} + V$ is finite-dimensional this is sufficient. Together with the first case, we have shown that for all $T \in (T_c(\ell_1), T_c)$ there exists a constant $\kappa(T) > 0$ such that

$$\left(\varphi, (K_T^{\Delta_{\ell_0}} + V)\varphi\right) \geq \kappa(T)\|\varphi\|^2_{L^2(\mathbb{R}^2)} \tag{4.61}$$

for all $\varphi \perp \mathcal{D}_{\ell_0}$. Given any compact interval $I \subset (T_c(\ell_1), T_c)$ we choose $\kappa = \min\{\kappa(T) \mid T \in I\}$ and we know that $\lambda_3(T) \geq \kappa > 0$ for all $T \in I$. This allows us to apply the perturbation argument repeatedly and we deduce that $K_T^{\Delta_{\ell_0}} + V$ is nonnegative on $(T_c(\ell_1), T_c)$. The continuity of the eigenvalues of $K_T^{\Delta_{\ell_0}} + V$ in T yields the full statement. □

Proof of Theorem 4.1. To summarize, we know from Lemma 4.1 that for ℓ_0 determined by $T_c(\ell_0) = \max_{\ell \in \mathbb{Z}} T_c(\ell)$, the functional $\mathcal{F}_{\ell_0}$ has a minimizer $\Gamma_{\ell_0} \in \mathcal{D}_{\ell_0}$. Proposition 4.1 and Proposition 4.2 show that

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) \geq 0 \tag{4.62}$$

holds for all $\Gamma \in \mathcal{D}^{per}$. Moreover, if $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) = 0$, then $\gamma = \gamma_{\ell_0}$ and $\alpha = \psi_1\alpha_{\ell_0} + \psi_2\alpha_{-\ell_0}$ for $\psi_1, \psi_2 \in \mathbb{C}$ by Proposition 4.2. It remains to show that either $|\psi_1| = 1$ and $\psi_2 = 0$ or $\psi_1 = 0$ and $|\psi_2| = 1$ holds. To do so, we realize that Eq. (4.62) also tells us that it is sufficient to look for minimizers of $\mathcal{F}^{per}$ in the set $\mathcal{D}$, on which $\mathcal{F}^{per}$ reduces to $\mathcal{F}$. But on this set there exists at least one minimizer and this minimizer solves the corresponding Euler-Lagrange equation. From Eq. (4.20) we conclude that $\Delta(p) = \frac{1}{\pi}\hat{V} * \hat{\alpha}(p)$ is a radial function. This together with Eq. (4.20) implies that $\hat{\alpha}$ is a radial function and lets us conclude that either $\psi_1 = 0$ or $\psi_2 = 0$ holds. Hence, we can restrict attention to the set $\mathcal{D}_{\ell_0} \cup \mathcal{D}_{-\ell_0}$, on which two minimizers exist, namely $(\gamma_{\ell_0}, \alpha_{\ell_0})$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0})$. In other words, we have found two minimizers of $\mathcal{F}^{per}$ and for both, the remaining ψ has absolute value one. If there exists another minimizer, it has to fulfil the Euler-Lagrange equation of $\mathcal{F}$. Eq. (4.20) together with the fact that $\left(\alpha_{\ell_0}, K_T^{\psi\Delta_{\ell_0}}\alpha_{\ell_0}\right)$ and $\left(\alpha_{-\ell_0}, K_T^{\psi\Delta_{\ell_0}}\alpha_{-\ell_0}\right)$ are strictly monotone in ψ shows that the constant in front of α_{ℓ_0} or $\alpha_{-\ell_0}$ for this minimizer must have absolute value one. Therefore, except for a phase, it cannot be distinct from the minimizers we have already found. This concludes the proof of Theorem 4.1. □

The proof of Theorem 4.2 is analogous to the proof of Theorem 4.1 with one exception.

Proof of Theorem 4.2. In case of $\ell_0 = 0$, all arguments given in the proof of Theorem 4.1 apply, except for Lemma 4.6, where we need to modify the assumptions on V slightly. One easily checks that $\hat{V} \in L^r(\mathbb{R}^3)$ with $r \in [1, 12/7)$ is a sufficient assumption in this case. □

Bibliography

[1] J. Bardeen, L. Cooper, J. Schrieffer, *Theory of superconductivity*, Phys. Rev. **108**, 1175-1204 (1957).

[2] L. Cooper, *Bound electron pairs in a degenerate Fermi gas*, Phys. Rev. **104**, 1189–1190 (1956).

[3] A. L. Fetter, J. D. Walecka, *Quantum theory of many-particle systems*, (Dover, New York, 2003).

[4] A. J. Leggett, *Diatomic molecules and Cooper pairs*, Modern trends in the theory of condensed matter, J. Phys. (Paris) Colloq. C7-19 (1980).

[5] L. D. Landau, E. M. Lifschitz, *Lehrbuch der Theoretischen Physik, Band 8*, (Akademie Verlag, Berlin, 1992).

[6] N. W. Ashcroft, N. D. Mermin, *Solid state physics*, (Holt, Rinehart and Winston, New York, 1976).

[7] R. D. Parks (Editor), *Superconductivity, Volume 1 and 2*, (Marcel Dekker, New York, 1969).

[8] V. Bach, E. Lieb, J. Solovej, *Generalized Hartree-Fock theory and the Hubbard model*, J. Stat. Phys. **76**, 3-89 (1994).

[9] F. Odeh, *An existence theorem for the BCS integral equation*, IBM J. Res. Develop. **8**, 187-188 (1964).

[10] P. Billard, G. Fano, *An existence proof for the gap equation in the superconductivity theory*, Commun. Math. Phys. **10**, 274-279 (1968).

[11] A. Vansevenant, *The gap equation in superconductivity theory*, Physica **17D**, 339-344 (1985).

[12] Y. Yang, *On the Bardeen-Cooper-Schrieffer integral equation in the theory of superconductivity*, Lett. Math. Phys. **22**, 27-37 (1991).

[13] J. B. McLeod, Y. Yang, *The uniqueness and approximation of a positive solution of the Bardeen-Cooper-Schrieffer gap equation*, J. Math. Phys. **41**, 6007 - 6025 (2000).

[14] Y. Yang, *Mathematical analysis of the multiband BCS gap equations in superconductivity*, Physica **200D**, 60-74 (2005).

[15] A. J. Leggett, *Quantum liquids, Bose condensation and Cooper pairing in condensed-matter systems*, (Oxford University Press, New York, 2006).

[16] C. Hainzl, E. Hamza, R. Seiringer, J. P. Solovej, *The BCS-functional for general pair interactions*, Math. Phys. **281**, 349-367 (2008).

[17] R. L. Frank, C. Hainzl, S. Naboko, R. Seiringer, *The critical temperature for the BCS equation at weak coupling*, J. Geom. Anal. **17**, 559-567 (2007).

[18] C. Hainzl, R. Seiringer, *Critical temperature and energy gap for the BCS equation*, Phys. Rev. B **77**, 184517 (2008).

[19] C. Hainzl, R. Seiringer, *Spectral properties of the BCS gap equation of superfluidity*, Mathematical results in quantum mechanics, World Sci. Publ., Hackensack, NJ (2008), 117–136.

[20] C. Hainzl, R. Seiringer, *The BCS critical temperature for potentials with negative scattering length*, Lett Math Phys **84**, 99-107 (2008).

[21] G. Bräunlich, C. Hainzl, R. Seiringer, *Translation invariant quasi-free states for fermionic systems and the BCS approximation*, Rev. Math. Phys. **26**, 1450012 (2014).

[22] C. Hainzl, R. Seiringer, *Low density limit of BCS theory and Bose-Einstein condensation of fermion pairs*, Lett. Math. Phys. **100**, 119-138 (2011).

[23] G. Bräunlich, C. Hainzl, R. Seiringer, *Bogolubov-Hartree-Fock theory for strongly interacting fermions in the low density limit*, arXiv:1511.08047 [math-ph] (2015).

[24] C. Hainzl, B. Schlein, *Dynamics of Bose-Einstein condensates of fermion pairs in the low density limit of BCS theory*, J. Funct. Anal. **265**, 399-423 (2013).

[25] A. Freiji, C. Hainzl, R. Seiringer, *The BCS gap equation for spin-polarized fermions*, J. Math. Phys. **53**, 012101 (2012).

[26] A. A. Abrikosov, *The magnetic properties of superconducting alloys*, J. Phys. Chem. Solids. Pergamon Press **2**, 199-208 (1957).

[27] L. P. Gor'kov, *Microscopic derivation of the Ginzburg-Landau equations in the theory of superconductivity*, Zh. Eksp. Teor. Fiz. **36**, 1918-1923 (1959); *English translation* Soviet Phys. JETP **9**, 1364-1367 (1959).

[28] R. L. Frank, C. Hainzl, R. Seiringer, J. P. Solovej, *Microscopic derivation of Ginzburg-Landau theory*, J. Amer. Math. Soc. **25**, 667-713 (2012).

[29] R. L. Frank, C. Hainzl, R. Seiringer, J. P. Solovej, *The external field dependence of the BCS critical temperature*, Commun. Math. Phys. 342, 189 (2016).

[30] R. L. Frank, C. Hainzl, R. Seiringer, J. P. Solovej, *Microscopic derivation of the Ginzburg-Landau model*, XVIIth International Congress on Mathematical Physics, A. Jensen (ed.), pp. 575-583, World Scientific (2013).

[31] C. Hainzl, R. Seiringer, *The BCS functional of superconductivity and its mathematical properties*, arXiv:1511.01995v1 [math-ph] (2015).

[32] R. L. Frank, M. Lemm, *Ginzburg-Landau theory with multiple order parameters: Microscopic derivation and examples*, Ann. Henri Poincaré, Online first (2016).

[33] R. L. Frank, C. Hainzl, B. Schlein, R. Seiringer, *Incompatibility of time-dependent Bogoliubov-De-Gennes and Ginzburg-Landau equations*, arXiv:1504.05885v2 [math-ph] (2015).

[34] C. Hainzl, J. Seyrich, *Comparing the full time-dependent Bogoliubov-de-Gennes equations to their linear approximation: A numerical investigation*, arXiv:1504.05881v2 [math-ph] (2015).

[35] W. Thirring, *Quantum mathematical physics*, (2^{nd} ed., Springer, New York, 2002).

[36] M. Lewin, J. Sabin, *A family of monotone quantum relative entropies*, Lett. Math. Phys. **104**, 691-705 (2014).

[37] A. Deuchert, C. Hainzl, R. Seiringer, *Note on a family of monotone quantum relative entropies*, Lett. Math. Phys. 105, 1449 (2015).

[38] R. Bhatia, *Matrix analysis*, (Springer-Verlag, New York, 1997).

[39] E. Carlen, *Trace inequalities and quantum entropy: an introductory course. Entropy and the quantum*, 73–140, Contemp. Math., 529, Amer. Math. Soc., Providence, RI, 2010.

[40] C. Hainzl, M. Lewin, R. Seiringer, *A Nonlinear Model for Relativistic Electrons at Positive Temperature*, Rev. Math. Phys. **20**, 1283-1307 (2008).

[41] M. Reed, B. Simon, *Methods of Modern Mathematical Physics. I. Functional Analysis*, (Academic Press, San Diego, 1980).

[42] M. Reed, B. Simon, *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*, (Academic Press, San Diego, 1975).

[43] V. L. Ginzburg, L. D. Landau, *On the theory of superconductivity*, Zh. Eksp. Teor. Fiz. **20**, 1064-1082 (1950).

[44] S. Gustafson, I. M. Sigal, T. Tzaneteas, *Statics and dynamics of magnetic vertices*

and of Nielson-Olesen (Nambu) strings, J. Math. Phys. **51**, 015217 (2010).

[45] A. Jaffe, C. Taubes, *Vortices and monopoles: Structure of static gauge theories*, Progress in Physics 2 (Birkhäuser, Bosten 1980).

[46] F. Bethuel, H. Brezis, F. Hélein, *Ginzburg-Landau vortices*, (Birkhäuser, MA, 1994).

[47] S. Fournais, B. Helffer, *Spectral methods in surface superconductivity*, (Birkhäuser, MA, 2010), Vol. 77.

[48] E. Sandier, S. Serfaty, *Vortices in the magnetic Ginzburg-Landau model*, (Birkhäuser, MA, 2007), Vol. 70.

[49] M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions with formulas, graphs and mathematical tables* (Dover, New York, 1964).

[50] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey (Editor), *Table of integrals, series and products*, (Academic Press, London, 1994).

[51] E. M. Stein, G. Weiss, *Fourier analysis on Euclidean spaces*, (Princeton University Press, Princeton, New Jersey, 1971).

[52] B. Simon, *Trace ideals and their application* (American Mathematical Society, Providence, Rhode Island, 2005).

[53] E. Hille, *Functional analysis and semi-groups*, (American Mathematical Society, New York, 1948).