Effective Models for Many Particle Systems: BCS Theory and the Kac Model

Alissa Geisinger

Mathematisches Institut
Universität Tübingen

2018

Effective Models for Many Particle Systems: BCS Theory and the Kac Model

DISSERTATION

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines Doktors der
Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von

Alissa Geisinger

aus Stuttgart

Tübingen
2018

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

Tag der mündlichen Qualifikation: 04.05.2018

Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Christian Hainzl
2. Berichterstatter:	Prof. Dr. Michael Loss

Es ist nicht das Wissen, sondern das Lernen, nicht das Besitzen, sondern das Erwerben, nicht das Dasein, sondern das Hinkommen, was den größten Genuss gewährt.

Carl Friedrich Gauß, 1777-1855.

Contents

CHAPTER 1

Preface

Although this may seem a paradox, all exact science is dominated by the idea of approximation. When a man tells you that he knows the exact truth about anything, you are safe in infering that he is an inexact man. Every careful measurement in science is always given with the probable error ... every observer admits that he is likely wrong, and knows about how much wrong he is likely to be.

Bertrand Russell, 1872 - 1970,
The Scientific Outlook, 1931.

In the beginning of the 20th century, the dependence of electrical resistance on temperature became a very active field of research in physics. Measurements on the resistance of many different metals at low temperatures, which were carried out by experimental physicists, amongst them James Dewar and John Ambrose Fleming, gave room for the idea that the resistance of metals could vanish completely at very low temperatures. Lord Kelvin [**63**] in 1902 made the prediction that resistance decreases with falling temperature up to a certain point. After this minimum is reached, he expected the resistance to increase again as temperature decreases further. This behavior turned out to be true for semiconductors. However, in the case of superconductivity things are different. At a certain temperature, the so-called critical temperature, resistance abruptly vanishes and stays zero at all temperatures below the critical temperature. This phenomenon was discovered by Heike Kamerlingh Onnes in 1911 and he was awarded the Nobel Prize for his discovery of superconducting materials in 1913.

It took almost 46 years of intensive research until, in 1957, John Bardeen, Leon Neil Cooper and John Robert Schrieffer published their famous paper with the title "Theory of Superconductivity" [**6**]. Their theory of superconductivity, usually referred to as BCS theory, is based on the idea of electron pairing driven by an effective attraction mediated by phonons. This publication was a breakthrough, presenting the first microscopic model for the remarkable effect of vanishing resistance and the three authors were awarded the Nobel prize for their discovery in 1972. Before, Cooper [**24**] had realized that a very tiny attractive interaction between particles in a Fermi gas suffices to cause pairing between electrons. While this interaction is possible because of interactions through the lattice in the case of metals, it is of local type in other situations, as for example for superfluid cold gases. The arising electron pairs are known as Cooper pairs. Approximately, Cooper pairs behave like Bosons and they form a condensed state which is not identical with a Bose-Einstein condensate, but nonetheless similar. Amongst the famous and typical properties of superconductors that can be explained with the appearance of this strongly correlated quantum state are, for example, infinite conductivity, the Meissner effect, flux quantization and the isotope effect, see the original work of Bardeen, Cooper and Schrieffer [**6**] and the book of Fetter and Walecka [**31**].

The second model, which will play a role in this thesis, is a model in kinetic theory and was introduced by Mark Kac in 1956 in his article "Foundations of kinetic theory" [**61**]. The Kac model is a linear, microscopic model to describe a gas of interacting particles in a probabilistic way. This model, and in particular the Kac master equation, due to its simplicity, have a special place among the models describing a large number of interacting particles. One goal, and the main motivation for Kac's work in [**61**], was to provide a satisfactory derivation of the spatially homogeneous Boltzmann equation. Indeed, Kac in [**61**] was able to derive the spatially homogeneous, non-linear Kac-Boltzmann equation. It was in this context that Kac introduced his notion of propagation of chaos as well as his definition of chaotic sequences, what he called 'sequences that have the Boltzmann property'. Both these concepts turned out to be very useful tools in his derivation. So far, the - much more difficult - derivation of the Boltzmann equation from the laws of classical mechanics has only been shown for situations with very few collisions, see the work of O. Lanford [**66, 67**]. In addition, Kac in [**61**] wished to lay a basis and give a mathematical setting for the study of approach to equilibrium, which he did by presenting his approach to the problem through master equations.

These two models, the BCS model and the Kac model, are the effective theories, which will be studied in this thesis.

Acknowledgements

First I would like to express my gratitude to my advisor Christian Hainzl for introducing me to exciting topics in mathematical physics and BCS theory and for giving me the opportunity to write this thesis. I am very thankful that I had the opportunity to profit from his knowledge, for his constant support, his advice and his help in various mathematical matters.

I am indebted to my teacher and co-author Michael Loss for sharing beautiful mathematical ideas. It has been a great pleasure to learn from him and to work with him. I am very grateful for the invitation to Georgia Institute of Technology and I would like to thank Michael for his kind hospitality and his support. I am grateful for his willingness to act as a referee in my thesis committee.

Special thanks go to my co-authors and collaborators Federico Bonetto, Andreas Deuchert, Rupert Frank, Tobias Ried and Tim Tzaneteas for interesting discussions and the collaboration that lead to new results and to Marcello Porta and Stefan Teufel for being members of my thesis committee.

Financial support by the Deutsche Forschungsgemeinschaft through Graduiertenkolleg 1838 is gratefully acknowledged. In particular, I thank the Graduiertenkolleg 1838 for the support of my research stay at the Georgia Institute of Technology, Atlanta.

It is a pleasure to thank all my friends and colleagues, and in particular the Mathematical Physics Group at the University in Tübingen. I am thankful for interesting and fruitful mathematical discussions, their help in LaTeX and the coffee breaks.

Last but not least, I would like to thank my family for their support and my husband for his constant encouragement.

CHAPTER 2

Summary

Von vielen Dingen spreche ich gar nicht, weil sie sich von selbst verstehen. Kluge Leute erraten das meiste.

Lea Mendelssohn, 1777 - 1842.

2.1. German summary

In dieser Arbeit werden Aspekte der BCS-Theorie und des Kac-Modells diskutiert und neue Ergebnisse zu diesen Forschungsgebieten vorgestellt. Im Rahmen der BCS-Theorie untersuchen wir die Frage nach Brechung oder Erhaltung der Translationssymmetrie. Ein weiteres Kapitel ist dem Zusammenhang von BCS-Theorie und dem makroskopischen Ginzburg-Landau-Modell gewidmet. Im Kontext des Kac-Modells in der kinetischen Theorie studieren wir die relative Entropie und stellen ein Ergebnis zum exponentiellem Abfall von Lösungen der Kac-Master-Gleichung vor.

Die BCS-Theorie wurde von ihren Namensgebern Bardeen, Cooper und Schrieffer, mit dem Artikel "Theory of Superconductivity" aus dem Jahre 1957 begründet. Die Autoren veröffentlichten dort erstmals ihre Idee eines mikroskopischen Modells der Supraleitung. Die Annahme von Translationssymmetrie stellt eine signifikante Vereinfachung des BCS-Modells dar. Deshalb ist es interessant, Situationen zu charakterisieren in denen die Translationssymmetrie nicht gebrochen ist und die Annahme der Translationssymmetrie damit gerechtfertigt ist. In Kapitel 5 betrachten wir diese Fragestellung im Fall von radialer Wechselwirkung in zwei Dimensionen. Für Temperaturen, die in einem bestimmten Intervall unterhalb der kritischen Temperatur liegen, zeigen wir, dass die Minimierer des BCS-Funktionals translationsinvariant sind. Das heißt, dass die Translationssymmetrie des Systems nicht gebrochen ist. Das Ergebnis lässt sich auf den Fall von drei Dimensionen übertragen, falls die Cooperpaarwellenfunktionen verschwindenden Drehimpuls haben. Des Weiteren lässt das Resultat den Schluss zu, auf die Eindeutigkeit des Minimierers des BCS-Funktionals zu schließen, bis auf komplexe Phasen der Cooperpaarwellenfunktion.

In einem weiteren Kapitel zur BCS-Theorie, Kapitel 6, geht es um die Verbindung zur Ginzburg-Landau-Theorie, siehe [**46**]. Diese phänomenologische Theorie wurde vor der bahnbrechenden Entdeckung des BCS-Trios herangezogen um einige Aspekte der Supraleitung zu erklären. Schon kurz nach der Veröffentlichung der Ideen von Bardeen, Cooper und Schrieffer, stellte Gorkov einen Zusammenhang zwischen den beiden Modellen fest. Das erste rigorose Ergebnis zu dieser Fragestellung erschien 2012, siehe [**37**]. An dieses wegweisende Resultat ist die Arbeit, die wir hier präsentieren angelehnt. Statt das magnetische Feld als Parameter zu betrachten, fassen wir es hier jedoch als zusätzliche Variable auf. Dies erlaubt uns auch die zweite Ginzburg-Landau-Gleichung aus der BCS-Theorie herzuleiten. Die technischen Schwierigkeiten, die hierdurch entstehen gehen wir mit einer rigorosen Approximation der Phase an, eine Methode, die erst in [**34**] vorgestellt wurde.

In Kapitel 7 stellen wir ein Ergebnis vor, bei dem es um die relative Entropie im Kac-Modell geht. Im Kac-Modell werden die Teilchen nur durch ihre jeweilige Geschwindigkeit beschrieben. Im thermischen Gleichgewichtszustand sind diese Geschwindigkeiten normalverteilt. Wir betrachten ein endliches Reservoir von Teilchen, das zu Beginn im Gleichgewichtszustand ist, und ein daran gekoppeltes System von weniger Teilchen. An die Anfangsgeschwindigkeitsverteilung der Teilchen im System machen wir nur sehr schwache Annahmen. So kann man das System auch als Störung eines großen Reservoirs verstehen. Diese Anfangsbedingungen sind entscheidend für unsere Betrachtung und das Ergebnis. Nach von Kac vorgegebenen Regeln werden nun die Kollisionen der Teilchen beschrieben. Die Zeitevolution des Systems ist bestimmt durch die Kac-Master-Gleichung. Wir zeigen, dass Lösungen der Kac-Master-Gleichung in relativer Entropie zum thermischen Gleichgewichtszustand exponentiell abfallen. Exponentiell meint hier, exponentiell in der Zeit. Die Rate wird explizit angegeben und ist unabhängig von der Teilchenzahl. Dieses Resultat stellt somit einen Gegenpol zu früheren Ergebnissen dar, die besagen, dass für beliebige Anfangsbedingung die Entropieproduktion invers proportional zur Teilchenzahl ist. Der Beweis, den diese Arbeit enthält, beruht auf einer Methoder der Abschätzung von Nelson, bekannt als Nelson's hypercontractive estimate, sowie einer geometrischen Form der Ungleichungen von Brascamp-Lieb, die auf Barthe zurückgeht. Das Resultat lässt sich in einem Spezialfall auf ein System mit dreidimensionalen Boltzmann - Kac - Kollisionen und Pseudo-Maxwellschen Molekülen übertragen.

2.2. List of publications

Included in this thesis are the following publications.

- *Persistence of translational symmetry in the BCS model with radial pair interaction*,
  A. Deuchert, A. Geisinger, C. Hainzl, M. Loss,
  accepted for publication in Ann. Henri Poincaré, *preprint:* arXiv:1612.03303.
- *Entropy decay for the Kac evolution*,
  F. Bonetto, A. Geisinger, M. Loss, T. Ried,
  submitted to Comm. Math. Phys., *preprint:* arXiv:1707.09584.

2.2.1. Personal Contribution. The results of Chapter 5 were obtained under the supervision of Prof. Christian Hainzl and Prof. Michael Loss. The author contributed mathematical ideas and was responsible for the proofs as well as for the manuscript. These responsibilities were shared by Andreas Deuchert and the author in equal parts.

The results of Chapter 6 were obtained under the supervision of Prof. Rupert Frank and Prof. Christian Hainzl. The author contributed mathematical ideas and was responsible for the proofs as well as for the manuscript. These responsibilities were shared by Tim Tzaneteas and the author in equal parts.

The results of Chapter 7 were obtained under the supervision of Prof. Federico Bonetto and Prof. Michael Loss. The author contributed mathematical ideas and was responsible for the proofs as well as for the manuscript. These responsibilities were shared by Tobias Ried and the author in equal parts.

CHAPTER 3

Mathematical aspects of the BCS model

Superconductivity is a peculiar phenomenon occurring in many metallic materials. Metals in their normal state have a certain electrical resistance, the magnitude of which varies with temperature. When a metal is cooled its resistance is reduced. In many metallic materials it happens that the electrical resistance not only decreases but also suddenly disappears when a certain critical temperature is passed which is a characteristic property of the material.

Stig Lundqvist
Nobel Lectures Physics 1971-1980,
World Scientific Publishing Co., Singapore, 1992.

3.1. Introduction

3.1.1. Superconductivity and BCS Theory. In the first mathematical works on BCS theory the authors considered the so-called *BCS gap equation*, a nonlinear integral equation for the gap function Δ, as the starting point. In particular, at that point, no benefit was derived from the fact that the BCS gap equation can be realized as the Euler-Lagrange equation of a functional, the so-called *BCS functional.* The connection between the BCS gap equation and the BCS functional was revealed by Leggett [**68**]. In d dimensions, at temperature $T \geq 0$ and with interaction potential V, the BCS gap equation reads,

$$\Delta(p) = -\int_{\mathbb{R}^d} V(p,q) \frac{\tanh\left(E(q)/(2T)\right)}{E(q)} \Delta(q)\,\mathrm{d}q, \tag{3.1}$$

where

$$E(q) = \left(\left(q^2 - \mu\right)^2 + |\Delta(q)|^2\right)^{1/2}$$

and $\mu \in \mathbb{R}$ is the chemical potential. The described system is said to be in a superconducting state if the BCS gap equation allows for a non-trivial solution, that is a solution $\Delta \neq 0$. In situations where the trivial solution is the only solution to (3.1), the system is said to be in the normal state. Whether one finds oneself in the first or in the latter case highly depends on the parameters V and T. Indeed, it turns out that under rather general assumptions one can prove the existence of a critical temperature $T_c \geq 0$, which is characterized by the property that the BCS gap equation only has a non-trivial solution at temperatures below T_c. Consequently, at temperatures at or above the critical temperature, i.e., for $T \geq T_c$, the only solution of the BCS gap equation is the trivial solution $\Delta \equiv 0$.

Among the most important of the first mathematical works on the BCS gap equation one has to mention the work of Odeh [**77**], who presented the first existence theorem for the BCS gap equation in the spherically symmetric case. Odeh considered the situation of a non-local interaction kernel describing the typical phonon-mediated effective interactions between electrons in metals and alloys. In the same setting, compactness of the integral operator associated to the BCS gap equation and the question of existence of solutions were studied by Billard and Fano and Vansevenant [**10, 81**]. Later, McLeod and Yang [**65, 83**] studied the existence of a critical temperature T_c in the case of a negative interaction kernel V, that is, in a situation where the gap function Δ is no longer expected to be radial.

A heuristic derivation of the BCS functional from quantum mechanics can be found in [**50, 57**]. Let us emphasize here, that the BCS model is a major simplification of the full many-body problem. The whole system is described in quantities that only depend on two variables, the reduced density matrix γ and the Cooper-pair wave function α. Very briefly, the first step in the derivation of the BCS functional, is to restrict the set of states to quasi-free states, so-called BCS states which have the property of satisfying Wick's rule, see [**2**]. Furthermore, one assumes translation invariance and $SU(2)$ rotation invariance. Situations in which the assumption of translation invariance is legitimate have been identified in one of the papers in this thesis [**26**], see Sections 3.2.1 and 5. The result is a functional, which is sometimes called the Bogoliubov-Hartree-Fock functional. Finally, by simply ignoring two terms, the so-called direct and the exchange term, one arrives at the BCS functional. Let us mention here, that, as pointed out by Leggett [**69**], in the physically relevant parameter regimes the two neglected terms are considered unimportant. Mathematical results concerning the last step, that is going from the Bogoliubov-Hartree-Fock functional to the BCS functional, in the derivation of the latter have been published by Bräunlich, Hainzl and Seiringer, see [**16**]. The authors investigated the role of the direct and exchange term, and showed that in the case of short-range interactions these terms only lead to a renormalization of the chemical potential, while the usual properties of the BCS functional are left unchanged. Hence, for the case of short-range interactions considered here, this publication can be seen as a rigorous justification of the last step of the BCS approximation. Let us emphasize here, that, in particular, [**16**] contains the first proof of pairing in the Bogoliubov-Hartree-Fock model in the continuum. In [**17**] the same authors considered the Bogoliubov-Hartree-Fock functional at temperature equal to zero and in the presence of an external electric potential in the low-density limit. By studying the ground state of this model, which consists of a Bose-Einstein condensate of tightly bound fermion pairs, they establish a connection to the Gross-Pitaevskii energy functional. In this context, let us also mention [**56**], where the authors derived the Gross-Piteavskii functional from the BCS functional.

Let us now introduce the BCS functional, before we continue with a short review of recent results in BCS theory. We consider a sample of a fermionic system in d, where

$d = 1, 2, 3$, spatial dimensions. As before, $\mu \in \mathbb{R}$ denotes the chemical potential and $T > 0$ the temperature of the sample. Let V be a local two-body potential through which the fermions interact. In BCS theory the state of the system can be conveniently described by its generalized one-particle density matrix, that is, in terms of a 2×2 operator valued matrix,

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

that satisfies $0 \leq \Gamma \leq 1$ as an operator on $L^2(\mathbb{R}^d) \oplus L^2(\mathbb{R}^d)$. The reduced one-particle density matrix γ is a positive trace class operator on the one-particle space, while α denotes the Cooper-pair wave function, that is, a two-particle wave function, which is only non-zero at temperatures below the critical temperature. The bar simply means complex conjugation. More precisely, $\overline{\alpha}$ has the integral kernel $\overline{\alpha(x, y)}$. Note that Γ is hermitian and thus γ, the one-particle density matrix, is hermitian, meaning that $\gamma(x, y) = \overline{\gamma(y, x)}$, as well. In addition, the Cooper-pair wave function α has to be symmetric, which in terms of kernels means that $\alpha(x, y) = \alpha(y, x)$. Let A denote a magnetic vector potential, then the BCS functional for the free energy of the system under consideration reads

$$\mathcal{F}(\Gamma) = \mathrm{Tr}\left[\left((-i\nabla + A(x))^2 - \mu\right)\gamma\right] + \int V(x-y)\,|\alpha(x,y)|^2 \,\mathrm{d}x\mathrm{d}y - TS\,(\Gamma)\,,$$

where the entropy is given by

$$S\,(\Gamma) = -\mathrm{Tr}\,[\Gamma \log \Gamma]$$

and the trace here is over $\mathbb{C}^2$ and $L^2(\mathbb{R}^d)$. In the situation where one considers the BCS gap equation, the main question is, whether or not there is a non-trivial solution to (3.1). The analogous question in terms of the BCS functional is, whether or not the minimizer of this functional has a non-trivial Cooper-pair wave function, i.e. whether it is the case that $\alpha \neq 0$. The connection between Δ and α will be explained later.

A fundamental work in the mathematical study of the BCS model by Hainzl, Hamza, Seiringer and Solovej [**50**] was published in 2008. The authors consider the translation invariant version of the BCS functional, that is, the BCS functional restricted to translation invariant states, in three dimensions. The study the case of general local pair interactions suitable for the description of interactions in cold Fermi gases. It is shown that the translation invariant BCS functional is bounded from below and attains its minimum [**50**, Proposition 2]. In the case of vanishing interaction potential $V \equiv 0$, a straightforward calculations shows that the minimizer is given by Γ_0 having the entries

$$\gamma = \gamma_0, \text{ with } \gamma_0(p) = \left(1 + e^{(p^2-\mu)/T}\right)^{-1}, \quad \text{and} \quad \alpha = 0.$$

This state, which does not display superconductivity, is referred to as the normal state. The crucial result in [**50**] is, however, the presentation of a *linear* characterization of the critical temperature T_c. Recall that at temperatures below the critical temperature the minimizer of the BCS functional has a non-vanishing Cooper-pair wave function, and, equivalently, the BCS gap equation has a nontrivial solution, while at T_c or above this is not the case. In view of the highly nonlinear BCS functional and BCS gap equation, this is a surprising result. To be more precise, the critical temperature T_c can be defined as being the value of T at which a certain linear operator, which is monotone increasing in T and which we will denote by $K_T + V$, has zero as its lowest eigenvalue. For a more detailed discussion and the explicit form of the operator K_T+V we refer the reader to 3.2.1. Indeed, the authors in [**50**] proved the following three statements to be equivalent,

- In the minimizer of the BCS functional $\alpha \neq 0$.
- The BCS gap equation has a non-trivial solution $\Delta \neq 0$.
- The operator $K_T + V$ has a negative eigenvalue.

It is worth mentioning that the operator K_T+V, the linear operator which characterizes the critical temperature, appears naturally in this context as the second variation of the BCS functional at the normal state Γ_0. Finally, it was also shown in [**50**] that the conditions that μ is positive and $V(x) \leq 0$ for almost all $x \in \mathbb{R}^3$, but V not identically zero, guarantee that T_c is positive. Shortly after, in [**35**], the linear characterization of T_c was used to study the asymptotic behavior of the critical temperature at weak coupling, that is as $\lambda \to 0$ for the local two-body interaction λV. In [**54**], the authors derive upper and lower bounds on the critical temperature and the energy gap at zero temperature for the BCS gap equation. Another work on the BCS critical temperature is [**53**], where the authors consider the case of potentials with negative scattering length and study T_c in the low density limit, that is as $\mu \to 0$. A summary of the above mentioned result can be found in [**55**]. In [**42**], the authors investigate the BCS gap equation for a situation with unequal population of spin-up and spin-down states.

It did not take long until the first works on the BCS functional with external fields were published after the mathematical study of the translational BCS functional and the critical temperature had gained momentum in 2008. The crucial paper in this context is by Frank, Hainzl, Seiringer and Solovej [**37**] and appeared in 2012, where an observation of Gorkov [**47**] was made rigorous. In 1959, only two years after Bardeen, Cooper and Schrieffer had come up their microscopic theory of superconductivity, Gorkov had presented a formal derivation of the Ginzburg-Landau equations from the BCS model for temperatures close to the critical temperature, where it was expected that the macroscopic Ginzburg-Landau theory is a good approximation to the microscopic theory of BCS. Ginzburg and Landau introduced their theory in 1950, see [**46**]. The Ginzburg-Landau equations are used to describe vortex lattices in superconductors in magnetic fields. In particular, these equations were one of the few phenomenological theories which, already before the innovative work and the discovery of the microscopic pairing mechanism of Bardeen, Cooper and Schrieffer, could describe some aspects of superconductivity. The state of the system is represented by a complex valued function of a single position variable which is called ψ, which is zero in the normal state and non-zero in the superconducting state. Hence, the order parameter ψ can be interpreted as a macroscopic wave function and its square $|\psi(x)|^2$ is then proportional to the density of superconducting particles. The work [**37**] is the basis for one of the manuscripts presented in this thesis, see Section 6. Summaries, pointing out the crucial steps in the proof in [**37**] are given in [**38, 39, 57**]. For an introduction to BCS theory and to Ginzburg-Landau theory, we refer to the book of Tinkham [**79**]. The techniques developed in [**37**] were applied afterwards to study several other related aspects. As a first example let us mention a work of the same authors [**40**], where the main focus is on the critical temperature in a situation with external electric and magnetic fields. It is shown, that, to leading order, the critical temperature is given by the critical temperature of the translation invariant BCS functional. However, it turns out that the next-to-leading order of the critical temperature equation in the limit as the ratio between microscopic and macroscopic scale tends to zero is determined by the lowest eigenvalue of the linearization of the Ginzburg-Landau equations. Closely related are the works of Hainzl and Seiringer [**56**] and Hainzl and Schlein [**52**]. While in [**37**] the authors consider periodic external fields and accordingly use the notion of free energy per unit volume, Deuchert in [**25**] made the first step following a different approach, which is based on the assumption that the external fields are sufficiently localized. In [**41**], Frank and Lemm presented an generalization of the derivation of the Ginzburg-Landau theory from the BCS model in the translation-invariant case, but allowing for multiple types of superconductivity, which leads them to the study of multi-component Ginzburg-Landau theory.

Let us mention here another interesting result in the context of the connections between the Ginzburg-Landau and the Bogoliubov-de Gennes equations, which determine the dynamics in BCS theory. In contrast to the natural relation between Ginzburg-Landau and BCS theory in the case of equilibrium states, the connection between the two models in the case of dynamics is not as apparent. In [**36**], Frank, Hainzl, Schlein and Seiringer pointed out that for initial states that are close to thermal equilibrium states at temperatures near the critical temperature, the time-dependent Ginzburg-Landau equation predicts a decay, while the Cooper-pair wave function does not decay in time. This incompatibility of the time-dependent Bogoliubov-de Gennes and the time-dependent Ginzburg-Landau equations was confirmed numerically in the one dimensional case, see [**58**].

In [**34**], Frank, Hainzl and Langmann investigate the influence of a weak homogeneous magnetic field on the critical temperature. It is shown that, within a linear approximation of BCS theory, the critical temperature is lowered by an explicit constant times the field strength, up to higher order terms, in such a magnetic field. For their proof, the authors developed a new method, a rigorous phase approximation, to control the effects of the magnetic field. In the work presented in Section 6 of this thesis, this method is a crucial ingredient.

3.2. Main results part I

3.2.1. Persistence of Symmetry in the BCS model. The goal in [**26**] was to study under which circumstances the minimizer of the full periodic BCS functional (without external fields) is in fact translation invariant. Knowing this, it would be legitimate to work with the translation-invariant version of the BCS functional which is much simpler than the full BCS model in these situations. Previous to the result we present here, persistence of translational symmetry was only known for the case $\hat{V} \leq 0$ and not identically zero, see [**57**]. The main result of the paper applies for the BCS functional in two dimension with radial pair interaction and says that the translational symmetry indeed does persist, at least for temperatures in a certain interval below T_c.

Before we discuss this result further, let us first introduce the full BCS functional. In order to avoid having to deal with boundary conditions we consider a periodic situation and, for simplicity, choose the lattice $\mathbb{Z}^d$ with the unit cell $\Omega = [0,1]^d$, where $d = 2, 3$ here. In this setting, we have to calculate all energies per unit volume. Let us thus define the trace per unit volume Tr_Ω of a periodic operator A by $\mathrm{Tr}_\Omega[A] = \mathrm{Tr}[\chi_\Omega A \chi_\Omega]$, where χ_Ω is the characteristic function of Ω. Periodic BCS states are most conveniently described by self-adjoint operators

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

on $L^2(\mathbb{R}^d) \oplus L^2(\mathbb{R}^d)$ with the property that $0 \leq \Gamma \leq 1$. Here, γ and α are periodic operators with period one and the bar denotes complex conjugation. A periodic BCS state is called admissible if $\mathrm{Tr}_\Omega(-\nabla^2 + 1)\gamma < \infty$ and in this case we write $\Gamma \in \mathcal{D}$. The full BCS functional at temperature $T \geq 0$, with chemical potential μ, interaction potential $V \in L^2(\mathbb{R}^d)$ and entropy

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

reads

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

We get the translation invariant version of the full BCS functional by restricting $\mathcal{F}$ to translation invariant admissible states, that is to states $\Gamma \in \mathcal{D}$ which satisfy additionally, that the kernels $\gamma(x,y)$ and $\alpha(x,y)$ of γ and α, respectively, are functions of $x - y$. The translation invariant BCS functional reads

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

To summarize and make the notions used more precise, let us note that we have a functional invariant under spatial translations in the sense that such translations do not change the energy of a state. However, this does not mean that minimizers of the functional are necessarily translation invariant. If there is a minimizer, which is not translation invariant, we say that the translational symmetry is broken. The question whether translational symmetry in the BCS model persists or if the translation symmetry is broken is important for various reasons. First let us mention that in the derivation of the BCS model from quantum mechanics one assumes the states to be translation invariant. Supposedly, this approximation is valid in the case of cold fermionic gases with a rotationally invariant pair interaction - and this was the starting point in the discussed project. Nevertheless, situations where the approximation is not valid are known, see for example [**2**] for an example in solid state physics. Apart from that, $\mathcal{F}^{\text{ti}}$ is much simpler than $\mathcal{F}$ and, for instance, is even suitable for numerical computations.

Before we finally state the results obtained in this context, let us very briefly discuss the definition of the critical temperature as presented in [**50**]. The critical temperature T_c can be defined in terms of the operator $K_T + V$, where K_T denotes the operator that acts by multiplication by

$$K_T(p) = \frac{p^2 - \mu}{\tanh((p^2 - \mu)/(2T))}$$

in Fourier space. The fact that K_T is increasing in T allows us to define the critical temperature as being the value of T such that the lowest eigenvalue of $K_T + V$ is exactly zero. In other words,

$$T_c = \inf \{T | K_T + V \geq 0\}.$$

The main results [**26**, Theorem 5.1 and Theorem 5.2] can be summarized as follows.

THEOREM 3.1. *Assume that $V \in L^2(\mathbb{R}^2)$, with $\hat{V} \in L^r(\mathbb{R}^2)$, where $r \in [1, 2)$, is radial and such that $T_c > 0$. In the case that the lowest eigenvalue of $K_{T_c} + V$ is at most twice degenerate, there exists $\widetilde{T} < T_c$, such that at all temperatures $T \in [\widetilde{T}, T_c)$ the minimizers of the periodic BCS functional $\mathcal{F}$ are translation invariant.*

The same holds true in the three dimensional case, if $V \in L^2(\mathbb{R}^3)$, with $\hat{V} \in L^r(\mathbb{R}^3)$, where $r \in [1, 12/7)$, is radial and such that $T_c > 0$ and such that the lowest eigenvalue of $K_{T_c} + V$ is non-degenerate.

More precisely, the proof shows that the minimizers of $\mathcal{F}$ in the case specified in the above theorem take the form $(\gamma_{\ell_0}, \alpha_{\ell_0})$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0})$, where

$$\alpha_{\pm}(p) = e^{\pm i \ell_0 \varphi} \sigma_{\ell_0}(p),$$

and $\ell_0 \in 2\mathbb{N}_0$, φ denotes the angle of p in polar coordinates and σ_{ℓ_0}, as well as γ_{ℓ_0} are radial functions. Note that ℓ_0 corresponds to the sector of angular momentum of the ground state of $K_{T_c} + V$. This means that in the three dimensional situation where we assume that the ground state of $K_{T_c} + V$ has vanishing angular momentum, we always have $\ell_0 = 0$ and the minimizer of the periodic BCS functional is radial. Furthermore, the proof tells us that the minimizers are unique, up to phases in front of the Cooper-pair wave function α.

The paper is complemented by Proposition 5.3, where it is shown that the strategy of the proof of Theorem 3.1, that is Theorems 5.1 and 5.2, cannot be used to prove a more general result in three dimensions.

3.2.2. Bogoliubov-de Gennes and Ginzburg-Landau equations. The result presented in Section 6 is closely related to the groundbreaking work of Frank, Hainzl, Seiringer and Solovej [**37**]. Already in 1959, Gorkov [**47**] had presented an idea, how, close to the critical temperature, the phenomenological Ginzburg-Landau theory arises from the microscopic BCS model. De Gennes presented a simpler version of Gorkov's argument in his textbook [**45**], see also the work of Eilenberger [**28**]. In [**37**], see also [**38, 39**], the authors identified the parameter regime where Ginzburg-Landau theory is a valid approximation of the BCS model. Moreover, they presented the first rigorous result to this question with quantitative error bounds. As in [**37**], we study the connection between BCS Theory and the Ginzburg-Landau model. However, while the authors in [**37**] treat the magnetic potential as a fixed parameter, we consider it as an independent variable, which allows us to also derive the second Ginzburg-Landau equation.

We consider a superconductor in a box $Q_h \subset \mathbb{R}^2$ with side length h^{-1} and impose periodic boundary conditions. In BCS theory, the state of the system is described by the generalized one-particle density matrix Γ and the magnetic potential A. To be more precise,

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

where, $0 \leq \Gamma(1-\Gamma) \leq 1$, and, as before, the bar denotes complex conjugation. Furthermore, γ is a self-adjoint operator on $L^2(\mathbb{R}^2)$ and α is an operator on $L^2(\mathbb{R}^2)$ satisfying $\alpha^* = \overline{\alpha}$. We assume that γ and α commute with translations of the lattice, which, in terms of kernels, means that $\alpha(x+h^{-1}t, y+h^{-1}t) = \alpha(x,y)$ for all $t \in \mathbb{Z}^2$, or more generally for any lattice. We assume that $A : \mathbb{R}^2 \to \mathbb{R}^2$ is also periodic, i.e., $A(x+h^{-1}t) = A(x)$ for all $t \in \mathbb{Z}^2$. In the situation described here, it is natural to consider energies per unit volume and the trace per unit volume is defined as in the previous section.

Let us denote the magnetic Laplacian by $-\Delta_A = (-i\nabla + A)^2$. We assume that Γ and A minimize the BCS energy functional at temperature $T \geq 0$,

$$\begin{aligned}\mathcal{F}_T^{\mathrm{BCS}}(\Gamma, A) = \operatorname{Tr}_{Q_h}\left((-\Delta_A - \mu)\,\gamma\right) - T \operatorname{Tr}_{Q_h} S(\Gamma) \\ - \int_{\mathbb{R}^2} \mathrm{d}x \int_{Q_h} \mathrm{d}y\, V(x-y)|\alpha(x,y)|^2 + \int_{Q_h} \mathrm{d}x\, |\operatorname{curl} A(x)|^2,\end{aligned}$$

where $\mu \in \mathbb{R}$ is the chemical potential, S is the entropy function $S(x) = -x\log x$, and where $-V : \mathbb{R}^2 \to \mathbb{R}$ is the interaction potential. We require V to satisfy the following assumptions.

ASSUMPTION 1. *We suppose that*

- *V is even and positive*
- *V is such that $\mathcal{T}_c > 0$ and that α_*, the ground state of $K_{\mathcal{T}_c} - V$, is non-degenerate*
- *$V \in L^\infty(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ and $|x|^2 V \in L^\infty(\mathbb{R}^2)$.*

Minimizers (Γ, A) of the BCS functional are critical points and hence they must be solutions of the corresponding Euler-Lagrange equations,

$$\frac{1}{2} H_A(-2V\alpha) = TS'(\Gamma), \tag{3.2a}$$

$$\operatorname{curl}^* \operatorname{curl} A + \operatorname{Re}\,(-i\nabla + A)\,\gamma|_{y=x} = 0, \tag{3.2b}$$

where $\operatorname{curl}^* = (\partial_2, -\partial_1)^T$ and $H_A(\Delta)$ is the matrix of operators

$$H_A(\Delta) = \begin{pmatrix} k_A & \Delta \\ \overline{\Delta} & -\overline{k}_A \end{pmatrix},$$

where $k_A := -\Delta_A^2 - \mu$ denotes the kinetic energy. Moreover, $V\alpha$ is the operator whose kernel is $V(x-y)\alpha(x,y)$, i.e., we think of V as a two-body multiplication operator. The notation $|_{y=x}$ denotes the diagonal of the operator.

We now turn to a very brief introduction to Ginzburg-Landau theory. The Ginzburg-Landau model is much simpler than the BCS model. The system is described by a function of only one variable. This function, which we will denote by ψ, only describes macroscopic aspects of the system, while the BCS state Γ contains macroscopic and microscopic information. We assume that the microscopic scale of the system is of order h, where $h \ll 1$, relative to the macroscopic scale. The magnetic field a appearing in the Ginzburg-Landau functional varies on the macroscopic scale, that is on the scale of the box $Q = Q_1$. Let us mention here that, on the other hand, the potential V lives on the microscopic scale. The superconductor is described by a complex valued order parameter ψ, that is a function $\psi : \mathbb{R}^2 \to \mathbb{C}$, and the magnetic potential $a : \mathbb{R}^2 \to \mathbb{R}^2$. These functions are again periodic, to be precise, $\psi(x+t) = \psi(x)$ and $a(x+t) = a(x)$ for all $t \in \mathbb{Z}^2$. We assume that ψ and a minimize the Ginzburg-Landau energy functional,

$$\mathcal{E}^{\mathrm{GL}}(\psi, a) := \frac{1}{2}\int_Q \overline{(-i\nabla + a)\psi} \cdot \mathbb{B}(-i\nabla + a)\psi - C_1|\psi|^2 + \frac{1}{2}C_2|\psi|^4 + \frac{1}{2}C_3\left|\operatorname{curl} a\right|^2,$$

where, $\mathbb{B}$ is a real symmetric 2×2 matrix, and C_1 , C_2, and C_3 are positive constants. The Ginzburg-Landau functional has been the subject of intensive mathematical study, see for instance [**33, 78**].

Critical points of $\mathcal{E}^{\mathrm{GL}}$ satisfy the Ginzburg-Landau equations,

$$(-i\nabla + a) \cdot \mathbb{B}\left(-i\nabla + a\right)\psi - C_1\psi + C_2\left|\psi\right|^2\psi = 0$$

$$C_3 \operatorname{curl}^* \operatorname{curl} a + \operatorname{Re}\overline{\psi}\mathbb{B}\left(-i\nabla + a\right)\psi = 0,$$

that is the corresponding Euler-Lagrange equations.

We now turn to the formulation of our main result. For a periodic operator α on $\mathbb{R}^2$ we define the $\mathcal{L}_h^2$ norm by

$$\|\alpha\|_{\mathcal{L}_h^2}^2 := h^2 \operatorname{Tr}_{Q_h}\left(\alpha^*\alpha\right).$$

In our main theorem we need to assume that

Assumption 2.

- $T_h = \mathcal{T}_c(1 - Dh^2)$ *for some constant* D
- $\|\alpha_h\|_{\mathcal{L}_h^2} \lesssim h$ *and* $\|(\nabla_x + \nabla_y)\alpha_h\|_{\mathcal{L}_h^2} \lesssim h^2$
- A_h *satisfies the gauge conditions* $\operatorname{div} A_h = 0$ *and* $\displaystyle\int_{Q_h} \mathrm{d}x\, A_h(x) = 0,$

- A_h *is smooth and such that* $\|A_h\|_{L^2(Q_h)} \lesssim h$ *and*

$$\left\| \, (\mathrm{curl}^* \;\; \mathrm{curl})^{k/2} A_h \, \right\|_{L^2(Q_h)} \lesssim h^{k+1} \quad \textit{for } k = 1, 2, 3, 4.$$

The main result is the following.

THEOREM 3.2. *Let* V *be such that it satisfies* Assumptions 1. *Suppose that* (Γ_h, A_h) *is a sequence of solutions of equations* (3.2a) *and* (3.2b) *at temperature* T_h *so that the* Assumptions 2 *are satisfied.*

Then for sufficiently small h *we have the decompositions*

$$\begin{aligned} V\alpha_h(x,y) &= hV(x-y)\alpha_*(x-y)\psi_h\left(h(x+y)/2\right) + \sigma_h(x,y), \\ A_h(x) &= ha_h(hx) \end{aligned}$$

where $\|\psi_h\|_{H^1(Q)} \lesssim 1$, $\|\sigma_h\|_{\mathcal{L}^2_h} = O(h^2 h^{-17/48})$, *and* $\|a_h\|_{H^2(Q)} \lesssim 1$ *as* $h \to 0$.

Moreover, if (ψ_*, a_*) *is a weak limit point of the sequence* $\{(\psi_h, a_h)\}$ *in* $H^1(Q) \times \vec{H}^1(Q)$*, then* $(\psi_*, 2a_*)$ *is a weak solution of the Ginzburg-Landau equations with appropriately chosen coefficients. It also follows that if*

$$\limsup_{h \to 0} h^{-1} \|\alpha_h\|_{\mathcal{L}^2_h} \gtrsim 1,$$

then there exists a non-trivial solution of the Ginzburg-Landau equations, i.e., a solution where $\psi \not\equiv 0$.

For the explicit form of the Ginzburg-Landau coefficients $\mathbb{B}, C_1, C_2$ and C_3, we refer to Theorem 6.1 in Section 6. Let us emphasize here that all that is needed to compute the coefficients is the BCS data, i.e., $V, \mu, \mathcal{T}_c$ and D.

Finally, note that by weak limit, we mean that $(\psi_{h_n}, a_{h_n}) \rightharpoonup (\psi_*, a_*)$ for some sequence $h_n \to 0$. We prove that weak limits $(\psi_*, 2a_*)$ of $\{\psi_h, 2a_h\}$ are weak solutions of the Ginzburg-Landau equation. By a standard bootstrap argument, one easily sees that ψ_* and a_* are in H^2. In fact, if V is regular enough, they are even smooth, and are indeed strong solutions.

CHAPTER 4

The Kac model and approach to equilibrium

The reader should be warned at the outset that more questions will be raised than answered. However, we hope to provide sharp formulations and thus perhaps pave the way toward further work on this fascinating borderline between mathematics and physics.

Mark Kac 1914 - 1984,
Foundations of Kinetic Theory, 1956.

4.1. Introduction

The Kac model and the Kac master equation. In his model which he introduced in 1956 in his article "Foundations of Kinetic Theory", Kac assumes a spatially homogeneous system. Moreover, the system is such that the state of the system is entirely specified by the velocities of the particles. Very briefly, in the Kac model, after waiting for an exponentially distributed time interval, a pair of particles is selected randomly and uniformly. One lets these two particles collide with a scattering angle that is also randomly selected. The time evolution of this system is now described by the so-called Kac master equation. The Kac master equation is a linear master equation that determines the time evolution for the probability distribution of finding the system in a given state, that is, of finding the particles in the system with certain velocities.

In contrast to the very successful Boltzmann equation, the model Kac developed is not based on mechanical principles, but it is based on clear and simple probabilistic assumptions. That means in particular that Kac's model has not been derived from first principles and therefore cannot be considered fundamental. In fact, the justification of the Kac model is a-posteriori, through a connection to the Boltzmann equation. It was actually Kac's main motivation in [**61**] to rigorously derive the non-linear spatially homogeneous Boltzmann equation, see also [**62**].

The simplicity of the Kac model allows for the study of interesting issues that are difficult to understand in more fundamental models like Newtonian mechanics, as for example the question of approach to equilibrium for the case of large particle systems. This will be explained and discussed in the next section. Here, in the first section of the introduction, we introduce the Kac model itself and we present recent results that connect to and build a context for [**13**].

We consider a system of N, where $N \in \mathbb{N}$, indistinguishable particles in one dimension. We assume that the state of the system is specified completely by the velocities of the particles. We denote the velocity of particle i, for $1 \leq i \leq N$, by v_i, where $v_i \in \mathbb{R}$, and collect the velocities of all particles $1, ..., N$ in the system in the velocity vector

$$\mathbf{v} = (v_1, v_2, ..., v_N) \,.$$

After a certain time, two randomly and uniformly selected particles collide with a random scattering angle. That means the velocities of these two particles, that is their pre-collisional velocities, are replaced by new, or post-collisional, velocities in such a way that the total energy

$$E = \sum_{k=1}^{N} v_i^2$$

is preserved. To be more precise, a collision of, say, particles i and j, where $1 \leq i < j \leq N$, is described as follows. We replace

$$\mathbf{v} = (v_1, ..., v_i, ..., v_j, ..., v_N) \quad \text{by} \quad \mathbf{v}^*_{ij}(\theta) = \left(v_1, ..., v_i^*(\theta), ..., v_j^*(\theta), ..., v_N\right),$$

where $\theta \in [-\pi, \pi)$ is a randomly selected scattering angle,

$$v_i^*(\theta) = v_i \cos\theta - v_j \sin\theta \quad \text{and} \quad v_j^*(\theta) = v_i \sin\theta + v_j \cos\theta.$$

The collisions happen after exponentially distributed time intervals. The average time between two successive collisions of a given particle is independent of the number of particles. Let us note here that three or more particle collisions are not taken into account in the Kac model. However, their effect on some properties of the time evolution are probably significant in reality.

Having understood the collision process in our model, we now aim for the time evolution of the system, that is the Kac master equation. Therefore, let us introduce a few notions. The state of the system is given by the function $f : \mathbb{R}^N \to \mathbb{R}_+$. To be precise, $f(\mathbf{v})$ is the probability density of finding the particles in the system with velocities $\mathbf{v}$. The result of a collision of particles i and j is described by the collision operator R_{ij} given by

$$R_{ij}\,[f]\,(\mathbf{v}) = \int_{-\pi}^{\pi} \rho(\theta)\,\mathrm{d}\theta\, f\left(\mathbf{v}^*_{ij}(\theta)\right),$$

where ρ is the probability density for the selection of the scattering angle θ in a collision process and, in particular, we have

$$\int_{-\pi}^{\pi} \rho(\theta)\,\mathrm{d}\theta = 1.$$

We denote by $f(\mathbf{v}, t)$ the probability density of the velocities of the particles in our system at time t, for $t \geq 0$. The time evolution is now given by the Kac master equation

$$\frac{\partial f}{\partial t} = \mathcal{L} f, \quad f(\mathbf{v}, 0) = f_0(\mathbf{v}), \tag{4.1}$$

where the infinitesimal generator $\mathcal{L}$ of this evolution is given by

$$\mathcal{L} = \frac{\lambda}{N-1} \sum_{1 \le i < j \le N} (R_{ij} - I) \, .$$

Here, I simply denotes the identity operator, whereas λ is the inverse of the average time between two collisions that involve a certain particle. More precisely, the times between collisions of a certain particle, say particle i for $1 \le i \le N$, are exponentially distributed with parameter λ. It is worth mentioning that the so-called mean free time $1/\lambda$ is the only parameter with physical significance in the model. This is because the evolution given by (4.1) is completely independent of the density of the particles, that is of the positions of the particles.

Approach to equilibrium. As was already mentioned above, one of the goals of Kac's original work [**61**] was to give a mathematical framework for the study of approach to equilibrium. Kac pointed out that for the case where the particle number N becomes large, one could show approach to equilibrium in a quantitative way by proving that the gap of the generator is bounded below uniformly in N. The lower uniform lower bound for the gap of the generator is known as Kac's conjecture. Some attempts to prove this conjecture were made in [**27**], where the authors show that the gap for a system of N particles, Δ_N, is bounded below by

$$\Delta_N \ge \frac{C}{N^2}.$$

Finally, in 2001 E. Janvresse in [**60**] proved Kac's conjecture applying H.-T. Yau's martingale method [**85, 86**]. Shortly thereafter, E. Carlen, M. Carvalho and M. Loss in [**19, 20**] computed the gap explicitly. For the case where the scattering angle θ is chosen uniformly, meaning that $\rho(\theta) = (2\pi)^{-1}$, they proved that

$$\Delta_N = \frac{N+2}{2(N-1)}.$$

Consequently, in the limit as N becomes large, $\Delta_N \to 1/2$. The strategy making use of the gap yields satisfactory results if applied once the system is already close to equilibrium. However, it turns out that this strategy does not work equally well if the system is far from equilibrium. To see why this is the case, let the function f, the probability density of the velocities of the particles, be given by the product of the N probability densities f_i, for $i = 1, ..., N$, of all the particles in the system, that is $f = \prod_{i=1}^{N} f_i$. Of course, f is normalized. Approach to equilibrium means that

$$\|f(\cdot, t) - 1\|_{L^2(\mathbb{R}^N)} \to 0.$$

However, it is reasonable to think that the functions f_i are pairwise almost independent, which implies

$$\|f\|_{L^2(\mathbb{R}^N)} \approx \prod_{j=1}^{N} \|f_j\|_{L^2(\mathbb{R})} = e^{C \cdot N}$$

for some positive constant C. Obviously, the same approximation is true for $\| \, f \, - \, 1 \, \|_2$. Therefore, by only using the gap estimate, one cannot get a better result than that it takes a time of order N to relax to the equilibrium distribution, which is obviously not what is expected from physics. Put differently, although measuring approach to equilibrium in terms of an L^2 distance as in the gap argument seems to be a natural strategy for the problem, the obtained result is not satisfactory.

Another natural approach to measure approach to equilibrium is given by the entropy. As we expect exponential decay of the entropy, what one would like to show is that the entropy production, that is the negative time derivative of the entropy, is proportional to the entropy itself. However, it turned out that this is not the case. C. Villani chose this approach and proved in [**82**] that the entropy decays exponentially. In the same work, he also showed that the exponential rate is bounded below, but this bound is inversely proportional to the particle number N. In this context, there is also a very interesting result by A. Einav. In [**29**], he studied an initial state, where most of the energy is concentrated in very few particles and most of the others particles have very little energy in contrast. The physical intuition that most of the particles are almost in some kind of equilibrium and hence, such a state still has low entropy production and at the same time is very improbable is made rigorous in his work. Einav shows the considered state to have entropy production essentially of order $1/N$. Let us mention here that low entropy production does not necessarily exclude exponential decay in entropy. This insight suggests, that one reconsiders the problem but restricts the set of initial states, excluding, in particular, highly improbable states like the one studied by Einav. The question raised here is: how can one characterize the states for which the entropy converges to zero on a reasonable time scale? S. Mischler and C. Mouhot followed that path and presented a series of results in their general investigation of Kac's program for gases of hard spheres and true Maxwellian molecules in three dimensions in [**73, 74**]. For these systems the authors could prove approach to equilibrium in relative entropy as well as in Wasserstein distance. The rate of relaxation they obtain is uniform in the particle number N, but polynomial in time. As mentioned above, this result does not hold for any initial state. The authors consider a natural class of chaotic states, which shifts the problem of finding the "right" initial conditions that allow for a proof of exponential decay to the level of the non-linear Boltzmann equation.

Another idea for a reasonable initial state is to couple a small system of particles out of equilibrium to a large system, that is a heat bath, in equilibrium. This path was taken in [**43**] by Fröhlich and Gang, where approach to equilibrium was shown for the spatially inhomogeneous Boltzmann equation coupled to a thermostat. In [**11, 12**] the authors studied particles in an electric field interacting with external scatterers. In this setting, the thermostat is given by a deterministic friction term, while stochasticity is provided by the collisions with the obstacles.

Another result in this direction is by Bonetto, Loss, and Vaidyanathan [**15**], where the authors prove in two ways that any initial distribution approaches the equilibrium distribution exponentially. They compute the gap of the generator of the evolution, but also prove exponential decay in relative entropy for a thermostated system. More precisely, they consider a system of interacting particles coupled to a thermostat, that is coupled to an infinite gas at thermal equilibrium at some inverse temperature β. The particles in the heat bath are at equilibrium, which means that their distribution is given by a Gaussian at inverse temperature β. A crucial point in this work is that the reservoir is not influenced by the interactions with the N particles in the system. Particles that are in the heat bath, and hence in equilibrium, stay in equilibrium forever. As also mentioned in [**15**], this infinite reservoir is not very realistic. It would be interesting to study the setting where the reservoir still is very large compared to the system, but finite. A first step in this direction was presented by the same authors together with Tossounian in [**14**], where a small system of M, for $M \in \mathbb{N}$ particles coupled to a large but finite reservoir initially in equilibrium at temperature β^{-1} of $N \gg M$ particles is considered. It is shown that the evolution of the system coupled to the finite reservoir can be approximated by the evolution of the thermostated system, where the infinite heat bath coupled to the system is in equilibrium at the same temperature β^{-1}. In the mathematical model describing this situation, one must have three rates at which the different kinds of collision occur. The rate λ_S belongs to the exponentially distributed waiting times between collisions of particles in the system, while λ_R is the rate at which particles in the reservoir will collide. The latter needs only to be taken into account in the case of finite heat bath, of course. The third kind of collision is an interaction between the system and the reservoir. The rate μ is chosen in such a way that the average time between two successive interaction of a specific system particle with the reservoir is independent of M, the number of particles in the system, and, in the case of finite reservoir, independent of N, the number of particles in the reservoir. A particle in the reservoir will scatter with a particle in the system at rate $\mu M/N$. Thus, in the case where the reservoir is very large compared to system, one can expect that the reservoir does not deviate from its initial equilibrium state. This intuition is made rigorous in [**14**]. The presented result is uniform in time and holds in L^2 as well as for the Gabetta-Toscani-Wennberg distance. The L^2 version of the result does not always give a satisfactory result, even when applied to some very reasonable initial distributions. For example, if one assumes that the system is initially also in equilibrium but at some different temperature $\beta_S^{-1} \neq \beta^{-1}$, the presented result [**14**, Theorem 1] tells us that the number of particles in the reservoir N has to be exponentially large in the number of system particles M. This is one of the reasons why the authors considered the Gabetta-Toscani-Wennberg metric, although it is technically much more complicated.

Let us summarize very shortly. From what is known and has been proved so far, there is no *mathematical* argument that indicates that one can expect exponential decay uniformly in N of the entropy in the Kac model. On the contrary, the work of A. Einav [**29**] mentioned above shows that, at least for physically improbable states, exponential decay with a rate independently of N cannot be true. With that said, our result in [**13**], where we show exponential decay for the Kac evolution in entropy relative to the thermal state can be seen as going in the exact opposite direction to earlier results, as presented, for example, by C. Villani, [**82**], and A. Einav, [**29**]. We will explain our result in the following chapter.

4.2. Main results part II

4.2.1. Entropy decay for the Kac evolution. In [**13**], we consider the situation of a small system of M particles coupled to a larger, but finite, reservoir of $N \geq M$ particles. The particles in the reservoir are at equilibrium at inverse temperature β. We show that solutions to the Kac master equation have exponential decay in entropy relative to the thermal state at temperature β. In particular, we get an explicit rate that is almost independent of the particle number. In order to state our main result, let us first explain the model in more detail.

As before, we denote by $\mathbf{v}$ the velocity vector of the system. The system contains M particles, hence $\mathbf{v} \in \mathbb{R}^M$. It is coupled to a reservoir of $N \geq M$ particles. The velocities of the particles in the reservoir are gathered in the vector $(w_{M+1}, ..., w_{M+N}) = \mathbf{w}$, $\mathbf{w} \in \mathbb{R}^N$. Note that the indexing is continued, so that the $M+1$-th entry of the velocity vector $(\mathbf{v}, \mathbf{w})$ of the whole setting, i.e. system and reservoir, is the first entry of the vector $\mathbf{w}$. So here, the probability distribution F of system *and* reservoir is a function on $\mathbb{R}^{M+N}$. In order to define the collision operators R_{ij}, where $1 \leq i < j \leq M+N$, elegantly in our setting, let us first introduce the notation $r_{ij}(\theta)$ for the rotation matrix for a rotation with angle $\theta \in [-\pi, \pi)$ acting as

$$r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w}) = (v_1, ..., v_i \cos(\theta) - v_j \sin(\theta), ..., v_i \sin\theta + v_j \cos\theta, ..., v_M, \mathbf{w}),$$

as long as $1 \leq i < j \leq M$. For $1 \leq i \leq M$, but $M+1 \leq j \leq N$, that is particle j is in the reservoir we have

$$r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w}) = (v_1, ..., v_i \cos(\theta) - w_j \sin(\theta), ..., w_{j-1}, v_i \sin\theta + w_j \cos\theta, ..., w_{M+N}),$$

respectively. The case where $M \leq i < j \leq M+N$, that means the collision takes place in the reservoir, works analogously. This notation is very useful for the definition of the collision operators, because it prevents further distinction of cases. For $1 \leq i < j \leq M+N$ the collision operator R_{ij} is given by

$$(R_{ij}F)(\mathbf{v}, \mathbf{w}) = \int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\, F\left(r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w})\right).$$

We assume the probability measure ρ to be smooth. Furthermore, we require that

$$\int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\, \sin\theta \cos\theta = 0\,. \tag{4.2}$$

The infinitesimal generator

$$\mathcal{L} = \frac{\lambda_S}{M-1} \sum_{1 \leq i < j \leq M} (R_{ij} - I) + \frac{\lambda_R}{N-1} \sum_{M < i < j \leq M+N} (R_{ij} - I) + \frac{\mu}{N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} (R_{ij} - I)$$

of the evolution consists of three parts in our setting. The first term describes the interactions within the system, the second term describes the interactions within the reservoir and the third term represents interactions between system and reservoir particles. Here λ_S and λ_R are the parameters of the exponential distributions of the waiting times between collisions within the system or the reservoir, respectively. Similarly, μ denotes the rate at which one certain particle in the system will scatter with the reservoir, that is, with any particle in the reservoir. We consider the evolution in $L^1(\mathbb{R}^{M+N})$ with Lebesgue measure. The reservoir is at equilibrium at inverse temperature β. Without loss of generality we assume $\beta = 2\pi$. Hence, we choose our initial condition as

$$F_0(\mathbf{v}, \mathbf{w}) = f_0(\mathbf{v}) e^{-\pi|\mathbf{w}|^2}, \tag{4.3}$$

where f_0 is an arbitrary probability distribution on $\mathbb{R}^M$ representing the initial velocity distribution of the particles in the system. The Kac master equation of our system is given by,

$$\frac{\partial F}{\partial t} = \mathcal{L}F\,, \quad F(\mathbf{v}, \mathbf{w}, 0) = F_0(\mathbf{v}, \mathbf{w}) = f_0(\mathbf{v}) e^{-\pi|\mathbf{w}|^2}.$$

We introduce the following notation

$$f(\mathbf{v}, t) = \int_{\mathbb{R}^N} \left[e^{\mathcal{L}t} F_0 \right] (\mathbf{v}, \mathbf{w}) \, \mathrm{d}\mathbf{w}$$

in order to define the entropy of f relative to the thermal state $e^{-\pi|\mathbf{w}|^2}$ as

$$S(f(\cdot, t)) := \int_{\mathbb{R}^M} f(\mathbf{v}, t) \log \left(\frac{f(\mathbf{v}, t)}{e^{-\pi|\mathbf{v}|^2}} \right) \mathrm{d}\mathbf{v}\,.$$

Our main result is the following.

THEOREM 4.1. *Let $N \geq M$ and let ρ be a probability distribution with an absolutely convergent Fourier series such that* (4.2) *holds. The entropy of f relative of to the thermal state $e^{-\pi|\mathbf{v}|^2}$ then satisfies*

$$S(f(\cdot, t)) \leq \left[\frac{M}{N+M} + \frac{N}{N+M} e^{-t\mu_\rho (N+M)/N} \right] S(f_0)\ ,$$

where

$$\mu_\rho = \mu \int_{-\pi}^{\pi} \rho(\theta) \, \mathrm{d}\theta \, \sin^2(\theta)\ ,$$

and f_0 is as introduced in (4.3).

Let us first emphasize that in Theorem 4.1, we do *not* consider the entropy relative to the equilibrium state, but the entropy relative to the thermal state. We do not a see a way to adapt our arguments in the proof to the case of the entropy relative to the equilibrium state. While the entropy relative to the equilibrium state does tend to zero as t tends to infinity, we do not see any indication that it does tend to zero at an exponential rate. The initial condition play a crucial role here and we suspect that if exponential decay can be shown, the rate would depend on the initial condition. The decay rate we get in Theorem 4.1 only depends on the parameters μ and the ρ. In particular, the interactions rates for the system λ_S and for the reservoir λ_R do not appear and hence the intensity of the collision processes within the system and the reservoir have no influence here.

Next let us briefly comment on the assumptions on ρ. Theorem 4.1 does not depend on the ergodicity of the evolution and, in particular, the result holds for the case were ρ is a delta measure with mass at the angles $\theta = \pm\pi/2$. Generally, the theorem applies to the case where ρ is a finite sum of Dirac measures.

Let us close the presentation of this result with a remark concerning the optimality of the derived rates and the connection to the work in [**14**]. As the number of particles N tends to infinity, Theorem 4.1 corresponds to the situation studied in [**14**]. Indeed, our results yields the exact same decay rate in this case. As the decay rate for the thermostat situation considered there is known to be optimal, as shown in the work of Tossounian and Vaidyanathan [**80**], the decay rate given in Theorem 4.1 is optimal as well.

The universality of the obtained decay rate, in the sense that, as explained above, the decay rate does depend only on μ and ρ allows for an application of the result to the standard Kac model. More precisely, in the language introduced above, we consider *a system* of $N + M$ particles evolving with the generator $\mathcal{L}_{\text{cl}}$ given by

$$\mathcal{L}_{\text{cl}} = \frac{2}{N+M-1} \sum_{1\le i<j\le N+M} (R_{ij} - I)\,.$$

In order to apply our result we split the particles in two groups, one containing M and one containing N particles. So, artificially, the generator $\mathcal{L}_{\text{cl}}$ now is of the wished form, i.e.,

$$\mathcal{L}_{\text{cl}} = \frac{2}{N+M-1} \sum_{1\le i<j\le M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{M+1\le i<j\le N+M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{i=1}^{M} \sum_{j=M+1}^{N+M} (R_{ij} - I)\,,$$

accordingly the collisions rate are now given by

$$\lambda_S = \frac{2(M-1)}{N+M-1}\ ,\ \lambda_R = \frac{2(N-1)}{N+M-1} \quad \text{and} \quad \mu = \frac{2N}{N+M-1}\,.$$

As a direct translation of Theorem 4.1 we get the following Corollary.

COROLLARY 4.2. *Let $N \ge M$ and consider the time evolution defined by $\mathcal{L}_{\text{cl}}$ with initial condition* (4.3). *Assume that the function f_0 in the initial condition has finite entropy. The entropy of the function*

$$f(\mathbf{v}, t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}_{\text{cl}} t} F_0\right](\mathbf{v}, \mathbf{w})\, \mathrm{d}\mathbf{w}$$

relative to the thermal state $e^{-\pi|\mathbf{v}|^2}$, satisfies

$$S(f(\cdot, t)) \le \left[\frac{M}{N+M} + \frac{N}{N+M} e^{-t\mu_\rho 2(N+M)/(N+M-1)}\right] S(f_0)\ ,$$

where

$$\mu_\rho = \int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\, \sin^2(\theta)$$

and ρ is a probability distribution such that (4.2) *holds.*

In a particular case the above results can also be extended to three-dimensional Boltzmann-Kac collisions.

From our viewpoint, the main achievement of [**13**] is the presentation of a simple argument which allows one to show exponential relaxation towards equilibrium. In our proof we combine, apart from some computations, an iterated version of Nelson's hypercontractive estimate, and a sharp version of the Brascamp-Lieb inequalities, see [**18**] and also [**70**]. The standard way of proving relaxation towards equilibrium through logarithmic Sobolev inequalities is not available in the present situation, since the generator $\mathcal{L}$ of the time evolution is not given by a Dirichlet form.

CHAPTER 5

Persistence of translational symmetry in the BCS model with radial pair interaction

A. Deuchert, A.Geisinger, C. Hainzl, M. Loss

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. In the case of vanishing angular momentum our results carry over to the three-dimensional case.

5.1. 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 [**77, 10, 81, 83, 65, 84**] for works on the gap equation with interaction kernels suitable to describe the physics of conduction electrons in solids and [**50, 35, 54, 55, 16, 42, 41**] for works that treat the translation-invariant BCS functional with a local pair interaction. The gap equation and the BCS functional are related in the way that the former is the Euler-Lagrange equation of the latter. One main question in the study of BCS theory is whether the gap equation

$$\Delta(p) = -\frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \hat{V}(p-q) \frac{\tanh\left(E(q)/2T\right)}{E(q)} \Delta(q)\, \mathrm{d}q, \tag{5.1}$$

with $E(q) = ((q^2-\mu)^2 + |\Delta(q)|^2)^{1/2}$ has a non-trivial solution, that is, one with $\Delta \neq 0$. If this is so, the system is said to be in a superconducting/superfluid state. The function Δ has the interpretation of a spectral gap of an effective mean-field Hamiltonian that is present only in the superconducting/superfluid phase, see the Appendix in [**50**] for further explanations. In [**50**] it has been demonstrated that, although the gap equation is highly non-linear, the question whether there exists a non-trivial solution 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 [**35, 55**] and [**53**], 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 [**56, 17, 52, 37, 40, 25, 36, 58**].

The gap equation in the form stated in Eq. (5.1) and the related BCS functional can be heuristically derived from Quantum Mechanics by a variational procedure under several simplifying assumptions, see [**50**] and the discussion in Section 5.2 below. One of these assumptions is that states used in this variational procedure are translation-invariant which leads to a strong simplification of the model. While this approximation is presumably valid in the case of cold fermionic gases with a rotationally-invariant pair interaction and is of great importance when it comes to numerical computations, it is in general hard to justify its validity. See [**2**] for examples in the context of solid state physics where this approximation is not valid. From a mathematical point of view one is faced with a functional that is invariant under translations in the sense that spatial translations do not change the energy of a state. Due to the non-linear nature of the functional, minimizers need not be translation-invariant, however. If they are not one says that the translational symmetry of the system is broken. The aim of this work is to prove the absence of translational symmetry breaking in two situations: We start by considering 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 persists. Afterwards, we realize that our analysis directly 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 [**57**].

5.2. Main results

We consider a sample of fermionic atoms in a cold gas in d-dimensional space ($d = 2, 3$) within the framework of BCS theory. It is convenient to think of the sample as infinite and periodic, since this setting avoids having to deal with boundary conditions at the boundary of the sample. To describe the periodicity we introduce the lattice $\mathbb{Z}^d$ with the unit cell $[0,1]^d = \Omega$. The special form of the lattice does not play any role for us and the proof carries over to an arbitrary Bravais lattice. To not artificially complicate the presentation, we therefore opt for the simplest choice. BCS states are most conveniently described by their generalized one-particle density matrix, that is, by a self-adjoint operator Γ on $L^2(\mathbb{R}^d) \oplus L^2(\mathbb{R}^d)$ of the form

$$\Gamma = \begin{pmatrix} \gamma & \alpha \\ \overline{\alpha} & 1 - \overline{\gamma} \end{pmatrix}, \tag{5.2}$$

with $0 \leq \Gamma \leq 1$. Here γ and α denote the one-particle density matrix and the Cooper-pair wave function of the state Γ, respectively. Both of them are represented by periodic operators with period one. In terms of kernels, the latter means that $\gamma(x+u, y+u) = \gamma(x,y)$ and $\alpha(x+u, y+u) = \alpha(x,y)$ for all $u \in \mathbb{Z}^d$ and all $x, y \in \mathbb{R}^d$. In (5.2), $\overline{\alpha} = C\alpha C$, where C denotes complex conjugation. Note that, in particular, $\alpha(x,y) = \alpha(y,x)$ for all $x, y \in \mathbb{R}^d$, due to the self-adjointness of Γ. In this setting, it is natural to consider energies per unit volume. Accordingly, we define for a periodic operator A, the trace per unit volume Tr_Ω by $\mathrm{Tr}_\Omega\,[A] = \mathrm{Tr}[\chi_\Omega A \chi_\Omega]$, where χ_Ω denotes the characteristic function of Ω. We call Γ of the form (5.2) an *admissible* BCS state if $\mathrm{Tr}_\Omega\,(-\nabla^2+1)\gamma < \infty$ and denote the set of admissible BCS states by $\mathcal{D}$. We will, by a slight abuse of notation, write $(\gamma, \alpha) \in \mathcal{D}$, meaning that the BCS state Γ given by (5.2) is admissible.

The BCS functional at temperature $T \geq 0$, with chemical potential $\mu \in \mathbb{R}$, interaction potential $V \in L^2(\mathbb{R}^d)$ and entropy

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

is then given by

$$\mathcal{F}(\Gamma) = \mathrm{Tr}_\Omega\,\left[\left(-\nabla^2 - \mu\right)\gamma\right] + \int_{\Omega \times \mathbb{R}^d} V(x-y)|\alpha(x,y)|^2\,\mathrm{d}(x,y) - TS(\Gamma). \tag{5.3}$$

Note that the same functional has been considered in [**37**], where the periodicity was introduced for ease of comparison with the translation-invariant functional. As already mentioned above, the BCS functional can be heuristically derived from Quantum Mechanics by a variational procedure. To that end, one considers the full free energy functional of the system and restricts attention to quasi-free states only. Due to the Wick rule, the energy and the entropy can then be expressed solely in terms of the generalized one-particle density matrix of the quasi-free state under consideration, see [**2**]. If one assumes additionally $SU(2)$-invariance as well as the above periodicity of the state and neglects the direct and the exchange term in the energy, one arrives at Eq. (5.3). For more details see the Appendix of [**50**].

The translation-invariant BCS functional $\mathcal{F}^{\mathrm{ti}}$ is obtained from $\mathcal{F}$ by restricting the set of admissible states to the translation-invariant ones. That is, the kernels of γ and α take the form $\gamma(x,y) = \gamma(x-y)$ and $\alpha(x,y) = \alpha(x-y)$, respectively. We describe translation-invariant BCS states via their momentum representations by 2×2 matrices of the form

$$\hat{\Gamma}(p) = \begin{pmatrix} \hat{\gamma}(p) & \hat{\alpha}(p) \\ \overline{\hat{\alpha}(p)} & 1 - \hat{\gamma}(-p) \end{pmatrix}, \tag{5.4}$$

for $p \in \mathbb{R}^d$, where the bar denotes complex conjugation and the hats indicate that those objects are Fourier transforms of integral kernels that depend only on $x - y$. Obviously, $\hat{\Gamma}(p)$ satisfies $0 \leq \hat{\Gamma}(p) \leq 1$ for all $p \in \mathbb{R}^d$. The latter translates to $|\hat{\alpha}(p)|^2 \leq \hat{\gamma}(p)(1-\hat{\gamma}(p))$ for $p \in \mathbb{R}^d$ in terms of $\hat{\gamma}$ and $\hat{\alpha}$. Note that the fact that Γ is self-adjoint implies that $\hat{\alpha}$ is an even function and that $\hat{\gamma}$ is real-valued. A translation-invariant BCS state Γ is admissible if and only if $\hat{\gamma} \in L^1(\mathbb{R}^d, (1+p^2)\,\mathrm{d}p)$ and $\alpha \in H^1(\mathbb{R}^d, \mathrm{d}x)$.

By $\mathcal{D}^{\mathrm{ti}}$ we denote the set of all admissible translation-invariant BCS states. For $T \geq 0$ the translation-invariant BCS functional with chemical potential $\mu \in \mathbb{R}$, interaction potential $V \in L^2(\mathbb{R}^d)$ and entropy S, which we can now write as

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

takes the form

$$\mathcal{F}^{\mathrm{ti}}(\Gamma) = \int_{\mathbb{R}^2} (p^2 - \mu) \hat{\gamma}(p) \, \mathrm{d}p + \int_{\mathbb{R}^2} V(x) |\alpha(x)|^2 \, \mathrm{d}x - TS(\Gamma). \tag{5.5}$$

Given a state Γ, we define the gap function Δ of that state as the Fourier transform of $2V(x)\alpha(x)$. One can then show that the gap function of any minimizing BCS state satisfies Eq. (5.1), see [**50**]. We note that $\mathcal{F}^{\mathrm{ti}}$ was studied in [**50**] without the constraint that α is reflection symmetric. The results there hold equally if one works only in the subspace of reflection-symmetric functions in $L^2(R^d)$, however. In the case of $V = 0$, the translation-invariant BCS functional $\mathcal{F}^{ti}$ is minimized by the pair $(\gamma_0, 0)$ where $\hat{\gamma}_0(p) = (1 + e^{\beta(p^2 - \mu)})^{-1}$. The same statement is true for the periodic BCS functional $\mathcal{F}$. The state $(\gamma_0, 0)$ is called the normal state and describes a situation where superfluidity is absent.

It was shown in [**50**, Theorem 1] that there exists a critical temperature $T_c \geq 0$ such for $T < T_c$, the minimizer of the translation-invariant BCS functional has a non-vanishing Cooper-pair wave function. On the other hand, for $T \geq T_c$, the normal state is the unique minimizer. Additionally, there is a characterization of T_c in terms of a linear operator. To make this statement more explicit, let us introduce the function $K_T : \mathbb{R}^d \to \mathbb{R}$ given by

$$K_T(p) = \frac{p^2 - \mu}{\tanh((p^2 - \mu)/(2T))}.$$

Then, $K_T = K_T(-i\nabla)$ defines an operator on $L^2(\mathbb{R}^d)$ acting by multiplication with $K_T(p)$ in Fourier space. The critical temperature of the translation-invariant BCS functional is given by

$$T_c = \inf\{T \geq 0 \mid K_T + V \geq 0\}.$$

In other words, T_c is the value of T such that the operator $K_T + V$ has zero as lowest eigenvalue. Observe that this definition makes sense because K_T is monotone increasing in T. The characterization of T_c in terms of a linear operator comes about because a minimizer of the translation-invariant BCS functional $\mathcal{F}^{ti}$ has a non-vanishing Cooper-pair wave function if and only if the normal state is unstable under pair formation. That is, if and only if the second variation of $\mathcal{F}^{ti}$ at $(\gamma_0, 0)$ has a negative eigenvalue. The operator $K_T + V$ is exactly the second variation of $\mathcal{F}^{ti}$ at the normal state in the direction of a perturbation with $\gamma = 0$ and $\alpha \not\equiv 0$.

In this paper, we treat the question whether there is translational symmetry breaking in the BCS model with radial pair interaction V. More precisely, we study the minimization problem

$$\inf \{\mathcal{F}(\Gamma) \mid \Gamma \in \mathcal{D}\}$$

and we are, in particular, concerned with the question whether the infimum of $\mathcal{F}$ is attained by the minimizers of the translation-invariant BCS functional. If $\hat{V} \le 0$ with $\hat{V}$ not identically zero this is already known to be the case, see [**37, 57**]. In order to study this question, we consider the BCS functional $\mathcal{F}_\ell^{\mathrm{ti}}$ on the sector of translation-invariant BCS states with Cooper-pair wave functions of angular momentum $\ell \in 2\mathbb{N}_0$, that we will define in the next paragraph. Our strategy consists of showing that there exists ℓ_0 such that the minimizers of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$ and $\mathcal{F}$ coincide under certain assumptions.

Let us now introduce the functionals $\mathcal{F}_\ell^{\mathrm{ti}}$ in the case $d = 2$. They are obtained from $\mathcal{F}^{\mathrm{ti}}$ by restricting the domain to Cooper-pair wave functions of the form

$$\hat{\alpha}_\ell(p) = e^{i\ell\varphi}\sigma_\ell(p), \tag{5.6}$$

for some $\ell \in 2\mathbb{Z}$, where φ denotes the angle of $p \in \mathbb{R}^2$ in polar coordinates and σ_ℓ is a radial function. Recall that α is an even function, which requires ℓ to be even. As we will see, the Euler-Lagrange equation of $\mathcal{F}^{\mathrm{ti}}$ implies that if (γ, α_ℓ) is a minimizer of $\mathcal{F}^{\mathrm{ti}}$, then $\hat{\gamma}$ has to be a radial function. Therefore, we define the BCS functional on the sector of Cooper-pair wave functions of angular momentum ℓ as follows. We make an angular decomposition for $(p, q) \mapsto \hat{V}(p-q)$, that is

$$\hat{V}(p-q) = \sum_{\ell\in\mathbb{Z}} \hat{V}_\ell(p,q)e^{i\ell\varphi},$$

where φ denotes the angle between p and q. In other words, this means that

$$\hat{V}_\ell(p,q) = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\ell\varphi}\hat{V}(p-q)\,\mathrm{d}\varphi. \tag{5.7}$$

Since $\hat{V}$ is a radial function, it only depends on the absolute value of its argument, that is, on $|p-q| = \sqrt{p^2+q^2-2|p||q|\cos(\varphi)}$ and we conclude that $\hat{V}_\ell$ is radial in both arguments. Furthermore, observe that $\hat{V}_\ell = \hat{V}_{-\ell}$.

Then, the BCS functional $\mathcal{F}_\ell^{\mathrm{ti}}$ on the sector of Cooper-pair wave functions of even angular momentum $\ell \in 2\mathbb{N}_0$ is given by

$$\mathcal{F}_\ell^{\mathrm{ti}}(\Gamma_\ell) = \int_{\mathbb{R}^2}(p^2-\mu)\gamma_\ell(p)\,\mathrm{d}p + \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\overline{\sigma_\ell(p)}\sigma_\ell(q)\hat{V}_\ell(p,q)\,\mathrm{d}p\mathrm{d}q - TS(\Gamma_\ell),$$

where V_ℓ is given in (5.7) and Γ_ℓ is determined by the pair $(\gamma_\ell, \sigma_\ell)$ with radial functions γ_ℓ and σ_ℓ. To be more precise, the domain of $\mathcal{F}_\ell^{\mathrm{ti}}$ is given by

$$\mathcal{D}_\ell := \left\{(\gamma_\ell, \sigma_\ell) \middle|\, \gamma_\ell, \sigma_\ell \text{ radial and } (\gamma_\ell, \alpha_\ell) \in \mathcal{D}^{\mathrm{ti}},\ \hat{\alpha}_\ell(p) = e^{i\ell\varphi}\sigma_\ell(p) \text{ for } p \in \mathbb{R}^2\right\}.$$

Equivalently, $\mathcal{F}_\ell^{\mathrm{ti}}$ can be understood as the restriction of $\mathcal{F}^{\mathrm{ti}}$ to pairs $(\gamma, \alpha) \in \mathcal{D}^{\mathrm{ti}}$ with the property that γ is radial and that α is of the form given in (5.6). In Section 5.3 we will show that $\mathcal{F}_\ell^{\mathrm{ti}}$ has a minimizer.

Next, we characterize the critical temperature $T_c(\ell)$ corresponding to the BCS functionals $\mathcal{F}_\ell^{\mathrm{ti}}$ on the sector of Cooper-pair wave functions of angular momentum $\ell \in 2\mathbb{N}_0$. For this purpose, let us introduce $\mathcal{H} = \{f \in H^1(\mathbb{R}^2, \mathrm{d}p) \mid f \text{ radial }\}$. Then the critical temperature $T_c(\ell)$ of $\mathcal{F}_\ell^{\mathrm{ti}}$ is given by

$$T_c(\ell) := \inf\left\{T \ge 0 \,\middle|\, (K_T + V_\ell)\big|_{\mathcal{H}} \ge 0\right\}. \tag{5.8}$$

The definition of V_ℓ in Eq. (5.7) and the fact that $K_T + V$ commutes with rotations, implies that

$$T_c = \max_{\ell \in 2\mathbb{N}_0} T_c(\ell)$$

holds.

Let us now assume that $T_c = T_c(\ell_0)$ and that the lowest eigenvalue of $K_{T_c} + V$ is at most twice degenerate. In other words, we assume the lowest eigenvalue of $K_{T_c} + V$ to be exactly twice degenerate in the case $\ell_0 \neq 0$ and we assume it to be non-degenerate in the case $\ell_0 = 0$. An exemplary situation satisfying this assumption is illustrated in Figure 1. The meaning of this schematic pictures is the following. Since $T_c = T_c(\ell_0)$, the lowest eigenvalue of $K_T + V$ lies in the sector with angular momentum ℓ_0. If we decrease the temperature this eigenvalue becomes negative and the second/third eigenvalue (depending on the degeneracy) will approach zero at some temperature $\widetilde{T} < T_c(\ell_0)$. For this eigenvalue, there are two possibilities: Either it also lies in the sector of angular momentum ℓ_0, which means that $\widetilde{T} \in (T_c(\ell_1), T_c)$ and this is the case illustrated in Figure 1, or the next eigenvalue lies in the next sector of angular momentum, which means that $\widetilde{T} = T_c(\ell_1)$.

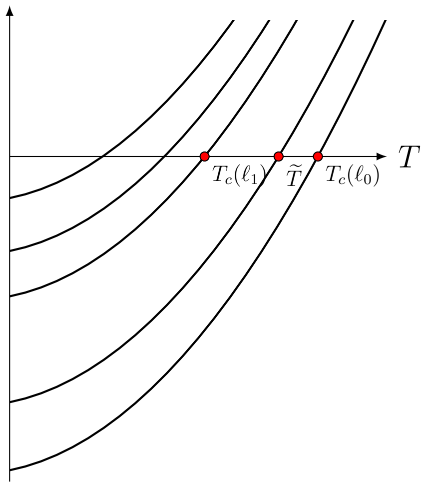

FIGURE 1. Schematic picture of the lowest eigenvalues of $K_T + V$ as a function of the temperature T. The lowest two lines represent eigenvalues in the sector of angular momentum ℓ_0. The third line corresponds to the lowest eigenvalue in the angular momentum ℓ_1 sector. The red dots highlight the temperatures at which one of the eigenvalues crosses the T-axis.

The following theorem shows that the translational symmetry in the BCS model persists if $T \in (\widetilde{T}, T_c)$. In particular, if $\ell_0 = 0$, the periodic (and the translation-invariant) BCS functional has a, up to a phase, unique radial minimizer (γ_0, α_0) for $T \in (\widetilde{T}, T_c)$. If $\ell_0 \neq 0$, the periodic (and the translation-invariant) BCS functional has two minimizers, namely $(\gamma_{\ell_0}, \alpha_{\ell_0})$ and $(\gamma_{\ell_0}, \alpha_{-\ell_0})$, with γ_{ℓ_0} radial and $\alpha_{\pm\ell_0}$ of the form $\hat{\alpha}_{\pm\ell_0}(p) = e^{\pm i\ell_0\varphi}\sigma_{\ell_0}(p)$.

Theorem 5.1. *Let $V \in L^2(\mathbb{R}^2)$ with $\hat{V} \in L^r(\mathbb{R}^2)$, where $r \in [1,2)$, be radial and such that $T_c > 0$. Suppose that $T_c = T_c(\ell_0)$ and that the lowest eigenvalue of $K_{T_c} + V$ is at most twice degenerate. If*

$$(\gamma_{\ell_0}, \sigma_{\ell_0}) \in \mathcal{D}_{\ell_0}$$

minimizes $\mathcal{F}^{\text{ti}}_{\ell_0}$, then there exists $\widetilde{T} < T_c$ such that

$$(\gamma_{\ell_0}, \alpha_{\ell_0}) \text{ and } (\gamma_{\ell_0}, \alpha_{-\ell_0}) \in \mathcal{D}^{\text{ti}},$$

where $\hat{\alpha}_{\pm\ell_0}(p) = e^{\pm i\ell_0\varphi}\sigma_{\ell_0}(p)$, minimize the BCS functional $\mathcal{F}$ for $T \in [\widetilde{T}, T_c)$. For $T \in (\widetilde{T}, T_c)$ these are the only minimizers of $\mathcal{F}$ up to phases in front of α_{ℓ_0} and $\alpha_{-\ell_0}$.

Remark. We want to emphasize that $\widetilde{T}$ is determined by the lowest nonzero eigenvalue of $K_{T_c} + V$. More precisely, $\widetilde{T}$ is given as the value of T such that the second eigenvalue (counted without multiplicities) of $K_T + V$ is zero, which is illustrated in Figure 1. In particular, if in addition to the assumption above, the second eigenvalue of $K_{T_c} + V$ lies in the sector of angular momentum $\ell_1 \neq \ell_0$, one can show that $\widetilde{T} = T_c(\ell_1)$.

Remark. The assumptions $V \in L^2(\mathbb{R}^2)$ and $\hat{V} \in L^r(\mathbb{R}^2)$ with $r \in [1,2)$ in Theorem 5.1 are of technical nature and we expect the Theorem to hold as long as $V \in L^{1+\epsilon}(\mathbb{R}^2)$ for $\epsilon > 0$. Note that this is the L^p regularity for which V is relatively form bounded with respect to the Laplacian in two space dimensions. The assumption on the Fourier transform of V is only needed in the proof of Proposition 5.7. In [**41**, Proposition 5.6] a similar result is proved in the case $d = 3$ under the assumption $V \in L^{3/2}(\mathbb{R}^3)$ which guarantees form boundedness relative to the Laplacian in this case. Although we expect the strategy of that proof to carry over to $d = 2$, our argument is much simpler than the one given in this reference and so we prefer to keep the additional assumption on $\hat{V}$.

Remark. The Fourier transform preserves angular momentum sectors, and hence the inverse Fourier transforms of the minimizing Cooper-pair wave functions $\hat{\alpha}_{\pm\ell_0}(p) = e^{\pm i\ell_0\varphi_p}\sigma_{\ell_0}(p)$ are of the form $e^{\pm i\ell_0\varphi_x} f_{\ell_0}(x)$ with f_{ℓ_0} radial. That is, the Cooper pairs have definite angular momentum also in position space.

Remark. An important step in the proof of Theorem 5.1 is to compare the minimizers of the BCS functional $\mathcal{F}^{\text{ti}}_{\ell_0}$ on the sector of Cooper-pair wave functions with angular momentum ℓ_0 with the minimizers of the periodic BCS functional $\mathcal{F}$. The crucial tool for this comparison will be the relative entropy inequality, [**37**, Lemma 5].

Remark. It is shown in [**41**], amongst other things, that for every $\ell \in 2\mathbb{N}_0$ one can find a radial potential such that the ground state of $K_{T_c} + V$ has angular momentum ℓ. This in particular implies $T_c = T_c(\ell)$ for such a potential. In the case of weak coupling, that is for $K_T + \lambda V$, where $\lambda \in \mathbb{R}$ is small enough, the methods of [**35, 55**] can be applied to determine the angular momentum ℓ_0 of the ground state of $K_{T_c} + V$. An application of these methods reduces the problem of finding the eigenvalues of $K_T + \lambda V$, for λ small enough, to finding the eigenvalues of a simple matrix, that only depends on the behavior of $\hat{V}$ on the Fermi sphere. This is easily solvable numerically.

In particular, one sees, that the eigenvalues are in one-to-one correspondence to the eigenvalues of the matrix $(\langle \psi_n, \hat{V}\psi_m\rangle)_{n,m\geq 0}$, where $\psi_n(p) = e^{in\varphi}$. Moreover, if the lowest eigenvalue of this matrix is at most twice degenerate one is in the situation described in Remark 5.2, i.e. $\widetilde{T} = T_c(\ell_1)$.

REMARK. In the non-interacting case, that is, for $V = 0$, the minimizer of the BCS function $\mathcal{F}$ is given by the normal state

$$\hat{\Gamma}_0 = \begin{pmatrix} \hat{\gamma}_0 & 0 \\ 0 & 1 - \overline{\hat{\gamma}_0} \end{pmatrix},$$

where $\hat{\gamma}_0 = (1 + \exp((-\nabla^2 - \mu)/T))^{-1}$. Let us assume that we are in the situation of Remark 5.2. 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 Γ_0 in the direction of α by

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{F}(\gamma_0, t\alpha)\bigg|_{t=0} = 2\langle \alpha, (K_T + V)\alpha\rangle,$$

one can understand Theorem 5.1 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 one could therefore hope to find a minimizer of $\mathcal{F}$ which lies approximately in this eigenspace. In fact, Theorem 5.1 states that the minimizers of $\mathcal{F}$ for temperatures T in a certain interval below T_c lie in exactly one specific sector of angular momentum $\pm\ell_0$. For $T = T_c(\ell_1)$ the next eigenvalue λ_1 and its eigenspace become important, since now also elements of the eigenspace of λ_1 are candidates to lower the energy.

In the special case $\ell_0 = 0$, Theorem 5.1 also holds in three dimensions.

THEOREM 5.2. *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 such that $T_c > 0$. Assume that zero is a non-degenerate eigenvalue of $K_{T_c} + V$, that is, the corresponding eigenfunction is radial. Then, there exists $\widetilde{T} < T_c$ such that the minimizer of the BCS functional $\mathcal{F}$ for $T \in [\widetilde{T}, T_c)$ is given by a pair (γ_0, α_0), where γ_0 and α_0 are radial functions. Moreover, (γ_0, α_0) is, up to phases, the only minimizer of $\mathcal{F}$ for $T \in (\widetilde{T}, T_c)$.*

REMARK. Note that $\hat{V} \leq 0$ implies that the ground state of $K_{T_c} + V$ is radial in all dimensions. Hence, the assumption that $K_{T_c} + V$ has a non-degenerate lowest eigenvalue is always satisfied for interaction potentials V with this property.

REMARK. As in the case of Theorem 5.1, we expect Theorem 5.2 to hold under the only assumption that V is relatively form bounded with respect to the Laplacian, that is, if $V \in L^{3/2}(\mathbb{R}^3)$.

We recall the gap function $\Delta(p) = 2(2\pi)^{-d/2}\hat{V} * \hat{\alpha}(p)$ with $d = 2, 3$. The Cooper-pair wave function of any minimizer of the translation-invariant BCS functional $\mathcal{F}^{\mathrm{ti}}$ satisfies the Euler-Lagrange equation

$$\left(K_T^\Delta + V\right)\alpha = 0. \tag{5.9}$$

Here K_T^Δ is the operator defined by multiplication in Fourier space with the function

$$K_T^\Delta(p) = \frac{E(p)}{\tanh(E(p)/(2T))}, \qquad \text{where} \qquad E(p) = \sqrt{(p^2-\mu)^2 + |\Delta(p)|^2}.$$

The key ingredient to the proof of Theorem 5.1 and Theorem 5.2 is that in both situations $K_T^\Delta + V \geq 0$ holds. The following Proposition tells us that this already implies that $|\hat{\alpha}(p)|$ is a radial function. Hence, our strategy of proof can only work if this is the case. In particular, it tells us that we cannot extend our results to situations where the absolute value of the Fourier transform of the ground state of $K_{T_c} + V$ is not radial.

PROPOSITION 5.3. *Let V be a radial function with $V \in L^2(\mathbb{R}^2)$ if $d = 2$ and $V \in L^{3/2}(\mathbb{R}^3)$ if $d = 3$. Assume that (γ, α) is a minimizer of the translation-invariant BCS functional $\mathcal{F}^{\rm ti}$ such that $|\hat{\alpha}(p)|$ is not a radial function. Then there exists a rotation $R \in SO(d)$ such that*

$$\left\langle U(R)\alpha, \left(K_T^\Delta + V\right) U(R)\alpha \right\rangle < 0, \tag{5.10}$$

where $(U(R)f)(p) = f(R^{-1}p)$.

5.3. Preparations

The proof of Theorem 5.2 works similarly to the proof of Theorem 5.1. In order to prove Theorem 5.1 we will show that there exists $\ell_0 \in 2\mathbb{N}_0$, such that the minimizers of $\mathcal{F}_{\ell_0}^{\rm ti}$ also minimize $\mathcal{F}$. The following lemma lays the basis for this approach.

In [**50**] it was shown that $\mathcal{F}^{\rm ti}$ is bounded from below and attains its infimum on $\mathcal{D}^{\rm ti}$ in three dimensions. The same results hold in two dimensions by analogous arguments, which provides a solution of the BCS gap equation in this case.

LEMMA 5.4. *The BCS functional $\mathcal{F}_\ell^{\rm ti}$ is bounded from below and attains its minimum.*

PROOF. Boundedness from below of $\mathcal{F}_\ell^{\rm ti}$ follows from the fact that $\mathcal{F}^{\rm ti}$ is bounded from below. As in the proof of [**50**, Lemma 1] we find a minimizing sequence $(\gamma_\ell^{(n)}, \sigma_\ell^{(n)})$ in $\mathcal{D}_\ell$ that converges strongly in $L^p(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ to (γ, σ) for some $p \in (1, \infty)$, as n tends to infinity. It is an easy consequence that $(\gamma, \sigma) \in \mathcal{D}_\ell$. □

The Euler-Lagrange equation of $\mathcal{F}_\ell^{\rm ti}$ takes the same form as the Euler-Lagrange equation of $\mathcal{F}^{\rm ti}$, which will play an important role in the proof. The derivation of the Euler-Lagrange equation of $\mathcal{F}^{\rm ti}$ given in [**57**, Proposition 3.1] translates to the case of $\mathcal{F}_\ell^{\rm ti}$. Therefore, we will not rewrite the proof here, but only give the Euler-Lagrange equation of $\mathcal{F}_\ell^{\rm ti}$ in its various forms.

Let us define the gap function Δ_ℓ related to the Cooper-pair wave function σ_ℓ by

$$\Delta_\ell(p) = \frac{1}{\pi} \int_{\mathbb{R}^2} V_\ell(p,q)\sigma_\ell(p) \mathrm{d}q. \tag{5.11}$$

Since $V_\ell(p,q)$ is radial in both arguments $\Delta_\ell(p)$ is a radial function, too. Also define

$$H_{\Delta_\ell}(p) = \begin{pmatrix} k(p) & \Delta_\ell(p) \\ \overline{\Delta_\ell(p)} & -k(p) \end{pmatrix} \tag{5.12}$$

with $k(p) = p^2 - \mu$. For $T > 0$, the Euler-Lagrange equation of the functional $\mathcal{F}^{\mathrm{ti}}_\ell$, is given by

$$\Gamma_\ell(p) = \begin{pmatrix} \gamma_\ell(p) & \sigma_\ell(p) \\ \overline{\sigma_\ell(p)} & 1 - \gamma_\ell(p) \end{pmatrix} = \frac{1}{1 + e^{H_{\Delta_\ell}(p)/T}}. \tag{5.13}$$

The right-hand side of Eq. (5.13) depends only on σ_ℓ through Δ_ℓ but not on γ_ℓ. That is, γ_ℓ is determined by σ_ℓ.

Let us define $E_\ell(p) = \sqrt{(p^2 - \mu)^2 + |\Delta_\ell(p)|^2}$ and the function $K_T^{\Delta_\ell}$, which for $T > 0$ is given by

$$K_T^{\Delta_\ell}(p) = \frac{E_\ell(p)}{\tanh\left(E_\ell(p)/(2T)\right)}.$$

Then $K_T^{\Delta_\ell} = K_T^{\Delta_\ell}(-i\nabla)$ defines an operator on $L^2(\mathbb{R}^2)$ acting by multiplication with $K_T^{\Delta_\ell}(p)$ in Fourier space. Calculations given explicitly in [**57**] show that (5.13) is equivalent to

$$\gamma_\ell(p) = \frac{1}{2} - \frac{p^2 - \mu}{2K_T^{\Delta_\ell}(p)}, \tag{5.14}$$

$$\sigma_\ell(p) = -\frac{\Delta_\ell(p)}{2K_T^{\Delta_\ell}(p)}. \tag{5.15}$$

Using Eq. (5.11), we see that Eq. (5.15) can be written as

$$\left(K_T^{\Delta_\ell} + V_\ell\right)\sigma_\ell = 0. \tag{5.16}$$

We will also make use of this equation in the form

$$\left(K_T^{\Delta_\ell} + V\right)\alpha_\ell = 0, \tag{5.17}$$

where α_ℓ is of the form (5.6).

5.4. Proof of Theorem 5.1 and Theorem 5.2

We begin with the proof of Theorem 5.1. Let $(\gamma_{\ell_0}, \sigma_{\ell_0}) \in \mathcal{D}_{\ell_0}$ be a minimizer of $\mathcal{F}^{\mathrm{ti}}_{\ell_0}$ and assume $T_c = T_c(\ell_0)$. Let Γ_{ℓ_0} be the BCS state given by the pair $(\gamma_{\ell_0}, \alpha_{\ell_0})$ with $\hat{\alpha}_{\ell_0}(p) = e^{i\ell_0\varphi}\sigma_{\ell_0}(p)$. Our aim is to show that the inequality $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) \geq 0$ holds for all $\Gamma \in \mathcal{D}$. We will use a generalization of the trace per unite volume, which for a periodic operator A on $L^2(\mathbb{R}^2, \mathbb{C}^2)$ is defined by

$$\mathrm{Tr}_0\left[A\right] = \mathrm{Tr}_\Omega\left[P_0 A P_0 + Q_0 A Q_0\right]$$

with

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } Q_0 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that if A is locally trace class, then $\mathrm{Tr}_0\left[A\right] = \mathrm{Tr}_\Omega\left[A\right]$.

We begin by calculating the difference $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell)$, where Γ_ℓ corresponds to a minimizer of $\mathcal{F}_\ell^{\mathrm{ti}}$ as described above. The state Γ is defined by the pair (γ, α). We find

$$\begin{aligned}
&\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) \\
&\quad = \mathrm{Tr}_\Omega \left[\left(-\nabla^2 - \mu\right)(\gamma - \gamma_\ell)\right] \\
&\qquad + \int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y)|^2 - |\alpha_\ell(x,y)|^2\right) \mathrm{d}(x,y) - T\left(S(\Gamma) - S(\Gamma_\ell)\right). \quad (5.18)
\end{aligned}$$

First, we complete the square in the difference of the interaction terms, which yields

$$\begin{aligned}
&\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y)|^2 - |\alpha_\ell(x,y)|^2\right) \mathrm{d}(x,y) \\
&\quad = \int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha(x,y) - \alpha_\ell(x,y)|^2\right) \mathrm{d}(x,y) \\
&\qquad - 2\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(|\alpha_\ell(x,y)|^2 - \mathrm{Re}\left(\overline{\alpha(x,y)}\alpha_\ell(x,y)\right)\right) \mathrm{d}(x,y).
\end{aligned}$$

Next, we combine the second term on the right hand side and the first term on the right hand side of (5.18). Let $\widetilde{\Delta}_\ell(p) = e^{i\ell\varphi}\Delta_\ell(p)$ where φ denotes the angle of $p \in \mathbb{R}^2$ in polar coordinates and Δ_ℓ is given by Eq. (5.11). Inserting the equation $\hat{\alpha}_\ell(p) = -\widetilde{\Delta}_\ell(p)/(2K_T^{\Delta_\ell}(p))$ which follows from Eq. (5.15), we see that

$$\begin{aligned}
&\mathrm{Tr}_\Omega \left[\left(-\nabla^2 - \mu\right)(\gamma - \gamma_\ell)\right] \\
&\quad + 2\,\mathrm{Re}\int_{\Omega\times\mathbb{R}^2} V(x-y)\left(\alpha_\ell(x,y)\overline{\alpha(x,y)} - |\alpha_\ell(x,y)|^2\right) \mathrm{d}(x,y) \\
&\quad = \frac{1}{2}\mathrm{Tr}_0\left[H_{\widetilde{\Delta}_\ell}(\Gamma - \Gamma_\ell)\right].
\end{aligned}$$

Here $H_{\widetilde{\Delta}_\ell}$ is given as in Eq. (5.12) with Δ_ℓ replaced by $\widetilde{\Delta}_\ell$.

At this point, it turns out to be convenient to introduce the relative entropy $\mathcal{H}$, which for two BCS states $\Gamma, \widetilde{\Gamma} \in \mathcal{D}$ is given by

$$\mathcal{H}(\Gamma, \widetilde{\Gamma}) = \mathrm{Tr}_0\left[\Gamma\left(\log\Gamma - \log\widetilde{\Gamma}\right) + (1-\Gamma)\left(\log(1-\Gamma) - \log\left(1-\widetilde{\Gamma}\right)\right)\right].$$

The fact that $H_{\widetilde{\Delta}_\ell}/T = \log(1-\Gamma_\ell) - \log\Gamma_\ell$ yields the following statement.

LEMMA 5.5. *Let $(\gamma_\ell, \sigma_\ell) \in \mathcal{D}_\ell$ be a minimizer of $\mathcal{F}_\ell^{\mathrm{ti}}$ and let Γ_ℓ be given by the pair $(\gamma_\ell, \alpha_\ell)$ where $\alpha_\ell(p) = e^{i\ell\varphi}\sigma_\ell(p)$. Then*

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) = \frac{T}{2}\mathcal{H}(\Gamma, \Gamma_\ell) + \int_{\Omega\times\mathbb{R}^2} V(x-y)|\alpha(x,y) - \alpha_\ell(x,y)|^2\,\mathrm{d}(x,y)$$

for all $\Gamma \in \mathcal{D}$, where $\alpha = (\Gamma)_{12}$.

Based on this identity, we estimate $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0})$ from below by applying the relative entropy inequality [**37, 51**].

Proposition 5.6. *Let $(\gamma_\ell, \sigma_\ell) \in \mathcal{D}_\ell$, be a minimizer of $\mathcal{F}_\ell^{\mathrm{ti}}$, let Γ_ℓ be as in Lemma 5.5 and denote $V_y(x) = V(x-y)$. Then, for all $\Gamma \in \mathcal{D}$, with $\alpha = (\Gamma)_{12}$,*

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) \geq \int_\Omega \left\langle \alpha, \left(K_T^{\Delta_\ell} + V_y(x)\right)_x \alpha \right\rangle_{L^2(\mathbb{R}^2, \mathrm{d}x)} \mathrm{d}y \\ + \mathrm{Tr}_\Omega\, K_T^{\Delta_\ell} (\Gamma - \Gamma_\ell)^2.$$

Here, we understand $(K_T^{\Delta_\ell} + V_y(x))_x$ as an operator acting on the x-coordinate of $\alpha(x,y)$.

Proof. The claimed estimate is a consequence of an inequality for the relative entropy that has been proven in [**37**, Lemma 5]. An application of this inequality yields

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_\ell) \geq \frac{1}{2}\mathrm{Tr}_\Omega \left[(\Gamma - \Gamma_\ell) \frac{H_{\widetilde{\Delta}_\ell}}{\tanh\left(H_{\widetilde{\Delta}_\ell}/(2T)\right)} (\Gamma - \Gamma_\ell) \right] \\ + \int_{\Omega \times \mathbb{R}^2} V(x-y) |\alpha(x,y) - \alpha_\ell(x,y)|^2 \, \mathrm{d}(x,y).$$

The fact that $x \mapsto x(\tanh(x/2))^{-1}$ is an even function and

$$H^2_{\widetilde{\Delta}_\ell}(p) = \mathbb{I}_{\mathbb{C}^2} E^2_\ell(p)$$

is diagonal, implies the statement. □

Next, we show that the operator $K_T^{\Delta_{\ell_0}} + V$ is nonnegative for $T \in [\widetilde{T}, T_c)$.

Proposition 5.7. *Assume $V \in L^2(\mathbb{R}^2)$ and $\hat{V} \in L^r(\mathbb{R}^2)$ for some $r \in [1,2)$. If the lowest eigenvalue of $K_{T_c} + V$ is at most twice degenerate then there exists $\widetilde{T} < T_c$ such that $K_T^{\Delta_{\ell_0}} + V$ is nonnegative as an operator on $L^2(\mathbb{R}^2)$ for all $T \in [\widetilde{T}, T_c)$.*

The proof of Proposition 5.7 is based on spectral perturbation theory and relies on the fact that $K_T^{\Delta_{\ell_0}} + V \to K_{T_c} + V$, while $\Delta_{\ell_0}(T) \to 0$, in norm resolvent sense for $T \to T_c$. We will derive this convergence from the following lemmas. In order to simplify the notation we write $a \lesssim b$ if there exists a constant $c > 0$ such that $a \leq cb$. Moreover, we denote by $\|\cdot\|$ the operator norm and by $\|\cdot\|_r$ the $L^r(\mathbb{R}^2)$-norm.

Lemma 5.8. *Let $T \in (0, T_c)$. The operators $K_{T_c} - K_T$ and $K_T^{\Delta_{\ell_0}} - K_T$ are bounded. More precisely, $\|K_{T_c} - K_T\| \lesssim (T_c - T)$ and $\|K_T^{\Delta_{\ell_0}} - K_T\| \lesssim \|\Delta_{\ell_0}\|_\infty$. Moreover, $K_{T_c} - K_T \geq 0$ and $K_T^{\Delta_{\ell_0}} - K_T \geq 0$.*

Proof. In the proof we abbreviate $A_T := K_{T_c} - K_T$ and $B_T := K_T^{\Delta_{\ell_0}} - K_T$. Notice that

$$K_T^{\Delta_{\ell_0}}(p) = \frac{\sqrt{k(p)^2 + |\Delta_{\ell_0}(p)|^2}}{\tanh\left(\sqrt{k(p)^2 + |\Delta_{\ell_0}(p)|^2}/(2T)\right)}$$

is an increasing function in T for fixed Δ_{ℓ_0} and vice versa. Hence $A_T \geq 0$ 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 $A_T(p)$ and $B_T(p)$ the symbols of A_T and B_T, respectively. In the following we abbreviate $T_c - T = \delta T$ and

$$I_T = \frac{1}{T} - \frac{1}{T_c}.$$

A simple calculation yields

$$A_T(p) = \int_0^1 \frac{I_T k(p)^2}{2\sinh^2\left(k(p)/(2T_c) + tI_T k(p)/2\right)}\,\mathrm{d}t.$$

Obviously, for large $|p|$ the smooth function $A : p \mapsto A(p)$ and all its derivatives have exponential decay. Moreover, $|I_T| \lesssim T_c - T$ implies $\|A_T\| \lesssim T_c - T$. In order to derive an analogous representation for $B_T(p)$ we define

$$f(x) := \frac{\mathrm{d}}{\mathrm{d}x}\frac{x}{\tanh(x/(2T))} = \frac{T\sinh(x/T) - x}{2T\sinh^2(x/(2T))} \tag{5.19}$$

as well as

$$\delta E_{\ell_0}(p) = \sqrt{k(p)^2 + |\Delta_{\ell_0}(p)|^2} - |k(p)|. \tag{5.20}$$

A straightforward calculation shows that

$$B_T(p) = \delta E_{\ell_0}(p) \int_0^1 f(|k(p)| + t\delta E_{\ell_0}(p))\,\mathrm{d}t. \tag{5.21}$$

Since the function f defined in (5.19) is bounded by 1, we find that $|B_T(p)| \le |\delta E_{\ell_0}(p)|$ for all $p \in \mathbb{R}^2$. It can be seen directly from the definition of $\delta E_{\ell_0}(p)$, see (5.20), that $|\delta E_{\ell_0}(p)| \le |\Delta_{\ell_0}(p)|$ for all $p \in \mathbb{R}^2$, which implies $|B_T(p)| \le |\Delta_{\ell_0}(p)|$ for all $p \in \mathbb{R}^2$. □

LEMMA 5.9. *Let $T \in (0, T_c)$. If α_{ℓ_0} is a solution of the BCS gap equation in the form of Eq.* (5.17)*, then* $\|(1+p^2)^{1/4}\hat{\alpha}_{\ell_0}\|_4^4 \lesssim \langle \alpha_{\ell_0}, (K_T^{\Delta_{\ell_0}} - K_T)\alpha_{\ell_0}\rangle$.

PROOF. We will make use of the following observation, which is implied by the fact that the function $|\Delta_{\ell_0}| \mapsto |\Delta_{\ell_0}|/K_T^{\Delta_{\ell_0}}$ is strictly increasing. Eq. (5.11) implies that

$$\|\Delta_{\ell_0}\|_\infty \le \|V\|_2 \, \|\hat{\alpha}_{\ell_0}\|_2 \,. \tag{5.22}$$

We will abbreviate $\|V\|_2 \, \|\hat{\alpha}_{\ell_0}\|_2$ by $c(\alpha_{\ell_0})$ in the following. Thus, together with (5.15), the just mentioned monotonicity of $|\Delta_{\ell_0}|/K_T^{\Delta_{\ell_0}}$ implies that

$$|\hat{\alpha}_{\ell_0}(p)| \le \frac{c(\alpha_{\ell_0})}{2K_T^{c(\alpha_{\ell_0})}(p)}$$

for all $p \in \mathbb{R}^2$. By taking the square and integrating, we see that

$$1 \le \frac{\|V\|_2^2}{4}\int_{\mathbb{R}^2}\left(K_T^{c(\alpha_{\ell_0})}(p)\right)^{-2}\mathrm{d}p.$$

Next, we use that $\tanh(x) \le 1$ for all x, which leads to

$$1 \le \frac{\|V\|_2^2}{4}\int_{\mathbb{R}^2}\left(\left(p^2-\mu\right)^2 + \|V\|_2^2\,\|\hat{\alpha}_{\ell_0}\|_2^2\right)^{-1}\mathrm{d}p$$

We may assume that $\|V\|_2^2\|\alpha_{\ell_0}\|_2^2 \geq \mu^2$ and conclude that

$$1 \leq \frac{\|V\|_2^2}{4} \int_{\mathbb{R}^2} \left(p^4/2 - \mu^2 + \|V\|_2^2 \|\hat{\alpha}_{\ell_0}\|_2^2\right)^{-1} \mathrm{d}p.$$

From this estimate one easily derives that

$$\|\hat{\alpha}_{\ell_0}\|_2^2 \leq \frac{\|V\|_2^2 \pi^4}{32} + \frac{\mu^2}{\|V\|_2^2}.$$

Making use of (5.22), we see that this directly implies that

$$\|\Delta_{\ell_0}\|_\infty^2 \leq \frac{\|V\|_2^4 \pi^4}{32} + \mu^2. \tag{5.23}$$

In other words, there exists a constant $m > 0$ that only depends on V and μ, such that $|\Delta_{\ell_0}(p)| < m$ for all $p \in \mathbb{R}^2$. In particular, m does not depend on T.

We have to estimate $K_T^{\Delta_{\ell_0}} - K_T$ from below. We recall that $|\Delta_{\ell_0}| \mapsto K_T^{\Delta_{\ell_0}}/|\Delta_{\ell_0}|^2$ is decreasing. Having in mind that $K_T^\Delta - K_T$ behaves like $|\Delta|^2$ for small $|\Delta|$ we thus estimate

$$\frac{K_T^{\Delta_{\ell_0}} - K_T}{|\Delta_{\ell_0}|^2}|\Delta_{\ell_0}|^2 \gtrsim \left(\frac{K_T^m - K_T}{m^2}\right)|\Delta_{\ell_0}|^2.$$

Abbreviating $y_t = \sqrt{k(p)^2 + tm^2}/(2T)$ we find that

$$\begin{aligned} K_T^{\Delta_{\ell_0}}(p) - K_T(p) &= 2T \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \frac{y_t}{\tanh(y_t)} \, \mathrm{d}t \\ &= \frac{m^2}{4T} \int_0^1 \left(\frac{1}{y_t \tanh(y_t)} - \frac{1}{\sinh^2(y_t)}\right) \mathrm{d}t. \end{aligned} \tag{5.24}$$

As one easily sees, the function

$$g(y) = \frac{1}{y}\frac{1}{\tanh(y)} - \frac{1}{\sinh^2(y)}$$

is decreasing, which implies

$$K_T^{\Delta_{\ell_0}}(p) - K_T(p) \gtrsim \frac{m^2}{4T}\left(\frac{1}{y_1 \tanh(y_1)} - \frac{1}{\sinh^2(y_1)}\right).$$

Moreover, g is bounded from below by $g(y) \geq 2/3\,(1+y)^{-1}$. Together with (5.24) this shows that

$$K_T^{\Delta_{\ell_0}}(p) - K_T(p) \gtrsim |\Delta_{\ell_0}(p)|^2 \frac{1}{1+p^2}. \tag{5.25}$$

Next, we make use of the Euler-Lagrange equation of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$, that is the relation $|\Delta_{\ell_0}(p)| = 2K_T^{\Delta_{\ell_0}}(p)|\hat{\alpha}_{\ell_0}(p)|$. Inserting this identity in (5.25) we see that

$$K_T^{\Delta_{\ell_0}}(p) - K_T(p) \gtrsim \left(K_T^{\Delta_{\ell_0}}(p)\right)^2 \frac{|\hat{\alpha}_{\ell_0}(p)|^2}{1+p^2} \gtrsim \left(1+p^2\right)|\hat{\alpha}_{\ell_0}(p)|^2,$$

which implies the statement. □

LEMMA 5.10. *Let $T \in (0, T_c)$. If α_{ℓ_0} is a solution of the BCS gap equation in the form* (5.17)*, then $\|\alpha_{\ell_0}\|_2 \lesssim (T_c - T)^{1/2}$. In particular, $\|\Delta_{\ell_0}\|_\infty \lesssim (T_c - T)^{1/2}$.*

PROOF. The gap equation, see (5.17), can be written as

$$\langle \alpha_{\ell_0}, (K_{T_c} + V)\, \alpha_{\ell_0} \rangle + \langle \alpha_{\ell_0}, B\alpha_{\ell_0} \rangle = \langle \alpha_{\ell_0}, A\alpha_{\ell_0} \rangle,$$

where we use the notation introduced in the proof of Lemma 5.8 but drop the subscript, i.e. $A = A_T$ and $B = B_T$ for brevity. Lemma 5.8 and the definition of T_c imply that

$$\langle \alpha_{\ell_0}, B\alpha_{\ell_0} \rangle \le \langle \alpha_{\ell_0}, A\alpha_{\ell_0} \rangle \lesssim (T_c - T)\, \|\alpha_{\ell_0}\|_2^2 \,. \tag{5.26}$$

From the combination of Lemma 5.9 and (5.26) we deduce that

$$\left\| \left(1 + p^2\right)^{1/4} \hat{\alpha}_{\ell_0} \right\|_4^4 \lesssim (T_c - T)\, \|\alpha_{\ell_0}\|_2^2 \,.$$

On the other hand, the $L^r(\mathbb{R}^2)$-norm of $\hat{\alpha}$ is bounded from above by

$$\|\hat{\alpha}_{\ell_0}\|_r \le \left\| \left(1 + p^2\right)^{-1/4} \right\|_s \left\| \left(1 + p^2\right)^{1/4} \hat{\alpha}_{\ell_0} \right\|_4 ,$$

where $r > 2$, due to the fact that we have to choose $s > 4$. Thus,

$$\|\hat{\alpha}_{\ell_0}\|_r^4 \lesssim (T_c - T)\, \|\hat{\alpha}_{\ell_0}\|_2^2 \,. \tag{5.27}$$

Furthermore, we conclude from the relation between Δ_{ℓ_0} and α_{ℓ_0} given by Eq. (5.11) that

$$\|\Delta_{\ell_0}\|_\infty \lesssim \left\| \hat{V} \right\|_t \|\hat{\alpha}_{\ell_0}\|_r \,, \tag{5.28}$$

where we choose $r > 2$ and $t \in [1, 2)$ appropriately. Note that the gap equation in the form (5.15) implies that $\|\hat{\alpha}_{\ell_0}\|_2 \lesssim \|\Delta_{\ell_0}\|_\infty$. Together with (5.27) and (5.28) this finally shows that

$$\|\hat{\alpha}_{\ell_0}\|_2 \lesssim (T_c - T)^{1/4}\, \|\hat{\alpha}_{\ell_0}\|_2^{1/2}$$

and hence proves the first part of the claim. In order to get the estimate on $\|\Delta_{\ell_0}\|_\infty$, we go back to (5.27) and insert $\|\alpha_{\ell_0}\|_2 \lesssim (T_c - T)^{1/2}$. Together with (5.28) this yields the statement. □

Let $T \in (0, T_c)$ and $z \in \mathbb{C} \setminus \mathbb{R}$. Taken together, Lemma 5.8 and Lemma 5.10 show that

$$\begin{aligned} &\left\| (z - (K_{T_c} + V))^{-1} - \left(z - \left(K_T^{\Delta_{\ell_0}} + V\right)\right)^{-1} \right\| \\ &\quad \le \left\| (z - (K_{T_c} + V))^{-1} \right\| \left\| K_T^{\Delta_{\ell_0}} - K_{T_c} \right\| \left\| \left(z - \left(K_T^{\Delta_{\ell_0}} + V\right)\right)^{-1} \right\| \\ &\quad \lesssim |\operatorname{Im}(z)|^{-2}\, (T_c - T)^{1/2} \,. \end{aligned}$$

In other words, $K_T^{\Delta_{\ell_0}} + V \to K_{T_c} + V$ 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 5.7.

PROOF OF PROPOSITION 5.7. We consider the case $\ell_0 \neq 0$. The proof for the case $\ell_0 = 0$ is analogous. As illustrated in Figure 1, we have by assumption that $T_c = T_c(\ell_0)$ and that the lowest eigenvalue of $K_{T_c} + V$ is exactly twice degenerate. Note that in the case that $\ell_0 = 0$ the smallest eigenvalue is non-degenerate. From the convergence of $K_T^{\Delta_{\ell_0}} + V$ to $K_{T_c} + V$ in norm resolvent sense one concludes that the lowest eigenvalue of $K_T^{\Delta_{\ell_0}} + V$ is stable.

In particular, this tells us that there exists $\widetilde{T} < T_c$ such that $K_T^{\Delta_{\ell_0}} + V$ with $T \in (\widetilde{T}, T_c]$ has exactly two eigenvalues $\lambda_1(T), \lambda_2(T) \in \{z \in \mathbb{C}| \ |z| < r\}$ for some radius $r > 0$. Combining this with the fact that the Euler-Lagrange equation (5.17) of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$ reads

$$\left(K_T^{\Delta_{\ell_0}} + V\right) \alpha = 0, \tag{5.29}$$

we conclude that $\lambda_1(T) = \lambda_2(T) = 0$. Having in mind that $K_T^{\Delta_{\ell_0}}$ is an increasing function of T and of Δ_{ℓ_0}, what we have seen by this argument is that the effects of these monotonicity properties exactly correspond. In other words, we have shown that there exists $\widetilde{T} < T_c$ such that $K_T^{\Delta_{\ell_0}} + V$ is nonnegative for all $T \in [\widetilde{T}, T_c]$. It is not hard to see that $\widetilde{T}$ can be chosen as pointed out in Remark 5.2. □

Proof of Theorem 5.1. We know from Lemma 5.4 that for ℓ_0 determined by $T_c(\ell_0) = \max_{\ell \in 2\mathbb{N}} T_c(\ell)$ the functional $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$ has a minimizer $(\gamma_{\ell_0}, \sigma_{\ell_0})$. Proposition 5.6 and Proposition 5.7 show that for Γ_{ℓ_0} given by $(\gamma_{\ell_0}, \alpha_{\ell_0})$, with α_{ℓ_0} as in (5.6),

$$\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) \geq 0,$$

holds for all $\Gamma \in \mathcal{D}$. Moreover, if $\mathcal{F}(\Gamma) - \mathcal{F}(\Gamma_{\ell_0}) = 0$, then $\gamma = \gamma_{\ell_0}$ and $\alpha \in \ker(K_T^{\Delta_{\ell_0}} + V_y)$ by Proposition 5.6. Consequently, α takes the form $\alpha = \psi_1 \alpha_{\ell_0} + \psi_2 \alpha_{-\ell_0}$, where $\alpha_{\pm\ell_0}(p) = e^{\pm i \ell \varphi} \sigma_{\ell_0}(p)$ and ψ_1 and ψ_2 denote complex constants. It remains to show that either $\psi_1 = 0$ and $|\psi_2| = 1$ or $|\psi_1| = 1$ and $\psi_2 = 0$. Observe that, in particular, $(\gamma_{\ell_0}, \alpha) \in \mathcal{D}^{\mathrm{ti}}$ and as we know that $\mathcal{F}^{\mathrm{ti}}$ has a minimizer, we conclude that $(\gamma_{\ell_0}, \alpha)$ satisfies the Euler-Lagrange equation of $\mathcal{F}^{\mathrm{ti}}$, that is

$$\gamma_{\ell_0}(p) = \frac{1}{2} - \frac{p^2 - \mu}{2K_T^{\Delta}(p)},$$

where $\Delta = \pi^{-1} \hat{V} * \hat{\alpha}$. Hence $|\Delta|$ is a radial function and consequently either $\psi_1 = 0$ or $\psi_2 = 0$. In other words, $(\gamma_{\ell_0}, \sigma_{\ell_0}) \in \mathcal{D}_{\ell_0}$. Thus, in order to find minimizers of $\mathcal{F}$, it is sufficient to find the minimizers of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$. As we know that $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$ has minimizers, the only thing left to show is that $(\gamma_{\ell_0}, \sigma_{\ell_0})$ is, up to a phase, the only minimizer of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$. The fact that other possible minimizers $(\gamma_{\ell_0}, \psi\sigma_{\ell_0})$, for some $\psi \in \mathbb{C}$, have to satisfy the gap equation (5.16) of $\mathcal{F}_{\ell_0}^{\mathrm{ti}}$ reads

$$\left(K_T^{\psi\Delta_{\ell_0}} + V_{\ell_0}\right) (\psi\sigma_{\ell_0}) = 0.$$

Together with the monotonicity of $K_T^{\psi\Delta_{\ell_0}}$ in ψ this implies that $|\psi| = 1$. □

The proof of Theorem 5.2 is analogous to the proof of Theorem 5.1 with one exception.

Proof of Theorem 5.2. In case $\ell_0 = 0$ all given arguments also apply in the three-dimensional case. The only exception is Lemma 5.10, where we need to modify the assumptions slightly. One easily sees that $\hat{V} \in L^r(\mathbb{R}^3)$ with $r \in [1, 12/7)$ is a sufficient assumption in this case. □

Proof of Proposition 5.3. We will carry out the proof for $d = 3$ and afterwards comment on the case $d = 2$. The Cooper-pair wave function of any minimizer of the translation-invariant BCS functional satisfies $\hat{\alpha}(p) = -\Delta(p)/(2K_T^\Delta(p))$ which is implied by the Euler-Lagrange equation of $\mathcal{F}$, see **[50]** or compare with Section 5.3. Hence, $|\hat{\alpha}|$ is radial if and only if $|\Delta|$ is radial. With Eq. (5.9) and the assumption that V is a radial function, one checks that it is sufficient to show

$$\left\langle U(R)\alpha, K_T^\Delta U(R)\alpha \right\rangle < \left\langle \alpha, K_T^\Delta \alpha \right\rangle . \tag{5.30}$$

Using the above relation between $\hat{\alpha}$ and Δ, we write

$$\begin{aligned}\left\langle U(R)\alpha, K_T^\Delta U(R)\alpha \right\rangle &= \frac{1}{4}\int_{\mathbb{R}^3} \frac{|\Delta(p)|^2}{K_T^\Delta(p)^2} K_T^\Delta(Rp)\,\mathrm{d}p \\ &= \frac{1}{4}\int_0^\infty \int_{\Omega_r} \frac{|\Delta(p)|^2}{K_T^\Delta(p)^2} K_T^\Delta(Rp)\,\mathrm{d}\omega(p)\, r^2\mathrm{d}r,\end{aligned}$$

where Ω_r denotes the sphere with radius r and $\mathrm{d}\omega(p)$ denotes the uniform measure on Ω_r. On Ω_r, that is for fixed radius $r = |p|$, we can understand $|\Delta(p)|^2/K_T^\Delta(p)^2$ as a function f that depends only on $|\Delta(p)|$. There also exists a function g such that $K_T^\Delta(Rp) = g(|\Delta(Rp)|)$ for all $p \in \Omega_r$. The functions f and g are both strictly increasing.

Consider the expression

$$M(R) := \int_{\Omega_r} [g(\Delta(Rp)) - g(\Delta(p))]\,[f(\Delta(Rp)) - f(\Delta(p))]\,\mathrm{d}\omega(p)$$

The functions f and g depend only on the magnitude of $\Delta(Rp)$, respectively $\Delta(p)$. Since f and g are strictly increasing we have that $M(R) > 0$ unless $|\Delta(Rp)| = |\Delta(p)|$ for a.e. p. To see this assume that $|\Delta(Rp)|$ and $|\Delta(p)|$ differ on a set of positive measure. Now consider the set $\{p : |\Delta(Rp)| > |\Delta(p)|\}$ and the set $\{p : |\Delta(Rp)| < |\Delta(p)|\}$ At least one of them must have positive measure. Hence on the union of these sets

$$[g(\Delta(Rp)) - g(\Delta(p))]\,[f(\Delta(Rp)) - f(\Delta(p))] > 0$$

since f and g are both strictly increasing. Using the rotation invariance of the measure ω, we find

$$\begin{aligned}0 < M(R) = 2\int_{\Omega_r} g(\Delta(p))f(\Delta(p))\,\mathrm{d}\omega(p) &- \int_{\Omega_r} g(\Delta(p))f(\Delta(Rp))\,\mathrm{d}\omega(p) \\ &- \int_{\Omega_r} g(\Delta(Rp))f(\Delta(p))]\,\mathrm{d}\omega(p)\end{aligned}$$

and hence one of the integrals

$$\int_{\Omega_r} g(\Delta(p))f(\Delta(Rp))\,\mathrm{d}\omega(p)$$

or

$$\int_{\Omega_r} g(\Delta(Rp))f(\Delta(p))\,\mathrm{d}\omega(p)$$

must be strictly below

$$\int_{\Omega_r} g(\Delta(p))f(\Delta(p))\,\mathrm{d}\omega(p).$$

Accordingly, there exists a $R \in SO(3)$ such that

$$\int_{\Omega_r} \frac{|\Delta(p)|^2}{K_T^\Delta(p)^2} K_T^\Delta(Rp)\,\mathrm{d}\omega(p) < \int_{\Omega_r} \frac{|\Delta(p)|^2}{K_T^\Delta(p)^2} K_T^\Delta(p)\,\mathrm{d}\omega(p). \tag{5.31}$$

To conclude that Eq. (5.30) holds, it suffices to note that Δ is a continuous function, see the first paragraph in the proof of [**50**, Proposition 3], which implies that both sides of Eq. (5.31) are continuous functions of r. If $d = 2$ the proof goes through in the same way with the only difference that the continuity of Δ is concluded from $\Delta(p) = \pi^{-1}\hat{V} * \hat{\alpha}(p)$, the assumption that $V \in L^2(\mathbb{R}^2)$ and the Riemann-Lebesgue Lemma. □

Acknowledgments. The paper was partially supported by the GRK 1838 and the Humboldt foundation. M.L. was partially supported by NSF grant DMS-1600560. Partial financial support by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 694227) is gratefully acknowledged (A.D.). We are grateful for the hospitality at the Department of Mathematics at the University of Tübingen (M.L.) and at the Georgia Tech School of Mathematics (A.G.).

CHAPTER 6

Derivation of the Ginzburg-Landau equations from the Bogoliubov-de Gennes equations

We show that minimizers of the BCS functional are to leading order given by approximate solutions of the Ginzburg-Landau equations. In contrast to earlier works we treat the magnetic field as an independent variable and also derive the second Ginzburg-Landau equation. In the proof we use a rigorous phase approximation which was recently introduced by Frank, Hainzl and Langmann [**34**].

6.1. Introduction and main result

6.1.1. Introduction. In 1957 Bardeen, Cooper and Schrieffer [**6**] presented their theory of superconductivity. This was the first microscopic explanation of superconductivity starting from a many-body Hamiltonian and a major breakthrough, for which they were awarded the Nobel Prize in 1972. Bardeen, Cooper and Schrieffer realized that the phenomenon of superconductivity can be described by a pairing mechanism. More precisely, superconducting paired states form due to an instability of the normal state in the presence of an attraction between the particles and only at temperatures below a certain critical value.

A breakthrough in the mathematical study of the BSC model was obtained in 2008 in [**50**]. Amongst other results, the authors give a linear criterion that characterizes the so-called critical temperature, which is the temperature below which the system is in a superconducting state. At temperatures equal to or above this critical temperature the system is in the normal state.

Earlier, in 1950 Ginzburg and Landau, see [**46**], had introduced a model of superconductivity which they developed in a phenomenological way, describing the macroscopic properties of a superconductor, without the need to understand the microscopic mechanism. This model has been extremely successful and is widely used in physics, and not only for the description of superconductivity. Due to its rich mathematical structure the theory has inspired the development of many interesting new concepts.

The justification of the macroscopic theory of Ginzburg and Landau (GL) in terms of the microscopic model of Bardeen, Cooper and Schrieffer (BCS) was first treated in the physics literature by Gorkov [**47**] in 1959 who realized, that close to the critical temperature, the model of GL arises from the theory of BCS. In [**45**], de Gennes later simplified Gorkov's arguments.

The first rigorous mathematical results appeared much later. In 2012 Frank, Hainzl, Seiringer and Solovej [**37**] showed that the free energy of the superconductor in the BCS model is given to leading order by the free energy as described by the GL model in the appropriate limit. They also showed that for any near-minimizer of the BCS free energy the corresponding Cooper-pair wave function is described on a macroscopic scale (to leading order) by a near-minimizer of the GL free energy.

Our work here is closely related to [**37**] but with some differences, the most important of which is that we treat the magnetic potential as an independent variable and not as a fixed parameter. This allows us to also derive the second Ginzburg-Landau equation. In order to do this, however, we work with the Euler-Lagrange equations of the energy functional, and therefore assume the existence of minimizers with appropriate scaling behaviour. The rigorous justification of these assumptions will be treated in another paper.

6.1.2. BCS Theory. We consider a superconductor in a 2-dimensional box $Q_h \subset \mathbb{R}^2$, with side length h^{-1} and impose periodic boundary conditions. In BCS theory, the state of the superconductor is described by the generalized one-particle density matrix Γ and the magnetic potential A.

More precisely, we have

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

where, $\overline{\alpha}$ is to be understood as $\overline{\alpha} = C\alpha C$, where C denotes complex conjugation. Furthermore, γ is a self-adjoint operator on $L^2(\mathbb{R}^2)$ and α is an operator on $L^2(\mathbb{R}^2)$ satisfying $\alpha^* = \overline{\alpha}$. On the level of kernels the latter translates to $\alpha(x,y) = \alpha(y,x)$. Furthermore, we have the requirement $0 \leq \Gamma(1-\Gamma) \leq 1$. We also assume that γ and α are periodic operators in the sense that they commute with translations of the lattice, which, in terms of kernels, means that $\alpha(x+h^{-1}t, y+h^{-1}t) = \alpha(x,y)$ for all $t \in \mathbb{Z}^2$, or more generally for any lattice. We assume that $A : \mathbb{R}^2 \to \mathbb{R}^2$ is also periodic, i.e., $A(x+h^{-1}t) = A(x)$ for all $t \in \mathbb{Z}^2$. In the situation described here, it is natural to consider energies per unit volume. Accordingly, we define for a periodic operator A, the trace per unit volume Tr_{Q_h} by $\mathrm{Tr}_{Q_h}[A] = \mathrm{Tr}[\chi_{Q_h} A \chi_{Q_h}]$, where χ_{Q_h} denotes the characteristic function of the box Q_h.

We assume that Γ and A minimize the BCS energy functional at temperature $T \geq 0$,

$$\begin{aligned} \mathcal{F}_T^{\mathrm{BCS}}(\Gamma, A) &= \mathrm{Tr}_{Q_h}\left((-\Delta_A - \mu + W)\,\gamma\right) - T\,\mathrm{Tr}_{Q_h} S(\Gamma) \\ &\quad - \int_{\mathbb{R}^2} \mathrm{d}x \int_{Q_h} \mathrm{d}y\, V(x-y)|\alpha(x,y)|^2 + \int_{Q_h} \mathrm{d}x\, \left|\mathrm{curl}\, A(x) - H_{\mathrm{ext}}\right|^2, \end{aligned}$$

where $-\Delta_A = (-i\nabla + A)^2 = \Delta - 2iA\cdot\nabla - i\,\mathrm{div}\,A + |A|^2$ is the magnetic Laplacian and $\mu \in \mathbb{R}$ is the chemical potential. The entropy function S in the functional is given by $S(x) = -x\log x$. In our notation, the interaction potential is $-V$ and we assume $V : \mathbb{R}^2 \to \mathbb{R}$ to be positive, meaning that $V \geq 0$. Furthermore we require V to be even, that is $V(-x) = V(x)$. Moreover, W is a periodic external potential and

$H_{\text{ext}} : \mathbb{R}^2 \to \mathbb{R}^2$ denotes a periodic applied external magnetic field. From here on we will set $W = 0$ and $H_{\text{ext}} = 0$ for simplicity, since W and H_{ext} have no significant influence on our results.

Minimizers are in particular critical points and must therefore be solutions of the corresponding Euler-Lagrange equations, the so-called Bogoliubov-de Gennes equations, which can be written in the form

$$\frac{1}{2} H_A(-2V\alpha) = TS'(\Gamma), \tag{6.1a}$$

$$\operatorname{curl}^* \operatorname{curl} A + \operatorname{Re}\left(-i\nabla + A\right) \gamma|_{y=x} = 0, \tag{6.1b}$$

where $\operatorname{curl}^* = (\partial_2, -\partial_1)^T$ and $H_A(\Delta)$ is the matrix of operators

$$H_A(\Delta) = \begin{pmatrix} k_A & \Delta \\ \overline{\Delta} & -\overline{k}_A \end{pmatrix},$$

where $k_A := -\Delta_A^2 - \mu$ denotes the kinetic energy. Moreover, $V\alpha$ is the operator whose kernel is $V(x-y)\alpha(x,y)$, i.e., we think of V as a two-body multiplication operator. The notation $|_{y=x}$ denotes the diagonal of the operator.

The translation invariant version of the BCS model is much simpler. In [**50**, Theorem 1], the authors proved the existence of a critical temperature $\mathcal{T}_c \geq 0$ such that for $T < \mathcal{T}_c$, the minimizer of the translation-invariant BCS functional has a non-vanishing Cooper-pair wave function α. On the other hand, for $T \geq \mathcal{T}_c$, the normal state is the unique minimizer of the functional. Additionally, there is a *linear* operator that characterizes $\mathcal{T}_c$. More precisely, let us introduce the function $K_T : \mathbb{R}^2 \to \mathbb{R}$ given by

$$K_T(p) = \frac{p^2 - \mu}{\tanh((p^2 - \mu)/(2T))}.$$

Then, $K_T(-i\nabla)$ defines an operator on $L^2(\mathbb{R}^2)$ acting by multiplication with $K_T(p)$ in Fourier space. The critical temperature of the translation-invariant BCS functional is now given by

$$\mathcal{T}_c = \inf\{T \mid K_T - V \geq 0\}.$$

Put differently, $\mathcal{T}_c$ is the value of T such that the operator $K_T - V$ has zero as lowest eigenvalue.

Now, our assumptions on the potential V are the following.

Assumption 3. *We suppose that*

- *V is even and positive*
- *V is such that $\mathcal{T}_c > 0$ and that α_*, the ground state of $K_{\mathcal{T}_c} - V$, is non-degenerate*
- *$V \in L^\infty(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ and $|x|^2 V \in L^\infty(\mathbb{R}^2)$.*

We note that we can choose α_* to be even and real, which implies, in particular, that the Fourier transform $\hat{V} * \hat{\alpha}_*$ is also real.

6.1.3. Ginzburg-Landau Theory. In Ginzburg-Landau theory the superconductor is described by a complex valued order parameter, that is a function $\psi : \mathbb{R}^2 \to \mathbb{C}$, and the magnetic potential $a : \mathbb{R}^2 \to \mathbb{R}^2$. These functions are again periodic, but here $\psi(x+t) = \psi(x)$ and $a(x+t) = a(x)$ for all $t \in \mathbb{Z}^2$.

We require that ψ and a minimize the Ginzburg-Landau energy functional, which reads

$$\mathcal{E}^{\mathrm{GL}}(\psi, a) := \frac{1}{2}\int_Q \overline{(-i\nabla + a)\psi} \cdot \mathbb{B}(-i\nabla + a)\psi + C_1 W|\psi|^2 - C_1|\psi|^2 + \frac{1}{2}C_2|\psi|^4 + \frac{1}{2}C_3 \left|\operatorname{curl} a - 2H_{\mathrm{ext}}\right|^2,$$

where $Q = Q_1$ is the unit cube. Also, $\mathbb{B}$ is a real symmetric 2×2 matrix, and C_1 , C_2, and C_3 are positive constants. Analogously to the BCS model, for our purpose, we set $W = 0$ and $H_{\mathrm{ext}} = 0$ in the following.

Critical points of the functional $\mathcal{E}^{\mathrm{GL}}$ satisfy the Ginzburg-Landau equations, that is,

$$(-i\nabla + a) \cdot \mathbb{B}\,(-i\nabla + a)\,\psi + C_1 W\psi - C_1\psi + C_2\,|\psi|^2\,\psi = 0 \tag{6.2a}$$

$$C_3 \operatorname{curl}^* \operatorname{curl} a + \operatorname{Re} \overline{\psi}\mathbb{B}\,(-i\nabla + a)\,\psi = 0. \tag{6.2b}$$

6.1.4. Main result. For a periodic operator α on $\mathbb{R}^2$ we define the $\mathcal{L}_h^2$ norm by

$$\|\alpha\|_{\mathcal{L}_h^2}^2 := h^2 \operatorname{Tr}_{Q_h}(\alpha^*\alpha).$$

In our main theorem we need the assumptions

Assumption 4.

- *$T_h = \mathcal{T}_c(1 - Dh^2)$ for some constant D*
- *$\|\alpha_h\|_{\mathcal{L}_h^2} \lesssim h$ and $\|(\nabla_x + \nabla_y)\alpha_h\|_{\mathcal{L}_h^2} \lesssim h^2$*
- *A_h satisfies the gauge conditions $\operatorname{div} A_h = 0$ and $\int_{Q_h} \mathrm{d}x\, A_h(x) = 0$,*
- *A_h is smooth and such that $\|A_h\|_{L^2(Q_h)} \lesssim h$ and*

$$\left\| \,(\operatorname{curl}^*\ \operatorname{curl})^{k/2} A_h \,\right\|_{L^2(Q_h)} \lesssim h^{k+1}$$

for $k = 1, 2, 3, 4$.

to be true for solutions (Γ_h, A_h) of the Euler-Lagrange equations (6.1a) and (6.1b) at temperature $T = T_h$, where $h > 0$. Here and in the following, we write $a \lesssim b$ if $a = O(b)$ as $h \to 0$ (in other words, there exists a positive constant $C > 0$ such that $a \le Cb$ for all sufficiently small h).

Before we finally state the main theorem, let us introduce a notation. We define $\eta : \mathbb{R}^2 \to \mathbb{R}$ to be the function

$$\eta(p) = \int_{\mathbb{R}^2} \mathrm{d}x\, e^{-ip\cdot x} V(x)\alpha_*(x).$$

Our main result is the following. The precise definitions of the norms are given below.

THEOREM 6.1. *Let V be such that it satisfies* Assumptions 3. *Suppose that* (Γ_h, A_h) *is a sequence of solutions of equations* (6.1a) *and* (6.1b) *at temperature* T_h *so that the* Assumptions 4 *are satisfied.*

Then for sufficiently small h we have the decompositions

$$\begin{aligned} V\alpha_h(x,y) &= hV(x-y)\alpha_*(x-y)\psi_h\left(h(x+y)/2\right) + \sigma_h(x,y), \\ A_h(x) &= ha_h(hx) \end{aligned}$$

where $\|\psi_h\|_{H^1(Q)} \lesssim 1$, $\|\sigma_h\|_{\mathcal{L}_h^2} = O(h^2 h^{-17/48})$, *and* $\|a_h\|_{H^2(Q)} \lesssim 1$ *as* $h \to 0$.

Moreover, if (ψ_*, a_*) *is a weak limit point of the sequence* $\{(\psi_h, a_h)\}$ *in* $H^1(Q) \times \vec{H}^1(Q)$, *then* $(\psi_*, 2a_*)$ *is a weak solution of the Ginzburg-Landau equations with the coefficients*

$$\begin{aligned} \mathbb{B}_{ij} = &\int_{\mathbb{R}^2} \mathrm{d}p \, \frac{|\eta(p)|^2}{(2\pi)^2} \left(\frac{\tanh(\beta_c \left(|p|^2-\mu\right)/2)}{4\left(|p|^2-\mu\right)^2} - \frac{\beta_c/8}{\left(|p|^2-\mu\right)\cosh^2\left(\beta_c\left(|p|^2-\mu\right)/2\right)} \right) \delta_{ij} \\ &+ \frac{\beta_c^2}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p \, |\eta(p)|^2 \frac{\tanh(\beta_c\left(|p|^2-\mu\right)/2)}{4\left(|p|^2-\mu\right)\cosh^2\left(\beta_c\left(|p|^2-\mu\right)/2\right)} p_i p_j, \\ C_1 = &\frac{D\beta_c}{2(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p \, \frac{|\eta(p)|^2}{\cosh^2(\beta_c\left(|p|^2-\mu\right)/2)} \\ C_2 = &\int_{\mathbb{R}^2} \mathrm{d}p \, \frac{|\eta(p)|^4}{(2\pi)^2} \left(\frac{2\tanh(\beta_c\left(|p|^2-\mu\right)/2)}{\left(|p|^2-\mu\right)^3} - \frac{\beta_c}{\left(|p|^2-\mu\right)^2\cosh^2\left(\beta_c\left(|p|^2-\mu\right)/2\right)} \right), \end{aligned}$$

and

$$C_3 = \frac{1}{2} - \frac{1}{24\pi\left(1+e^{-\beta_c\mu}\right)}.$$

It also follows that if

$$\limsup_{h\to 0} h^{-1} \|\alpha_h\|_{\mathcal{L}_h^2} \gtrsim 1,$$

then there exists a non-trivial solution of the Ginzburg-Landau equations, i.e., a solution where $\psi \not\equiv 0$.

REMARK. We prove that weak limits $(\psi_*, 2a_*)$ of $\{\psi_h, 2a_h\}$, where we by weak limit mean that $(\psi_{h_n}, a_{h_n}) \rightharpoonup (\psi_*, a_*)$ for some sequence $h_n \to 0$, are weak solutions of the Ginzburg-Landau equation. However, a standard bootstrap argument shows that ψ_* and a_* are in H^2, in fact, they are even smooth, and are indeed strong solutions.

REMARK. The requirement that A_h be smooth can probably be replaced by assuming that A_h is in H^1 and using the equations to show, via a bootstrap argument, that A_h is in fact in H^4 with the correct scaling. We need control up to the H^4 norm of A_h in order to apply the phase approximation method.

6.2. Preliminaries and outline of proof

In this chapter, we will first introduce the norms used in the theorem and the proof, section 6.2.1, and present some preliminaries for the proof, in particular a useful reformulation of the problem, section 6.2.2. After that, we present three important theorems, Theorem A, Theorem B and Theorem C, see sections 6.2.3, 6.2.4 and 6.2.5. Finally, we show how these three theorems imply our main result.

6.2.1. Norms. We first note that we will often work with the relative and center of mass coordinates $r = x - y$ and $X = (x+y)/2$. It is therefore convenient to introduce the notation

$$\zeta_X^r = X + \frac{r}{2}. \tag{6.3}$$

For example we can then work with $\alpha\left(\zeta_X^r, \zeta_X^{-r}\right)$ instead of $\alpha(x,y)$.

We will use a number of norms for periodic operators on $\mathbb{R}^2$. Note that these norms implicitly depend on the parameter h. First we have the $\mathcal{L}_h^p$ norms for α, which are defined by

$$\|\alpha\|_{\mathcal{L}_h^p} := \left(h^2 \operatorname{Tr}_{Q_h}(\alpha^*\alpha)^{p/2}\right)^{1/p}.$$

We also define $\|\cdot\|_\infty$ to be the operator norm. We will often use the fact that these norms satisfy Hölder's inequality, i.e.,

$$\|\alpha\widetilde{\alpha}\|_{\mathcal{L}_h^r} \leq \|\alpha\|_{\mathcal{L}_h^p} \|\widetilde{\alpha}\|_{\mathcal{L}_h^q}, \tag{6.4}$$

whenever $p^{-1}+q^{-1} = r^{-1}$. For the proof of this inequality and a summary of properties of the trace per unit volume we refer the reader to [**37**]. We note that the $\mathcal{L}_h^2$-norm can also be expressed as follows:

$$\|\alpha\|_{\mathcal{L}_h^2}^2 = h^2 \int_{\mathbb{R}^2} \mathrm{d}x \int_{Q_h} \mathrm{d}y \; |\alpha(x,y)|^2 = h^2 \int_{\mathbb{R}^2} \mathrm{d}r \int_{Q_h} \mathrm{d}X \; \left|\alpha\left(\zeta_X^r, \zeta_X^{-r}\right)\right|^2,$$

where the last equality uses the periodicity of α.

The following lemmas will be useful in the proof. For the proofs we refer the reader to the Appendix. In the first lemma we present some estimates on $V^{1/2}\phi$.

Lemma 6.2. *Suppose that $V \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$. Then, we have*

$$\left\|V^{1/2}\phi\right\|_{\mathcal{L}_h^p} \lesssim \|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X\phi\|_{\mathcal{L}_h^2},$$

for all even integers $p \geq 4$. In the case that $V \in L^\infty(\mathbb{R}^2)$ is compactly supported, we have

$$\left\|V^{1/2}\phi\right\|_{\mathcal{L}_h^p} \lesssim h^{-(p-2)/p}\|\phi\|_{\mathcal{L}_h^2} \quad \text{and} \quad \left\|V^{1/2}\phi\right\|_\infty \lesssim h^{-1}\|\phi\|_{\mathcal{L}_h^2},$$

again for all even integers $p \geq 4$.

Furthermore, if $\phi(x,y) = \varphi(x-y)\Psi((x+y)/2)$, then

$$\left\|V^{1/2}\phi\right\|_\infty \lesssim \|\varphi\|_{L^2}\|\Psi\|_{L^\infty}.$$

It will also be useful to know the following properties about kernels.

Lemma 6.3. *Suppose that $K : \mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{C}$ satisfies $|K(X+r/2, X-r/2)| \leq g(r)$, where $g \in L^1(\mathbb{R}^2)$. Then K defines a bounded operator on $L^2(\mathbb{R}^2)$ satisfying $\|K\|_\infty \leq \|g\|_{L^1}$. Similarly, if $K_j : \mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{C}$, $j = 1,2$, satisfies $|K_j(X+r/2, X-r/2)| \leq g_j(r)$, where $g_j \in L^1(\mathbb{R}^2)$. Then $|(K_1K_2)(X+r/2, X-r/2)| \leq (g_1 * g_2)(r)$.*

The estimate present in the next lemma will be of great use in the proof of our main theorem.

LEMMA 6.4. *Suppose that $K : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}$ defines a $h^{-1}\mathbb{Z}^2$-periodic operator. Then*

$$\left\| K|_{x=y} \right\|_{L^2(Q_h)} \lesssim \| K(1-\Delta) \|_{\mathcal{L}^2_h},$$

where $|_{x=y}$ indicates, that we consider the diagonal of the operator.

When doing estimates using the phase approximation method, it will be necessary to have L^∞ bounds on A and its derivatives. First we define the Fourier transform for $h^{-1}\mathbb{Z}^2$-periodic functions f to be the function $\hat{f}$ on $(2\pi\mathbb{Z})^2$ given by

$$\hat{f}(m) = h^2 \int_{Q_h} \mathrm{d}x\, e^{-ihm\cdot x} f(x).$$

Note, that this means

$$f(x) = \sum_{m \in (2\pi\mathbb{Z})^2} e^{ihm\cdot x} \hat{f}(m).$$

Turning to A, we note that the gauge conditions on A ensure that $\hat{A}(0) = 0$ and that $\operatorname{curl}^2 A = -\Delta A$. This means

$$|A(x)| \le \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0}} \left|\hat{A}(m)\right| \le \left(\sum_{m \neq 0} |m|^{-4} \right)^{1/2} \left(\sum_{m \neq 0} |m|^4 \left|\hat{A}(m)\right|^2 \right)^{1/2}$$
$$\lesssim h^{-2} \left\| \operatorname{curl}^2 A \right\|_{L^2(Q_h)}.$$

From this it follows that we have the bounds

$$\left\| \operatorname{curl}^k A \right\|_{L^\infty} \lesssim h^{-2} \left\| \operatorname{curl}^{k+2} A \right\|_{L^2(Q_h)} \lesssim h^{k+1},$$

for solutions A_h.

6.2.2. Reformulation. It will be convenient to reformulate the problem in terms of $\phi = V^{1/2}\alpha$. This stems from applying the Birman-Schwinger principle. We begin by noting that one can invert (6.1a) and write

$$\Gamma = \frac{1}{1 + e^{\beta H_A(-2V\alpha)}},$$

but as the right-hand side is a function of α and A alone, we see that γ can be seen as a function of these two variables, i.e.,

$$\gamma = \gamma_A(-2V\alpha) = \left(\frac{1}{1 + e^{\beta H_A(-2V\alpha)}} \right)_{11},$$

where the index $_{ij}$ denotes the ij-entry of the matrix. This means that it suffices to consider the equation

$$\alpha = \left(\frac{1}{1 + e^{\beta H_A(-2V\alpha)}} \right)_{12}. \tag{6.5}$$

Now using the facts that

$$\frac{1}{1+e^x} + \frac{1}{1+e^{-x}} = 1 \quad \text{and} \quad \frac{1}{1+e^x} - \frac{1}{1+e^{-x}} = -\tanh\left(\frac{x}{2}\right),$$

we see that

$$\left(\frac{1}{1+e^{\beta H_A(-2V\alpha)}}\right)_{12} = \frac{1}{2}\left(\frac{1}{1+e^{\beta H_A(-2V\alpha)}} - \frac{1}{1+e^{-\beta H_A(-2V\alpha)}} + 1\right)_{12}$$
$$= -\frac{1}{2}\left(\tanh\frac{\beta}{2}H_A(-2V\alpha)\right)_{12}.$$

We now rewrite the Bogoliubov-de Gennes equations as equations for $\phi = V^{1/2}\alpha$. If α is a solution of (6.5), then ϕ solves the equation

$$\phi + \frac{1}{2}V^{1/2}\left(\tanh\frac{\beta}{2}H_A\left(-2V^{1/2}\phi\right)\right)_{12} = 0, \tag{6.6}$$

Conversely, if ϕ solves (6.6) and we define

$$\alpha = -\frac{1}{2}\left(\tanh\frac{\beta}{2}H_A\left(-2V^{1/2}\phi\right)\right)_{12},$$

then

$$-2V\alpha = 2V^{1/2}V^{1/2}\frac{1}{2}\left(\tanh\frac{\beta}{2}H_A\left(-2V^{1/2}\phi\right)\right)_{12} = -2V^{1/2}\phi,$$

which means α solves (6.5).

This allows us to write the equations we are treating in the following form.

$$F_T^{\mathrm{BCS}}(\phi, A) := \phi + \frac{1}{2}V^{1/2}\left(\tanh\frac{\beta}{2}H_A\left(-2V^{1/2}\phi\right)\right)_{12} = 0,$$
$$G_T^{\mathrm{BCS}}(\phi, A) := \operatorname{curl}^*\operatorname{curl} A + J_{T,A}\left(-2V^{1/2}\phi\right) = 0,$$

where $J_{T,A}(\phi) = \operatorname{Re}\pi_A\gamma_A(\phi)|_{y=x}$ and $\pi_A = (-i\nabla + A)$.

6.2.3. Decomposition of ϕ and Theorem A. We now introduce a decomposition of ϕ in terms of a projection on $\mathcal{L}_h^2$. For convenience we define $\varphi_* = V^{1/2}\alpha_*$ and normalize α_* so that

$$\int_{\mathbb{R}^2} \mathrm{d}r\ |\varphi_*(r)|^2 = \int_{\mathbb{R}^2} \mathrm{d}r\, V(r)\,|\alpha_*(r)|^2 = 1.$$

Recall the notation introduced in (6.3). We now define a projection $P : \mathcal{L}_h^2 \to \mathcal{L}_h^2$ by

$$P\phi\left(\zeta_X^r, \zeta_X^{-r}\right) = \varphi_*(r)\int_{\mathbb{R}^2} \mathrm{d}s\,\varphi_*(s)\phi\left(\zeta_X^s, \zeta_X^{-s}\right). \tag{6.8}$$

Note that P simply projects onto φ_* in the $r = x - y$ direction. We let $P^\perp = 1 - P$, and we use this notation for all projections.

Besides this projection, we will also need a momentum cut-off in the center-of-mass coordinate. The abbreviation $\varepsilon = h^{17/48}$ turns out to be convenient here. Furthermore, we define the projection χ_ε on $\mathcal{L}_h^2$ by

$$\chi_\varepsilon\phi\left(\zeta_X^r, \zeta_X^{-r}\right) = h^{-2}\sum_{m\in(2\pi\mathbb{Z})^2} e^{im\cdot hX}\chi\left(h|m| \le \varepsilon\right)\int_{Q_h} \mathrm{d}Y\, e^{-im\cdot hY}\phi\left(\zeta_Y^r, \zeta_Y^{-r}\right),$$

where, as one would expect, $\chi(h|m| \le \varepsilon) = 1$ if $h|m| \le \varepsilon$ and 0 otherwise. We now define $P_\varepsilon : \mathcal{L}_h^2 \to \mathcal{L}_h^2$ to be the projection

$$P_\varepsilon = \chi_\varepsilon P. \tag{6.9}$$

Note that we have

$$P_\varepsilon^\perp = 1 - \chi_\varepsilon P = P^\perp + P - \chi_\varepsilon P = P^\perp + \chi_\varepsilon^\perp P,$$

i.e., $P_\varepsilon^\perp$ is the sum of a projection onto the subspace orthogonal to φ_* in the relative coordinate and a projection onto the high center-of-mass momenta in the direction of φ_*

Our first main goal will be the proof of the following theorem, which shows that the solution ϕ_h is given to leading order by $P_\varepsilon \phi_h$.

THEOREM A. *Suppose that ϕ_h, A_h, T_h satisfy Assumption 4 with α_h replaced by ϕ_h and that $F^{\mathrm{BCS}}_{T_h}(\phi_h, A_h) = 0$ and $G^{\mathrm{BCS}}_{T_h}(\phi_h, A_h) = 0$. Then*

$$\|\phi_h - P_\varepsilon \phi_h\|_{\mathcal{L}^2_h} \lesssim h^3 \varepsilon^{-2}.$$

Moreover,

$$\left\| P_\varepsilon F^{\mathrm{BCS}}_{T_h}(P_\varepsilon \phi_h, A_h) \right\|_{\mathcal{L}^2_h} \lesssim h^3 \varepsilon^3 h^{-1},$$

and

$$\left\| G^{\mathrm{BCS}}_{T_h}(P_\varepsilon \phi_h, A_h) \right\|_{L^2(Q_h)} \lesssim h^3 \varepsilon^{-1} h^{1/2} \left(-\frac{h}{\varepsilon} \log \frac{h}{\varepsilon} \right)^{1/2},$$

where $\varepsilon = h^{17/48}$.

6.2.4. Ginzburg-Landau equations and Theorem B. The Ginzburg-Landau equations are defined for $\psi : \mathbb{R}^2 \to \mathbb{C}$ and $a : \mathbb{R}^2 \to \mathbb{R}^2$ that are $\mathbb{Z}^2$-periodic, i.e., ψ and a are in particular not $h^{-1}\mathbb{Z}^2$-periodic like ϕ. We therefore introduce the following rescaled projections. First we have $\mathcal{Q}_h : \operatorname{range} P \to L^2(Q)$ defined by the condition

$$P\phi(x,y) = h\varphi_*(x-y) \left(\mathcal{Q}_h P\phi\right)(h(x+y)/2),$$

which means

$$\left(\mathcal{Q}_h P\phi\right)(X) = h^{-1} \int_{\mathbb{R}^2} \mathrm{d}s\, \varphi_*(s) \phi\left(\zeta^s_{X/h}, \zeta^{-s}_{X/h}\right).$$

We note that

$$\|\mathcal{Q}_h\|_{\mathcal{L}^2_h \to L^2(Q)} = h^{-1}.$$

We also define $\vec{\mathcal{Q}}_h : L^2(Q_h) \to L^2(Q)$ by

$$\left(\vec{\mathcal{Q}}_h A\right)(X) = h^{-1} A\left(X/h\right).$$

We again have

$$\left\| \vec{\mathcal{Q}}_h \right\|_{L^2(Q_h) \to L^2(Q)} = h^{-1}.$$

For convenience, we introduce the following notation for (6.2a) and (6.2b).

$$F^{\mathrm{GL}}(\psi, a) = (-i\nabla + a) \cdot \mathbb{B} \left(-i\nabla + a\right) \psi + C_1 W \psi - C_1 \psi + C_2 \left|\psi\right|^2 \psi,$$

$$G^{\mathrm{GL}}(\psi, a) = C_3 \operatorname{curl}^* \operatorname{curl} a + \operatorname{Re} \overline{\psi} \mathbb{B} \left(-i\nabla + a\right) \psi$$

THEOREM B. *Suppose that ϕ_h, A_h, T_h satisfy Assumption 4 with α_h replaced by ϕ_h. Let $\psi_h = \mathcal{Q}_h P_\varepsilon \phi_h$ and $a_h = \vec{\mathcal{Q}}_h A_h$. Then*

$$\left\| h^{-2}\mathcal{Q}_h P_\varepsilon F^{\mathrm{BCS}}_{T_h}\left(P_\varepsilon\phi_h, A_h\right) - F^{\mathrm{GL}}\left(\psi_h, 2a_h\right)\right\|_{L^2(Q)} \lesssim \varepsilon^3 h^{-1},$$

and

$$\left\| h^{-2}\vec{\mathcal{Q}}_h G^{\mathrm{BCS}}_{T_h}\left(P_\varepsilon\phi_h, A_h\right) - G^{\mathrm{GL}}\left(\psi_h, 2a_h\right)\right\|_{L^2(Q)} \lesssim \varepsilon^{1/2} h^{1/2},$$

where $\varepsilon = h^{17/48}$.

6.2.5. Exact solutions of the Ginzburg-Landau equation, Theorem C. We have the following theorem.

THEOREM C. *Suppose that $\psi_n \rightharpoonup \psi_*$ in $H^1(Q)$ and that $a_n \rightharpoonup a_*$ in $H^1(Q)$. If it is also true that*

$$\left\| F^{\mathrm{GL}}\left(\psi_n, 2a_n\right)\right\|_{L^2(Q)} \to 0 \quad \textit{and} \quad \left\| G^{\mathrm{GL}}\left(\psi_n, 2a_n\right)\right\|_{L^2(Q)} \to 0,$$

then $F^{\mathrm{GL}}(\psi_*, 2a_*) = 0$ and $G^{\mathrm{GL}}(\psi_*, 2a_*) = 0$ in the weak sense.

6.2.6. Proof of Theorem 6.1. Note that the first part of Theorem 6.1, that is the decomposition

$$V\alpha_h(x,y) = hV(x-y)\alpha_*(x-y)\psi_h\left(h(x+y)/2\right) + \sigma_h(x,y),$$

while $A_h(x) = ha_h(hx)$ and the estimates $\|\psi_h\|_{H^1(Q)} \lesssim 1$, $\|\sigma_h\|_{\mathcal{L}^2_h} = O(h^2 h^{-17/48})$ as $h \to 0$, and $\|a_h\|_{H^2(Q)} \lesssim 1$, stated in Theorem 6.1 follow directly from Theorem A. Let us mention here, that $\sigma_h = V^{1/2} P_\varepsilon^\perp \phi_h$. From Theorem B follow the explicit forms of the coefficients. Finally, Theorem C implies that $(\psi_*, 2a_*)$ is a weak solution of the Ginzburg-Landau equations. □

6.3. Phase approximation

One of the main technical tools we will use is the phase approximation method for the resolvent of k_A, recall that $k_A = -\Delta_A^2 - \mu$, and in this section we develop this method for smooth magnetic potentials $A : \mathbb{R}^2 \to \mathbb{R}^2$ that are $h^{-1}\mathbb{Z}^2$-periodic and divergence-free. We also assume that $z \in \mathbb{C}$ is always such that $\operatorname{Im} z \neq 0$.

6.3.1. Basic definitions. We first define $G_A^z : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}$ to be such that $G_A^z(x,y)$ is the kernel of $(z - k_A)^{-1}$. Since

$$\left(z + \overline{k}_A\right)^{-1} = -(-z - k_{-A})^{-1},$$

it follows that $-G_{-A}^{-z}(x,y)$ is the kernel of $\left(z + \overline{k}_A\right)^{-1}$. We also define $G^z : \mathbb{R}^2 \to \mathbb{C}$ is such that $G^z(x-y)$ is the kernel of $(z - k_0)^{-1}$. We want to mention here that $(z-k_0)^{-1}$ commutes with translations and therefore its kernel is a function of $x - y$. As before, it then follows that $-G^{-z}(x-y)$ is the kernel of $(z + k_0)^{-1}$. We note that G^z is a radial function and has the following decay property, see **[1]**,

$$\left\| |x|^s G^z \right\|_{L^1(\mathbb{R}^2)} \le C_s \left|\operatorname{Im} z\right|^{-1} \tag{6.11}$$

for all $s \in \mathbb{N}_0$, where C_s is an explicit constant depending on s.

We now define the phase $\Phi_A : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ to be the line integral

$$\Phi_A(x,y) = -\int_y^x A(u) \cdot \mathrm{d}u := -\int_0^1 \mathrm{d}t\, A(y + t(x-y)) \cdot (x-y). \tag{6.12}$$

We can now define the kernel $K_A^z : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}$ to be

$$K_A^z(x,y) = e^{i\Phi_A(x,y)} G^z(x-y). \tag{6.13}$$

The obvious bound $|K_A^z\left(\zeta_X^r, \zeta_X^{-r}\right)| \le |G^z(r)|$ implies that K_z^A defines a bounded operator on $L^2(\mathbb{R}^2)$ and, moreover, (6.11) together with Young's inequality shows that

$$\|K_A^z\|_\infty \lesssim |\mathrm{Im}\, z|^{-1}.$$

We now define $\widetilde{A} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ to be

$$\widetilde{A}(x,y) = \int_0^1 \mathrm{d}t\, t \,\mathrm{curl}\, A\left(y + t\left(x-y\right)\right)\left(x-y\right)^\perp. \tag{6.14}$$

Moreover, we introduce the kernel $T_A^z : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}$ to be

$$T_A^z(x,y) = e^{i\Phi_A(x,y)} G^z(x-y) \left(i\, \mathrm{div}_x\, \widetilde{A}(x,y) - \left|\widetilde{A}(x,y)\right|^2 \right).$$

and define $h_A^z : \mathbb{R}^2 \to \mathbb{R}$ by

$$h^z(r) = \left(|r| + |r|^2\right) |G^z(r)|.$$

One can then show that $|T_A^z\left(\zeta_X^r, \zeta_X^{-r}\right)| \lesssim M_A h^z(r)$, and this leads to the bound

$$\|T_A^z\|_\infty \lesssim M_A \,|\mathrm{Im}\, z|^{-1},$$

where

$$M_A := \max\left\{ \left\|\mathrm{curl}^2 A\right\|_{L^\infty(\mathbb{R}^2)}, \|\mathrm{curl}\, A\|^2_{L^\infty(\mathbb{R}^2)} \right\}. \tag{6.15}$$

The main lemma for the approximation of G_A^z is based on the work of G. Nenciu, see [**76**].

Lemma 6.5. *We have that*

$$(z - k_A)\, K_A^z = 1 + T_A^z.$$

Therefore,

$$(z-k_A)^{-1} = K_A^z\left(1 + T_A^z\right)^{-1} = \sum_{k=0}^\infty (-1)^k K_A^z \left(T_A^z\right)^k,$$

if $\|T_A^z\| < 1$.

Proof. This lemma will follow from a number of calculations. Note that, for ease of notation, we think of all our two-dimensional vectors as three-dimensional by adding a zero-entry as third component. We first use the fact that

$$\nabla(X \cdot Y) = (X \cdot \nabla)Y + (Y \cdot \nabla)X + X \wedge \mathrm{curl}\, Y + Y \wedge \mathrm{curl}\, X.$$

With $X = \int_0^1 \mathrm{d}t\, A(y + t(x-y))$ and $Y = x - y$, we have that $\nabla Y = 0$ and see that

$$\nabla_x \int_y^x A(u) \cdot \mathrm{d}u = \int_0^1 \mathrm{d}t\, A\left(y + t(x-y)\right) + ((x-y)\cdot\nabla) \int_0^1 \mathrm{d}t\, A(y + t(x-y)) \\ + (x-y) \wedge \int_0^1 \mathrm{d}t\, t \,\mathrm{curl}\, A(y + t(x-y)).$$

Now for the second term we simply use integration by parts to see that

$$((x-y)\cdot\nabla) \int_0^1 \mathrm{d}t\, A(y + t(x-y)) = \int_0^1 \mathrm{d}t\, t\, \frac{\mathrm{d}}{\mathrm{d}t} A(y + t(x-y)) \\ = A(x) - \int_0^1 \mathrm{d}t\, A(y + t(x-y)).$$

Recalling the notations introduced in (6.12) and (6.14), we find that

$$\nabla_x \Phi_A(x,y) = -A(x) + \widetilde{A}(x,y).$$

It now follows that for $u : \mathbb{R}^2 \to \mathbb{C}$, we have

$$(-i\nabla + A)^2 e^{i\Phi_A(x,y)} u(x) = (-i\nabla + A) \cdot e^{i\Phi_A(x,y)} \left(-i\nabla u(x) + \widetilde{A}(x,y) u(x)\right) \\ = e^{i\Phi_A(x,y)} \left(\left(-i\nabla_x + \widetilde{A}(x,y)\right) \cdot \left(-i\nabla u(x) + \widetilde{A}(x,y)u(x)\right)\right),$$

and by expanding, we get that

$$(-i\nabla + A)^2 e^{i\Phi_A(x,y)} u(x) \\ = e^{i\Phi_A(x,y)} \left(-\Delta u(x) - 2i\widetilde{A}(x,y)\cdot\nabla u(x) - i\,\mathrm{div}_x\, \widetilde{A}(x,y) u(x) + \left|\widetilde{A}(x,y)\right|^2 u(x)\right).$$

Note that $\widetilde{A}(x,y)$ is orthogonal to $x - y$, and therefore $\widetilde{A}(x,y)\cdot\nabla G^z(x-y) = 0$, since G^z is a radial function. By definition

$$(z - k_A)\, K_A^z u(x) = (z - (-i\nabla + A)^2 + \mu) \int_{\mathbb{R}^2} \mathrm{d}y\, e^{i\Phi_A(x,y)} G^z(x-y) u(y)$$

and hence

$$(z - k_A)\, K_A^z u(x) \\ = \int_{\mathbb{R}^2} \mathrm{d}y\, e^{i\Phi_A(x,y)} \Big(z + \Delta G^z(x-y) + i\,\mathrm{div}_x\, \widetilde{A}(x,y) G^z(x-y) \\ - \left|\widetilde{A}(x,y)\right|^2 G^z(x-y) + \mu G^z(x-y)\Big)\, u(y) \\ = \int_{\mathbb{R}^2} \mathrm{d}y\, e^{i\Phi_A(x,y)} \delta(x-y) u(y) \\ + \int_{\mathbb{R}^2} \mathrm{d}y\, e^{i\Phi_A(x,y)} \left(i\,\mathrm{div}_x\, \widetilde{A}(x,y) - |\widetilde{A}(x,y)|^2\right) G^z(x-y) u(y) \\ = u(x) + \int_{\mathbb{R}^2} \mathrm{d}y\, T_A^z(x,y) u(y).$$

This proves the first part of the lemma, and the remaining claim is clear. □

Corollary 6.6. *If M_A is sufficiently small, then the function*

$$H_A^z(r) := \sum_{k=1}^{\infty}(-1)^k M_A^k \left(G^z * h^z * \cdots * h^z\right)(r)$$

is well defined and satisfies

$$\|H_A^z\|_{L^1(\mathbb{R}^2)} \leq \sum_{k=1}^{\infty} M_A^k \|G^z\|_{L^1(\mathbb{R}^2)} \|h^z\|_{L^1(\mathbb{R}^2)}^k = \frac{\|G^z\|_{L^1(\mathbb{R}^2)}}{1 - M_A \|h^z\|_{L^1(\mathbb{R}^2)}} M_A \|h^z\|_{L^1(\mathbb{R}^2)} .$$

Furthermore, we have the estimate

$$\left|G_A^z\left(\zeta_X^r, \zeta_X^{-r}\right) - K_A^z\left(\zeta_X^r, \zeta_X^{-r}\right)\right| \leq H_A^z(r)$$

for all X and r.

Proof. We begin by writing

$$G_A^z(x,y) - K_A^z(x,y) = \left(K_A^z \sum_{k=1}^{\infty}(-1)^k (T_A^z)^k\right)(x,y).$$

First, note that $\left|(T_A^z)^k\left(\zeta_X^r, \zeta_X^{-r}\right)\right| \leq M_A^k (h^z * \cdots * h^z)(r)$. Furthermore, since h^z is a bounded function, we have by Young's inequality that

$$\left(h^z * \cdots * h^z\right)(r) \leq \|h^z\|_{L^1}^{k-1} \|h^z\|_{L^\infty} .$$

This means that if M_A is sufficiently small, the function

$$r \mapsto \sum_{k=1}^{\infty}(-1)^k M_A^k (h^z * \cdots * h^z)(r)$$

is well defined, and the same holds for H_A^z. The estimate for the L^1-norm of H_A^z also follows from Young's inequality. Altogether, we then have

$$\left|G_A^z\left(\zeta_X^r, \zeta_X^{-r}\right) - K_A^z\left(\zeta_X^r, \zeta_X^{-r}\right)\right| \leq H_A^z(r).$$

This proves the corollary. □

We will also make use of the following result concerning Φ_A.

Lemma 6.7. *We have*

$$\left|\Phi_A\left(\zeta_X^r, \zeta_{X+Y}^s\right) - \Phi_A\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right) - \Phi_{2A}\left(\zeta_X, \zeta_{X+Y}\right) + \frac{1}{4}(r-s)\cdot DA(X)(r+s)\right| \lesssim \left\|D^2 A\right\|_{L^\infty}\left(|s|^2 + |r|^2\right),$$

where DA denote the Jacobian matrix of A, i.e., $(DA)_{jk} = \partial_j A_k$.

Proof. We first see that

$$\begin{aligned}\Phi_A\left(\zeta_X^r, \zeta_{X+Y}^s\right) - \Phi_A\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right) &= \Phi_A\left(\zeta_X^r, \zeta_{X+Y}^s\right) + \Phi_A\left(\zeta_X^{-r}, \zeta_{X+Y}^{-s}\right) \\ &= -\int_0^1 \mathrm{d}t\, A\left(\zeta_{X+Y}^s + t\zeta_{-Y}^{r-s}\right)\cdot \zeta_{-Y}^{r-s} - \int_0^1 \mathrm{d}t\, A\left(\zeta_{X+Y}^{-s} + t\zeta_{-Y}^{-(r-s)}\right)\cdot \zeta_{-Y}^{-(r-s)}.\end{aligned}$$

Sorting these terms for Y and $(r-s)/2$, we get two terms, namely

$$-\int_0^1 \mathrm{d}t\ \left[A\left(\zeta^s_{X+Y}+t\zeta^{r-s}_{-Y}\right)+A\left(\zeta^{-s}_{X+Y}-t\zeta^{r-s}_{Y}\right)\right]\cdot(-Y) \tag{6.16}$$

and

$$-\int_0^1 \mathrm{d}t\ \left[A\left(\zeta^s_{X+Y}+t\zeta^{r-s}_{-Y}\right)-A\left(\zeta^{-s}_{X+Y}-t\zeta^{r-s}_{Y}\right)\right]\cdot\frac{r-s}{2}. \tag{6.17}$$

We can simplify these terms by noting that

$$\begin{aligned}&A\left(X+hY\pm hs/2+th\left(-Y\pm(r-s)/2\right)\right)\\&\qquad\qquad=A\left(X+hY-thY\pm hs/2\pm th(r-s)/2\right)\end{aligned}$$

and therefore, by a Taylor expansion, the last expression equals

$$A\left(X+Y-tY\right)\pm\frac{1}{2}DA\left(X+Y-tY\right)\left(s+t(r-s)\right)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2)).$$

Therefore, going back to (6.16) and considering the difference, we see that the term with the gradient cancels out and we get that

$$\begin{aligned}&-\int_0^1 \mathrm{d}t\ \left[A(\zeta^s_{X+Y}+t\zeta^{r-s}_{-Y})+A\left(\zeta^{-s}_{X+Y}-t\zeta^{r-s}_{Y}\right)\right]\cdot(-Y)\\&=-\int_0^1 \mathrm{d}t\,2A(X+Y-tY)\cdot(-Y)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2))\\&=\Phi_{2A}\left(X,X+Y\right)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2)).\end{aligned}$$

In the case of (6.17), we have that

$$\begin{aligned}&-\int_0^1 \mathrm{d}t\ \left[A(\zeta^s_{X+Y}-t\zeta^{r-s}_{-Y})-A\left(\zeta^{-s}_{X+Y}-t\zeta^{r-s}_{Y}\right)\right]\cdot\frac{1}{2}\left(r-s\right)\\&=\frac{1}{2}\int_0^1 \mathrm{d}t\,DA\left(X+Y-tY\right)\left(s+t\left(r-s\right)\right)\cdot(r-s)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2))\\&=\frac{1}{2}\int_0^1 \mathrm{d}t\,DA\left(X\right)\left(s+t\left(r-s\right)\right)\cdot(r-s)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2)),\end{aligned}$$

where in the last step we expanded DA around the point X. Integration now yields that (6.17) equals

$$\frac{1}{4}(r-s)\cdot DA(X)(r+s)+O(\|D^2A\|_{L^\infty}(|r|^2+|s|^2)).$$

The lemma now follows. ☐

6.3.2. An operator equality. We will often make use of the operator equality

$$e^{i\Phi_A(x,x+Y)}e^{iY\cdot(-i\nabla)}=e^{iY\cdot(-i\nabla+A)}, \tag{6.18}$$

which follows as a special case by a theorem of Feynman [**32**]. Here the operators act on $L^2(\mathbb{R}^2_x)$, $Y\in\mathbb{R}^2$, and $A:\mathbb{R}^2\to\mathbb{R}^2$ is a smooth magnetic potential. In order to derive this equation, we require the fact that for two operators P and Q, one can show that

$$\exp\left(\int_0^1 \mathrm{d}t\, e^{tP} Q e^{-tP}\right) e^P = e^{P+Q}$$

if $[e^{tP}Qe^{-tP}, e^{sP}Qe^{-sP}] = 0$ for all s and t. If we now take $P = iY \cdot (-i\nabla)$ and $Q = iY \cdot A$ and a function $u \in L^2(\mathbb{R}^2)$, we find that

$$\begin{aligned}(e^{tP}Qe^{-tP}u)(x) = \left(Qe^{-tP}u\right)(x+tY) = iY\cdot A\,(x+tY)\left(e^{-tP}u\right)(x+tY)\\ = iY\cdot A\,(x+tY)\,u(x).\end{aligned}$$

Consequently,

$$\begin{aligned}\int_0^1 \mathrm{d}t\, e^{tP}Qe^{-tP}u(x) &= \int_0^1 \mathrm{d}t\, iY\cdot A\,(x+tY)\,u(x)\\ &= -i\int_0^1 \mathrm{d}t\, A\,(x+Y-tY)\cdot(-Y)u(x) = i\Phi_A\,(x, x+Y)\,u(x),\end{aligned}$$

which establishes (6.18).

6.4. Linearization

In this section we will study the linear operator that arises from the linearization of F_T^{BCS} about $\phi = 0$. We denote this operator by $L_{T,A}$. It is explicitly given by

$$L_{T,A}\alpha = \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\,(z-k_A)^{-1}\,\alpha\left(z+\overline{k}_A\right)^{-1} = \frac{2}{\beta}\sum_{n\text{ odd}}(z_n-k_A)^{-1}\,\alpha\left(z_n+\overline{k}_A\right)^{-1},$$

where $\rho(z) = \tanh(z/2)$ and $\mathcal{C}$ is $\{r \pm i\pi/(2\beta_c)\,|\,r\in\mathbb{R}\}$, and $z_n = (\pi i n)/\beta$ are the poles of $\tanh(\beta z/2)$. Note that, for h small enough the contour does not enclose any poles. Let us also introduce the non-linear map $N_{T,A}$, defined by

$$\begin{aligned}N_{T,A}(\alpha) &= \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\,(z-k_A)^{-1}\,\alpha\left(z+\overline{k}_A\right)^{-1}\overline{\alpha}\,(z-k_A)^{-1}\,\alpha\,(z-H_A(\alpha))^{-1}_{22}\\ &= \frac{2}{\beta}\sum_{n\text{ odd}}(z_n-k_A)^{-1}\,\alpha\left(z_n+\overline{k}_A\right)^{-1}\overline{\alpha}\,(z_n-k_A)^{-1}\,\alpha\,(z_n-H_A(\alpha))^{-1}_{22}.\end{aligned}$$

We can then write

$$\frac{1}{2}V^{1/2}\left(\tanh\tfrac{\beta}{2}H_A\left(-2V^{1/2}\phi\right)\right)_{12} = -V^{1/2}L_{T,A}V^{1/2}\phi + \frac{1}{2}V^{1/2}N_{T,A}\left(-2V^{1/2}\phi\right),$$

which means

$$F_T^{\mathrm{BCS}}(\phi, A) = \left(1 - V^{1/2}L_{T,A}V^{1/2}\right)\phi + \frac{1}{2}V^{1/2}N_{T,A}\left(-2V^{1/2}\phi\right).$$

6.4.1. Approximating operators. In this section we present a convenient representation of the operator $L_{T,A}$, which will also allows us to define a series of operators that approximate $L_{T,A}$. We begin by noting that

$$\begin{aligned} L_{T,A}\alpha\left(\zeta_X^r, \zeta_X^{-r}\right) &= \frac{2}{\beta}\sum_{n \text{ odd}}\left((z_n - k_A)^{-1}\alpha\left(z_n + \overline{k}_A\right)^{-1}\right)\left(\zeta_X^r, \zeta_X^{-r}\right) \\ &= -\frac{2}{\beta}\sum_{n \text{ odd}}\int_{\mathbb{R}^4} \mathrm{d}u\mathrm{d}v\, G_A^{z_n}\left(\zeta_X^r, u\right) G_{-A}^{-z_n}\left(v, \zeta_X^{-r}\right)\alpha(u,v). \end{aligned}$$

By a change of variables, or to be more precise, setting $u = Y + s/2$, $v = Y - s/2$, we further see that

$$\begin{aligned} &L_{T,A}\alpha\left(\zeta_X^r, \zeta_X^{-r}\right) \\ &= -\frac{2}{\beta}\sum_{n \text{ odd}}\int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, G_A^{z_n}\left(\zeta_X^r, \zeta_Y^s\right) G_{-A}^{-z_n}\left(\zeta_Y^{-s}, \zeta_X^{-r}\right)\alpha\left(\zeta_Y^s, \zeta_Y^{-s}\right) \\ &= -\frac{2}{\beta}\sum_{n \text{ odd}}\int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, G_A^{z_n}\left(\zeta_X^r, \zeta_{X+Y}^s\right) G_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right) e^{iY\cdot(-i\nabla_X)}\alpha\left(\zeta_X^s, \zeta_X^{-s}\right). \end{aligned}$$

Therefore, let us define

$$F_{T,A}(X,Y,r,s) := -\frac{2}{\beta}\sum_{n \text{ odd}} G_A^{z_n}\left(\zeta_X^r, \zeta_{X+Y}^s\right) G_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right),$$

which allows us to write

$$L_{T,A}\alpha\left(\zeta_X^r, \zeta_X^{-r}\right) = \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, F_{T,A}\left(X,Y,r,s\right) e^{iY\cdot(-i\nabla_X)}\alpha\left(\zeta_X^s, \zeta_X^{-s}\right).$$

In view of the phase approximation method, it will be useful to also introduce the operator $\widetilde{L}_{T,A}$, where we replace G_A^z by K_A^z in $L_{T,A}$. For the definition of $K_A^{z_n}$ we refer to (6.13). To be precise, we define

$$\begin{aligned} \widetilde{F}_{T,A}(X,Y,r,s) &:= -\frac{2}{\beta}\sum_{n \text{ odd}} K_A^{z_n}\left(\zeta_X^r, \zeta_{X+Y}^s\right) K_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right) \\ &= -\frac{2}{\beta}\sum_{n \text{ odd}} G^{z_n}\left(\zeta_Y^{-r+s}\right) G^{-z_n}\left(\zeta_Y^{r-s}\right) e^{i\Phi_A\left(\zeta_X^r, \zeta_{X+Y}^s\right) - i\Phi_A\left(\zeta_{X+Y}^{-s}, \zeta_X^{-r}\right)}, \end{aligned}$$

and let the operator $\widetilde{L}_{T,A}$ be defined by

$$\widetilde{L}_{T,A}\alpha\left(\zeta_X^r, \zeta_X^{-r}\right) = \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, \widetilde{F}_{T,A}\left(X,Y,r,s\right) e^{iY\cdot(-i\nabla_X)}\alpha\left(\zeta_X^s, \zeta_X^{-s}\right). \tag{6.19}$$

For the case $A = 0$, we drop the index A and write L_T for $L_{T,0}$. It turns out that $F_{T,0}$ only depends on Y and $r - s$, so we write $F_{T,0}(X,Y,r,s) = F_T(Y, r-s)$. Indeed we see that

$$\begin{aligned}
&F_{T,0}(X,Y,r,s)\\
&\quad= -\frac{2}{\beta}\sum_{n \text{ odd}} G^{z_n}\left(\zeta_Y^{-r+s}\right) G^{-z_n}\left(\zeta_Y^{r-s}\right)\\
&\quad= \frac{1}{(2\pi)^4}\int_{\mathbb{R}^4} \mathrm{d}p\mathrm{d}q \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\,\frac{e^{-i(p-q)\cdot(r-s)/2}e^{i(p+q)\cdot Y}}{(z-(|p|^2-\mu))\,(z+(|q|^2-\mu))}\\
&\quad= \frac{1}{(2\pi)^4}\int_{\mathbb{R}^4} \mathrm{d}p\mathrm{d}q \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\,\frac{e^{-ip\cdot(r-s)}e^{iq\cdot Y}}{(z-(|p+q/2|^2-\mu))\,(z+(|p-q/2|^2-\mu))}.
\end{aligned}$$

Now integrating over z, we have

$$\begin{aligned}
&F_{T,0}(X,Y,r,s)\\
&\quad= \frac{1}{(2\pi)^4}\int_{\mathbb{R}^4} \mathrm{d}p\mathrm{d}q\,\frac{\rho\left(\beta(|p+q/2|^2-\mu)\right)+\rho\left(\beta(|p-q/2|^2-\mu)\right)}{(|p+q/2|^2-\mu)+(|p-q/2|^2-\mu)}e^{-ip\cdot(r-s)}e^{iq\cdot Y}\\
&\quad= \frac{1}{(2\pi)^4}\int_{\mathbb{R}^4} \mathrm{d}p\mathrm{d}q\, f_T\left(p+q/2,p-q/2\right)e^{ip\cdot(r-s)}e^{iq\cdot Y}\\
&\quad=: F_T(Y,r-s),
\end{aligned}$$

where the function $f_T:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ is given by

$$f_T(p,q) = \frac{\tanh\left(\beta\left(|p|^2-\mu\right)/2\right)+\tanh\left(\beta\left(|q|^2-\mu\right)/2\right)}{(|p|^2-\mu)+(|q|^2-\mu)}.$$

One can easily verify that $F_T(-Y,r)=F_T(Y,r)$, from which it follows that we can also write

$$L_T\alpha\left(\zeta_X^r,\zeta_X^{-r}\right) = \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T\left(Y,r-s\right)\cos\left(Y\cdot\pi_0\right)\alpha\left(\zeta_X^s,\zeta_X^{-s}\right).$$

We will see below that L_T is not in all cases a sufficiently good approximation of $L_{T,A}$. Therefore, on the basis of Lemma 6.7, we introduce the intermediate approximations

$$\begin{aligned}
\widetilde{M}_{T,A}\alpha\left(\zeta_X^r,\zeta_X^{-r}\right) &= \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T\left(Y,r-s\right)e^{i(r-s)/4\cdot DA(X)(r+s)}e^{iY\cdot\pi_{2A}}\alpha\left(\zeta_X^s,\zeta_X^{-s}\right)\\
&= \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T\left(Y,r-s\right)e^{i(r-s)/4\cdot DA(X)(r+s)}\cos\left(Y\cdot\pi_{2A}\right)\alpha\left(\zeta_X^s,\zeta_X^{-s}\right), \quad (6.20)
\end{aligned}$$

and

$$\begin{aligned}
M_{T,A}\alpha\left(\zeta_X^r,\zeta_X^{-r}\right) &= \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T\left(Y,r-s\right)e^{iY\cdot\pi_{2A}}\alpha\left(\zeta_X^s,\zeta_X^{-s}\right)\\
&= \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T\left(Y,r-s\right)\cos\left(Y\cdot\pi_{2A}\right)\alpha\left(\zeta_X^s,\zeta_X^{-s}\right). \quad (6.21)
\end{aligned}$$

Here, and below, π_{2A} is always to be understood as $\pi_{2A}=(-i\nabla_X+2A(X))$.

Concerning the function f_T, we mention here the fact that $f_T(p,p)=1/K_T(p)$. We also have the following important inequality.

LEMMA 6.8. *For all $p, q \in \mathbb{R}^2$,*

$$f_T(p,q) \leq \frac{1}{2}\left(f_T(p,p) + f_T(q,q)\right),$$

for $T > 0$.

PROOF. We want to show that

$$\frac{\tanh a + \tanh b}{a+b} \leq \frac{1}{2}\left(\frac{\tanh a}{a} + \frac{\tanh b}{b}\right).$$

We assume $a \geq |b|$ without loss of generality. Consequently, $\tanh(b)/b \geq \tanh(a)/a$. The claimed inequality then follows easily by rearrangement,

$$\begin{aligned}
&\frac{1}{2}\left(\frac{a\tanh(a)/a + b\tanh(b)/b}{a+b} + \frac{\tanh(a)+\tanh(b)}{a+b}\right)\\
&\leq \frac{1}{2}\left(\frac{b\tanh(a)/a + a\tanh(b)/b}{a+b} + \frac{\tanh(a)+\tanh(b)}{a+b}\right)\\
&= \frac{1}{2}\left(\frac{\tanh(a)}{a} + \frac{\tanh(b)}{b}\right).
\end{aligned}$$

This completes the proof. □

6.4.2. Basic estimates. We now prove some basic estimates concerning the above linear operators. We recall that we have $T_h = \mathcal{T}_c(1 - Dh^2)$ and $\mathcal{T}_c > 0$, which implies $T_h \lesssim 1$.

LEMMA 6.9. *Suppose that $A \in \dot{W}^{1,\infty}(\mathbb{R}^2) \cap \dot{W}^{2,\infty}(\mathbb{R}^2)$. We then have*

$$\left\|V^{1/2}\left(L_{T,A} - M_{T,A}\right)V^{1/2}\alpha\right\|_{\mathcal{L}_h^2} \lesssim \left(\|D^2A\|_{L^\infty} + \|DA\|_{L^\infty}^2\right)\|\alpha\|_{\mathcal{L}_h^2}$$

and

$$\left\|V^{1/2}\left(L_{T,A} - L_T\right)V^{1/2}\alpha\right\|_{\mathcal{L}_h^2} \lesssim \left(\|DA\|_{L^\infty} + \|D^2A\|_{L^\infty} + \|DA\|_{L^\infty}^2\right)\|\alpha\|_{\mathcal{L}_h^2},$$

for all $\alpha \in \mathcal{L}^2_{h,\mathrm{sym}}$.

PROOF. Our first goal is to estimate $L_{T,A} - \widetilde{L}_{T,A}$, where the latter operator is defined above by (6.19). To do this it is useful to first estimate

$$\int_{\mathbb{R}^2} \mathrm{d}Y \operatorname*{ess\,sup}_{X \in Q_h} \left|F_{T,A}(X,Y,r,s) - \widetilde{F}_{T,A}(X,Y,r,s)\right|.$$

We begin by observing that

$$
\begin{aligned}
&\frac{\beta}{2}\left|F_{T,A}(X,Y,r,s)-\widetilde{F}_{T,A}(X,Y,r,s)\right| \\
&=\left|\sum_{n\text{ odd}} G_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right) G_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)-\sum_{n\text{ odd}} K_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right) K_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)\right| \\
&\leq \sum_{n\text{ odd}}\left|G_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right)-K_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right)\right|\left|K_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)\right| \\
&\quad+\sum_{n\text{ odd}}\left|K_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right)\right|\left|G_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)-K_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)\right| \\
&\quad+\sum_{n\text{ odd}}\left|G_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right)-K_A^{z_n}\left(\zeta_X^r,\zeta_{X+Y}^s\right)\right|\left|G_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)-K_{-A}^{-z_n}\left(\zeta_{X+Y}^{-s},\zeta_X^{-r}\right)\right|
\end{aligned}
$$

Applying Corollary 6.6, we finally get the estimate

$$
\begin{aligned}
&\left|F_{T,A}(X,Y,r,s)-\widetilde{F}_{T,A}(X,Y,r,s)\right| \\
&\qquad\leq \frac{2}{\beta}\sum_{n\text{ odd}}\left[H_A^{z_n}\left(\zeta_{-Y}^{r-s}\right)\left|G^{-z_n}\left(\zeta_Y^{r-s}\right)\right|+\left|G^{z_n}\left(\zeta_{-Y}^{r-s}\right)\right| H_{-A}^{-z_n}\left(\zeta_Y^{r-s}\right)\right. \\
&\qquad\qquad\left.+\,H_A^{z_n}\left(\zeta_{-Y}^{r-s}\right) H_{-A}^{-z_n}\left(\zeta_Y^{r-s}\right)\right].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\int_{\mathbb{R}^2}\mathrm{d}Y\ \operatorname*{ess\,sup}_{X\in Q_h}\left|F_{T,A}(X,Y,r,s)-\widetilde{F}_{T,A}(X,Y,r,s)\right| \\
&\qquad\leq \frac{2}{\beta}\sum_{n\text{ odd}}\int_{\mathbb{R}^2}\mathrm{d}Y\ \left[H_A^{z_n}(r-s-Y)\left|G^{-z_n}(Y)\right|+\left|G^{z_n}(r-s-Y)\right| H_{-A}^{-z_n}(Y)\right. \\
&\qquad\qquad\left.+H_A^{z_n}(r-s-Y)H_{-A}^{-z_n}(Y)\right].
\end{aligned}
$$

We set

$$
q_{T,A}^{(1)}(r):=\frac{2}{\beta}\sum_{n\text{ odd}}\left(\left(H_A^{z_n}*\left|G^{-z_n}\right|\right)(r)+\left(\left|G^{z_n}\right|*H_{-A}^{-z_n}\right)(r)+\left(H_A^{z_n}*H_{-A}^{-z_n}\right)(r)\right)
$$

and conclude that

$$
\int_{\mathbb{R}^2}\mathrm{d}Y\ \operatorname*{ess\,sup}_{X\in Q_h}\left|F_{T,A}(X,Y,r,s)-\widetilde{F}_{T,A}(X,Y,r,s)\right|\leq q_{T,A}^{(1)}(r-s). \tag{6.22}
$$

With Corollary 6.6, it is straightforward to verify that $\|q^{(1)}_{T,A}\|_{L^1} \lesssim \|D^2A\|_{L^\infty} + \|DA\|^2_{L^\infty}$. We now see that for $\xi, \phi \in \mathcal{L}^2_h$, we have

$$\begin{aligned}
&\left|\left\langle \xi \left| \left(L_{T,A} - \widetilde{L}_{T,A}\right) \phi \right.\right\rangle\right| \\
&= h^2 \left| \int_{\mathbb{R}^2} \mathrm{d}r \int_{Q_h} \mathrm{d}X\, \overline{\xi\left(\zeta^r_X, \zeta^{-r}_X\right)} \int_{\mathbb{R}^2} \mathrm{d}s\mathrm{d}Y\, \left(F_{T,A}(X,Y,r,s) - \widetilde{F}_{T,A}(X,Y,r,s)\right) \right. \\
&\qquad \left. \times e^{iY\cdot(-i\nabla_X)} \phi\left(\zeta^s_X, \zeta^{-s}_X\right) \right| \\
&\lesssim \int_{\mathbb{R}^2} \mathrm{d}r\mathrm{d}s\mathrm{d}Y\, \operatorname*{ess\,sup}_{X\in Q_h} \left| F_{T,A}(X,Y,r,s) - \widetilde{F}_{T,A}(X,Y,r,s) \right| \\
&\qquad \times \left| \left\langle \xi\left(\cdot + r/2, \cdot - r/2\right) \left| e^{iY\cdot\pi_{2A}} \phi\left(\cdot + s/2, \cdot - s/2\right) \right.\right\rangle_{L^2(Q_h)} \right| .
\end{aligned}$$

Finally, by (6.22) we find that

$$\begin{aligned}
&\left|\left\langle \xi \left| \left(L_{T,A} - \widetilde{L}_{T,A}\right) \phi \right.\right\rangle\right| \\
&\lesssim \int_{\mathbb{R}^4} \mathrm{d}r\mathrm{d}s\, q^{(1)}_{T,A}(r-s) \left\|\xi\left(\cdot + r/2, \cdot - r/2\right)\right\|_{L^2(Q_h)} \left\|\phi\left(\cdot + s/2, \cdot - s/2\right)\right\|_{L^2(Q_h)} \\
&\lesssim \left\| q^{(1)}_{T,A} \right\|_{L^1} \|\xi\|_{\mathcal{L}^2_h} \|\phi\|_{\mathcal{L}^2_h} ,
\end{aligned}$$

where the last inequality follows from Young's inequality for convolutions. Thus we have shown that

$$\left\| V^{1/2} \left(L_{T,A} - \widetilde{L}_{T,A}\right) V^{1/2} \phi \right\|_{\mathcal{L}^2_h} \lesssim \left(\|D^2A\|_{L^\infty} + \|DA\|^2_{L^\infty}\right) \|\phi\|_{\mathcal{L}^2_h} .$$

We now wish to estimate $V^{1/2}(\widetilde{L}_{T,A} - \widetilde{M}_{T,A})V^{1/2}$. Using (6.18), we rewrite

$$\begin{aligned}
\widetilde{L}_{T,A}\alpha\left(\zeta^r_X, \zeta^{-r}_X\right) = &\int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, \widetilde{F}_{T,A}\left(X,Y,r,s\right) e^{-i\Phi_{2A}(X,X+Y)} e^{-i(r-s)/4\cdot DA(X)(r+s)} \\
&\times e^{i(r-s)/4\cdot DA(X)(r+s)} e^{iY\cdot\pi_{2A}} \alpha\left(\zeta^s_X, \zeta^{-s}_X\right) .
\end{aligned}$$

We now estimate

$$\begin{aligned}
&\int_{\mathbb{R}^2} \mathrm{d}Y\, \operatorname*{ess\,sup}_{X\in Q_h} \left| \widetilde{F}_{T,A}(X,Y,r,s) e^{-i\Phi_{2A}(X,X+Y)} e^{-i(r-s)/4\cdot DA(X)(r+s)} - F_T(Y, r-s) \right| \\
&\lesssim \sum_{n \text{ odd}} \int_{\mathbb{R}^2} \mathrm{d}Y\, \left| G^{z_n}\left(\zeta^{-(r-s)}_Y\right) G^{-z_n}\left(\zeta^{r-s}_Y\right) \right| \\
&\qquad \times \operatorname*{ess\,sup}_{X\in Q_h} \left| \left(e^{i\Phi_A\left(\zeta^r_X, \zeta^s_{X+Y}\right) - i\Phi_A\left(\zeta^{-s}_{X+Y}, \zeta^{-r}_X\right) - i\Phi_{2A}(X,X+Y) - i(r-s)/4\cdot DA(X)(r+s)} - 1 \right) \right| \\
&\lesssim \sum_{n \text{ odd}} \int_{\mathbb{R}^2} \mathrm{d}Y\, \left| G^{z_n}\left(\zeta^{-(r-s)}_Y\right) G^{-z_n}\left(\zeta^{r-s}_Y\right) \right| \|D^2A\|_{L^\infty} (|r|^2 + |s|^2),
\end{aligned}$$

where in the last step we applied Lemma 6.7. As before we then obtain

$$\left\langle \xi \left| V^{1/2}(\widetilde{L}_{T,A} - \widetilde{M}_{T,A})V^{1/2} \phi \right.\right\rangle \lesssim \|D^2A\|_{L^\infty} \|\xi\|_{\mathcal{L}^2_h} \|\phi\|_{\mathcal{L}^2_h} .$$

We now need to estimate $V^{1/2}(\widetilde{M}_{T,A} - M_{T,A})V^{1/2}$, and here we will need to use the fact that α is assumed to be symmetric. Indeed, if α and ξ are symmetric, then one can show that in fact

$$\begin{aligned}
&\left\langle \xi \,\middle|\, V^{1/2}\widetilde{M}_{T,A}V^{1/2}\alpha \right\rangle \\
&\quad = h^2 \int_{Q_h} \mathrm{d}X \int_{\mathbb{R}^6} \mathrm{d}r\mathrm{d}s\mathrm{d}Y\, F_T\left(Y, r-s\right) \cos\left(\frac{1}{4}(r-s)\cdot DA(X)(r+s)\right) \\
&\qquad\qquad \times \overline{\xi\left(\zeta_X^r, \zeta_X^{-r}\right)} e^{iY\cdot\pi_{2A}}\alpha\left(\zeta_X^s, \zeta_X^{-s}\right).
\end{aligned}$$

If we now expand

$$\cos\left(\frac{1}{4}(r-s)\cdot DA(X)(r+s)\right) = 1 + O(\|DA\|_{L^\infty}^2(|r|^4 + |s|^4)),$$

we see that the first term gives simply $M_{T,A}$, and the error can be controlled as above to show that

$$\left\langle \xi \,\middle|\, V^{1/2}(\widetilde{M}_{T,A} - M_{T,A})V^{1/2}\phi \right\rangle \lesssim \|DA\|_{L^\infty}^2 \|\xi\|_{\mathcal{L}_h^2} \|\phi\|_{\mathcal{L}_h^2}.$$

This completes the proof of the first estimate.

For the second estimate, we only have to estimate the difference $M_{T,A} - L_T$. To do this we first write

$$\begin{aligned}
&\left(M_{T,A} - L_T\right)\phi\left(\zeta_X^r, \zeta_X^{-r}\right) \\
&= \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, F_T(Y, r-s)\left(e^{iY\cdot\Pi_{2A}} - e^{iY\cdot\Pi_0}\right)\phi\left(\zeta_X^s, \zeta_X^{-s}\right) \\
&= \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, F_T(Y, r-s)\left(e^{i\Phi_{2A}(X,X+Y)} - 1\right)e^{iY\cdot\Pi_0}\phi\left(\zeta_X^s, \zeta_X^{-s}\right).
\end{aligned}$$

Note that we have the simple estimate

$$\left|e^{i\Phi_{2A}(X,X+Y)} - 1\right| \le |\Phi_{2A}(X, X+Y)| \lesssim \|A\|_{L^\infty}|Y|.$$

This means, as above, that

$$\begin{aligned}
&\left|\left\langle \xi \,\middle|\, \left(M_{T,A} - L_T\right)\phi \right\rangle\right| \\
&= \left| \int_{\mathbb{R}^2} \mathrm{d}r \int_{Q_h} \mathrm{d}X\, \overline{\xi\left(\zeta_X^r, \zeta_X^{-r}\right)} \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, F_T(Y, r-s) \right. \\
&\qquad\qquad \left. \times \left(e^{i\Phi_{2A}(X,X+Y)} - 1\right)e^{iY\cdot\Pi_0}\phi\left(\zeta_X^s, \zeta_X^{-s}\right) \right| \\
&\lesssim h \int_{\mathbb{R}^6} \mathrm{d}r\mathrm{d}s\mathrm{d}Y \int_{Q_h} \mathrm{d}X\, |YF_T(Y, r-s)| \left|\xi\left(\zeta_X^r, \zeta_X^{-r}\right)\right| \left|e^{iY\cdot\Pi_0}\phi\left(\zeta_X^s, \zeta_X^{-s}\right)\right| \\
&\lesssim h \|\xi\|_{\mathcal{L}_h^2} \|\phi\|_{\mathcal{L}_h^2}.
\end{aligned}$$

That completes the proof of the lemma. □

LEMMA 6.10. *Suppose that $A \in W^{1,\infty}(\mathbb{R}^2)$. Then for all $\Psi \in L^2(Q_h)$ we have*

$$\left\|\left(M_{T,A} - L_T\right)\varphi_*\Psi\right\|_{\mathcal{L}_h^2} \lesssim \left(\|DA\|_{L^\infty} + \|A\|_{L^\infty}^2\right)\|\Psi\|_{L^2(Q_h)}.$$

PROOF. We first write

$$\begin{aligned}&(M_{T,A}-L_T)\,\phi\left(\zeta_X^r,\zeta_X^{-r}\right)\\&\qquad=\int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\,F_T(Y,r-s)\left(\cos\left(Y\cdot\pi_{2A}\right)-\cos\left(Y\cdot\pi_0\right)\right)\phi\left(\zeta_X^s,\zeta_X^{-s}\right).\end{aligned}$$

We now claim that for $\Psi\in L^2(Q_h)$, we have

$$\begin{aligned}&\left\|\left(\cos\left(Y\cdot\pi_{2A}\right)-\cos\left(Y\cdot\pi_0\right)\right)\Psi\right\|_{L^2(Q_h)}\\&\qquad=\frac{1}{2}\left\|\left(e^{iY\cdot\pi_{2A}}+e^{-iY\cdot\pi_{2A}}-e^{iY\cdot\pi_0}-e^{-iY\cdot\pi_0}\right)\Psi\right\|_{L^2(Q_h)}\\&\qquad=\frac{1}{2}\left\|\left(e^{iY\cdot\pi_0}\left(e^{-iY\cdot\pi_0}e^{iY\cdot\pi_{2A}}-1\right)+e^{-iY\cdot\pi_0}\left(e^{iY\cdot\pi_0}e^{-iY\cdot\pi_{2A}}-1\right)\right)\Psi\right\|_{L^2(Q_h)}\\&\qquad\lesssim|Y|^2\left(\|DA\|_{L^\infty}+\|A\|_{L^\infty}^2\right)\|\Psi\|_{L^2(Q_h)}.\end{aligned}$$

To see this we define

$$\Psi_t=\left(e^{iY\cdot\pi_0}\left(e^{-itY\cdot\pi_0}e^{itY\cdot\pi_{2A}}-1\right)+e^{-iY\cdot\pi_0}\left(e^{itY\cdot\pi_0}e^{-itY\cdot\pi_{2A}}-1\right)\right)\Psi$$

and calculate that

$$\partial_t\Psi_t=\left(e^{iY\cdot\pi_0}e^{-itY\cdot\pi_0}(2iY\cdot A)e^{itY\cdot\pi_{2A}}+e^{-iY\cdot\pi_0}e^{itY\cdot\pi_0}(-2iY\cdot A)e^{-itY\cdot\pi_{2A}}\right)\Psi.$$

Note that the derivative vanishes at $t=0$. We also calculate

$$\begin{aligned}\partial_t^2\Psi_t=&\left(e^{iY\cdot\pi_0}e^{-itY\cdot\pi_0}\left((-iY\cdot\pi_0)(2iY\cdot A)+(2iY\cdot A)(iY\cdot\pi_{2A})\right)e^{itY\cdot\pi_{2A}}\right.\\&\left.+e^{-iY\cdot\pi_0}e^{itY\cdot\pi_0}\left((iY\cdot\pi_0)(-2iY\cdot A)+(-2iY\cdot A)(-iY\cdot\pi_{2A})\right)e^{-itY\cdot\pi_{2A}}\right)\Psi.\end{aligned}$$

Now using the fact that $[Y\cdot\pi_0,Y\cdot A]=Y\cdot(DA)Y$ and therefore

$$\begin{aligned}(-iY\cdot\pi_0)(2iY\cdot A)+(2iY\cdot A)(Y\cdot\pi_{2A})&=2[Y\cdot\pi_0,Y\cdot A]-4(Y\cdot A)^2\\&=2Y\cdot(DA)Y-4(Y\cdot A)^2,\end{aligned}$$

we see that

$$\|\partial_t^2\Psi_t\|_{L^2(Q_h)}\lesssim|Y|^2\left(\|DA\|_{L^\infty}+\|A\|_{L^\infty}^2\right)\|\Psi\|_{L^2(Q_h)}.$$

But this means that

$$\begin{aligned}&\left\|\left(e^{iY\cdot\pi_{2A}}+e^{-iY\cdot\pi_{2A}}-e^{iY\cdot\pi_0}-e^{-iY\cdot\pi_0}\right)\Psi\right\|_{L^2(Q_h)}=\|\Psi_1-\Psi_0\|_{L^2(Q_h)}\\&\qquad\le\int_0^1\mathrm{d}t\,(1-t)\left\|\partial_t^2\Psi_t\right\|_{L^2(Q_h)}\lesssim|Y|^2\left(\|DA\|_{L^\infty}+\|A\|_{L^\infty}^2\right)\|\Psi\|_{L^2(Q_h)},\end{aligned}$$

which establishes the claim.

Now, since $\int_{\mathbb{R}^2}\mathrm{d}Y\,|Y|^2F_T(Y,r-s)$ is an integrable function of $r-s$, we can show, as in the previous lemma, that for any $\xi\in\mathcal{L}_h^2$, we have

$$|\langle\xi\,|(M_{T,A}-L_T)\,\chi_\varepsilon\phi\rangle|\lesssim\|\xi\|_{\mathcal{L}_h^2}\left(\|DA\|_{L^\infty}+\|A\|_{L^\infty}^2\right)\|\Psi\|_{L^2(Q_h)},$$

and the lemma now follows. ☐

Lemma 6.11. *If $\mathcal{T}_c > 0$, then we have*

$$\|L_T - L_{\mathcal{T}_c}\|_{\mathcal{L}^2_h} \lesssim |\mathcal{T}_c - T| \quad \text{and} \quad \|M_{T,A} - M_{\mathcal{T}_c,A}\|_{\mathcal{L}^2_h} \lesssim |\mathcal{T}_c - T|\,,$$

for all T sufficiently close to $\mathcal{T}_c$.

Proof. The strategy is the same as in the previous lemma. We write the proof only for the first estimate, that is, for the difference of the L_T, as the second case involving $M_{T,A}$ is the same. We first note that

$$\begin{aligned}&F_T(Y, r-s) - F_{\mathcal{T}_c}(Y, r-s)\\&= (2\pi)^{-4} \int_{\mathbb{R}^4} \mathrm{d}p\mathrm{d}q\, \left(f_T(p+q/2, p-q/2) - f_{\mathcal{T}_c}(p+q/2, p-q/2)\right) e^{ip\cdot(r-s)} e^{iq\cdot Y}.\end{aligned}$$

Now, since $\beta = 1/T$, we have $\beta = \upbeta_c + \upbeta_c(\mathcal{T}_c - T)$ and therefore $\rho(\beta z) = \rho(\upbeta_c z) + O(|\mathcal{T}_c - T|)$, where the remainder decays exponentially for $z \in \mathcal{C}$. This means that the difference $f_T(p+q/2, p-q/2) - f_{\mathcal{T}_c}(p+q/2, p-q/2)$ is also $O(|\mathcal{T}_c - T|)$ with exponential decay in p and q. It follows that

$$\int_{\mathbb{R}^2} \mathrm{d}Y \operatorname*{ess\,sup}_{X\in Q_h} |F_T(Y, r-s) - F_{\mathcal{T}_c}(Y, r-s)| \le h^2 q(r-s),$$

where $q \in L^1(\mathbb{R}^2)$ and $\|q\|_{L^1} \lesssim 1$. We then have that for any $\xi \in \mathcal{L}^2_h$ that

$$|\langle \xi\, |(L_T - L_{\mathcal{T}_c})\, \phi\rangle| \lesssim |\mathcal{T}_c - T|\, \|q\|_{L^1} \|\xi\|_{\mathcal{L}^2_h} \|\phi\|_{\mathcal{L}^2_h}\,,$$

and the lemma now follows. □

Lemma 6.12. *We have*

$$\left\|\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) \varphi_* \Psi\right\|_{\mathcal{L}^2_h} \lesssim \|\Delta\Psi\|_{L^2(Q_h)}$$

for all $\Psi \in H^2(Q_h)$.

Proof. We first calculate that

$$\begin{aligned}&V^{1/2} L_{\mathcal{T}_c} V^{1/2} P\phi\left(\zeta^r_X, \zeta^{-r}_X\right)\\&= V^{1/2}(r) \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}Y\, \left(V\alpha_*\right)(s) F_{\mathcal{T}_c}(Y, r-s) e^{iY\cdot(-i\nabla_X)} \Psi(X)\\&= \frac{1}{(2\pi)^4} \int_{\mathbb{R}^8} \mathrm{d}s\mathrm{d}Y\mathrm{d}p\mathrm{d}q\, V^{1/2}(r)\left(V\alpha_*\right)(s) f_{\mathcal{T}_c}\left(p+q/2, p-q/2\right) e^{ip\cdot(r-s)} e^{iq\cdot Y}\\&\qquad\qquad \times \sum_{m\in(2\pi\mathbb{Z})^2} e^{ihm\cdot X} e^{ihm\cdot Y} \hat{\Psi}(m)\\&= \frac{1}{(2\pi)^2} \sum_{m\in(2\pi\mathbb{Z})^2} e^{ihm\cdot X} \int_{\mathbb{R}^4} \mathrm{d}s\mathrm{d}p\, e^{ip\cdot(r-s)} V^{1/2}(r)\left(V\alpha_*\right)(s)\\&\qquad\qquad \times f_{\mathcal{T}_c}\left(p+hm/2, p-hm/2\right) \hat{\Psi}(m).\end{aligned}$$

On the other hand, recalling that $(1 - V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2})\varphi_* = 0$, or, in other words, $\varphi_* = V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\varphi_*$, we see that we can write

$$
\begin{aligned}
P\phi\left(\zeta_X^r, \zeta_X^{-r}\right) &= \varphi_*(r)\Psi(X) \\
&= \left(V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\varphi_*\right)(r)\Psi(X) \\
&= \frac{1}{(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2} e^{ihm\cdot X}\int_{\mathbb{R}^4}\mathrm{d}p\mathrm{d}s\, e^{ip\cdot(r-s)}V^{1/2}(r)\left(V\alpha_*\right)(s)f_{\mathcal{T}_c}(p,p)\Psi(X).
\end{aligned}
$$

Using the inequality $|f_{\mathcal{T}_c}(p,p) - f_{\mathcal{T}_c}(p + hm/2, p - hm/2)| \lesssim \min\{1, h^2|m|^2\}$, where the right-hand side decays exponentially in p, we see that

$$
\left\|\left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)\varphi_*\Psi\right\|_{\mathcal{L}_h^2}^2 \lesssim \sum_{m\in(2\pi\mathbb{Z})^2} h^4|m|^4|\hat{\Psi}(m)|^2 = \|\Delta\Psi\|_{L^2(Q_h)}^2,
$$

and that proves the lemma. □

6.4.3. Coercivity estimates. In this section we state and prove two important estimates, that are presented in the following lemma.

LEMMA 6.13. *If h is sufficiently small, then we have*

$$
\varepsilon^2\left\|P_\varepsilon^\perp\xi\right\|_{\mathcal{L}_h^2} \lesssim \left\|P_\varepsilon^\perp\left(1 - V^{1/2}L_{T_h,A_h}V^{1/2}\right)P_\varepsilon^\perp\xi\right\|_{\mathcal{L}_h^2}
$$

and

$$
\left\|P^\perp\xi\right\|_{\mathcal{L}_h^2} \lesssim \left\|P^\perp\left(1 - V^{1/2}L_{T_h,A_h}V^{1/2}\right)P^\perp\xi\right\|_{\mathcal{L}_h^2}
$$

for any $\xi \in \mathcal{L}_h^2$.

PROOF. The above lemmas show, in effect, that

$$
\left\|\left(1 - V^{1/2}L_{T,A_h}V^{1/2}\right) - \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)\right\|_{\mathcal{L}_h^2} \lesssim h.
$$

This means that the lemma will be proven once we have shown that

$$
P^\perp\left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)P^\perp \geq CP^\perp, \tag{6.23}
$$

and

$$
P_\varepsilon^\perp\left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)P_\varepsilon^\perp \geq C\varepsilon^2P_\varepsilon^\perp, \tag{6.24}
$$

where $C > 0$ is some positive constant independent of h. We follow here the argument given in [**37**]. To do this, let U be the unitary map on $\mathcal{L}_h^2$ given by $U = e^{-ir/2\cdot(-i\nabla_X)}$. We will first show that

$$
V^{1/2}L_{\mathcal{T}_c}V^{1/2} \leq \frac{1}{2}U^*V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}U + \frac{1}{2}UV^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}U^*, \tag{6.25}
$$

where we view $K_{\mathcal{T}_c}$ as an operator acting only on the $x - y$ coordinate. To do this, we recall the notation $\hat{\phi}_m$ for the Fourier transform in the center-of-mass coordinate, as well as the relation

$$
\phi\left(\zeta_X^r, \zeta_X^{-r}\right) = \sum_{m\in(2\pi\mathbb{Z})^2} e^{ihm\cdot X}\hat{\phi}_m(r),
$$

which implies

$$\langle\phi|\eta\rangle = \sum_{m\in(2\pi\mathbb{Z})^2} \int_{\mathbb{R}^2} \mathrm{d}r\, \overline{\hat{\phi}_m(r)}\, \hat{\eta}_m(r).$$

We begin by writing

$$\begin{aligned}\langle\phi\,|L_{\mathcal{T}_c}\phi\rangle &= \int_{\mathbb{R}^2}\mathrm{d}r \int_{Q_h}\mathrm{d}X\, \overline{\phi\left(\zeta_X^r,\zeta_X^{-r}\right)} \int_{\mathbb{R}^4}\mathrm{d}s\mathrm{d}Y\, F_{\mathcal{T}_c}(r-s,Y)e^{iY\cdot(-i\nabla_X)}\phi\left(\zeta_X^s,\zeta_X^{-s}\right)\\ &= \sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^6}\mathrm{d}r\mathrm{d}s\mathrm{d}Y\, F_{\mathcal{T}_c}(r-s,Y)e^{ihm\cdot Y}\overline{\hat{\phi}_m(r)}\hat{\phi}_m(s),\end{aligned}$$

which is easily seen to be equal to

$$\begin{aligned}&\frac{1}{(2\pi)^4}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^{10}}\mathrm{d}r\mathrm{d}s\mathrm{d}Y\mathrm{d}p\mathrm{d}q\, f_{\mathcal{T}_c}\left(p+q/2,p-q/2\right)e^{ip\cdot(r-s)}e^{-iq\cdot Y}e^{ihm\cdot Y}\overline{\hat{\phi}_m(r)}\hat{\phi}_m(s)\\ &= \frac{1}{(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^2}\mathrm{d}p\, f_{\mathcal{T}_c}\left(p+hm/2,p-hm/2\right)\left|\int_{\mathbb{R}^2}\mathrm{d}r\, e^{ip\cdot r}\hat{\phi}_m(r)\right|^2.\end{aligned}$$

We now use Lemma 6.8 to see that

$$\begin{aligned}&\langle\phi|L_{\mathcal{T}_c}\phi\rangle\\ &\le \frac{1}{2(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^2}\mathrm{d}p\\ &\qquad\times\left(f_{\mathcal{T}_c}\left(p+hm/2,p+hm/2\right)+f_{\mathcal{T}_c}\left(p-hm/2,p-hm/2\right)\right)\left|\int_{\mathbb{R}^2}\mathrm{d}r\, e^{ip\cdot r}\hat{\phi}_m(r)\right|^2\\ &= \frac{1}{2(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^6}\mathrm{d}r\mathrm{d}s\mathrm{d}p\, f_{\mathcal{T}_c}(p,p)e^{ip\cdot(r-s)}\overline{e^{ir/2\cdot hm}\hat{\phi}_m(r)}e^{is/2\cdot hm}\hat{\phi}_m(s)\\ &\quad+\frac{1}{2(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^6}\mathrm{d}r\mathrm{d}s\mathrm{d}p\, f_{\mathcal{T}_c}(p,p)e^{ip\cdot(r-s)}\overline{e^{-ir/2\cdot hm}\hat{\phi}_m(r)}e^{-is/2\cdot hm}\hat{\phi}_m(s).\end{aligned}$$

As a consequence,

$$\langle\phi\,|L_{\mathcal{T}_c}\phi\rangle \le \frac{1}{2}\left\langle\phi\,\middle|\left(UK_{\mathcal{T}_c}^{-1}U^*+U^*K_{\mathcal{T}_c}^{-1}U\right)\phi\right\rangle.$$

The bound (6.25) now follows, since $V^{1/2}$ is a multiplication operator in the relative coordinate and therefore commutes with U.

Since α_* is a non-degenerate ground state of $K_{\mathcal{T}_c}-V$, it follows that there exists some $\kappa>0$ such that $1-V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\ge\kappa P^\perp$. We can now see that

$$\begin{aligned}1-V^{1/2}L_{\mathcal{T}_c}V^{1/2} &\ge \frac{1}{2}U\left(1-V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\right)U^*+\frac{1}{2}U^*\left(1-V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\right)U\\ &\ge\frac{\kappa}{2}\left(UP^\perp U^*+U^*P^\perp U\right)\\ &=\kappa\left(1-\frac{1}{2}UPU^*-\frac{1}{2}U^*PU\right).\end{aligned}$$

We now wish to derive a lower bound for the operator $1-(UPU^*-U^*PU)/2$. We begin by calculating that

$$\begin{aligned}
&\langle \phi \,|(UPU^* + U^*PU)\, \phi \rangle \\
&= \sum_{m\in(2\pi\mathbb{Z})^2} \int_{\mathbb{R}^2} \mathrm{d}r\, \overline{\hat{\phi}_m(r)} \Big(e^{-ihr/2\cdot m} \varphi_*(r) \int_{\mathbb{R}^2} \mathrm{d}s\, \varphi_*(s) e^{ihs/2\cdot m} \hat{\phi}_m(s) \\
&\qquad\qquad + e^{ihr/2\cdot m} \varphi_*(r) \int_{\mathbb{R}^2} \mathrm{d}s\, \varphi_*(s) e^{-ihs/2\cdot m} \hat{\phi}_m(s) \Big) \\
&= \sum_{m\in(2\pi\mathbb{Z})^2} \Big[\Big\langle \hat{\phi}_m \,\Big| e^{-ihm/2\cdot r} \varphi_* \Big\rangle \Big\langle e^{-ihm/2\cdot r} \varphi_* \Big| \,\hat{\phi}_m \Big\rangle \\
&\qquad\qquad + \Big\langle \hat{\phi}_m \,\Big| e^{ihm/2\cdot r} \varphi_* \Big\rangle \Big\langle e^{ihm/2\cdot r} \varphi_* \Big| \,\hat{\phi}_m \Big\rangle \Big] \\
&= 2 \left| \langle \varphi_* \,| \varphi_* \rangle \right|^2 + \sum_{\substack{m\in(2\pi\mathbb{Z})^2 \\ m\neq 0}} \Big\langle \hat{\phi}_m \,| Q_m | \,\hat{\phi}_m \Big\rangle ,
\end{aligned}$$

where Q_m denotes the rank 2 operator

$$\left| e^{ihm/2\cdot r} \varphi_* \right\rangle \left\langle e^{ihm/2\cdot r} \varphi_* \right| + \left| e^{-ihm/2\cdot r} \varphi_* \right\rangle \left\langle e^{-ihm/2\cdot r} \varphi_* \right| .$$

In order to study the quadratic form associated to Q_m we distinguish two cases. In the first case we assume that $\hat{\phi}_m$ is orthogonal to φ_*. We can thus consider the operator

$$(1 - |\varphi_*\rangle \langle \varphi_*|)\, Q_m\, (1 - |\varphi_*\rangle \langle \varphi_*|) .$$

In the basis

$$\left\{ (1 - |\varphi_*\rangle \langle \varphi_*|)\, e^{ihm/2\cdot r} \varphi_*, (1 - |\varphi_*\rangle \langle \varphi_*|)\, e^{-ihm/2\cdot r} \varphi_* \right\}$$

this operator is represented by the matrix

$$\begin{pmatrix} 1-\lambda_m^2 & \lambda_{2m} - \lambda_m^2 \\ \lambda_{2m} - \lambda_m^2 & 1-\lambda_m^2 \end{pmatrix},$$

where

$$\lambda_m = \int_{\mathbb{R}^2} \mathrm{d}r\, e^{ihm/2\cdot r} \left| \varphi_*(r) \right|^2 = \int_{\mathbb{R}^2} \mathrm{d}r\, \cos\left(hm/2 \cdot r\right) \left| \varphi_*(r) \right|^2 .$$

The eigenvalues of this matrix are

$$1 - \lambda_m^2 \pm \left| \lambda_{2m} - \lambda_m^2 \right| ,$$

and we conclude that if $\hat{\phi}_m \perp \varphi_*$, as assumed, then

$$\left\langle \hat{\phi}_m | Q_m | \hat{\phi}_m \right\rangle \leq \left(1 - \lambda_m^2 + \left| \lambda_{2m} - \lambda_m^2 \right| \right) \left\| \hat{\phi}_m \right\|_{L^2(\mathbb{R}^2)}^2 .$$

For the general case, where $\hat{\phi}_m$ is not necessarily orthogonal to φ_*, we must consider Q_m directly. In the basis

$$\left\{ e^{ihm/2\cdot r} \varphi_*, e^{-ihm/2\cdot r} \varphi_* \right\}$$

this operator is represented by the matrix

$$\begin{pmatrix} 1 & \lambda_{2m} \\ \lambda_{2m} & 1 \end{pmatrix}.$$

The eigenvalues of this matrix are $1 \pm |\lambda_{2m}|$, and thus

$$\left\langle \hat{\phi}_m | Q_m | \hat{\phi}_m \right\rangle \leq (1 + |\lambda_{2m}|) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)}.$$

First, suppose that $\phi = P^\perp \phi$. Then $\hat{\phi}_m \perp \varphi_*$ for all m. We then have

$$\begin{aligned}
&\left\langle \phi \left| \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) \phi \right. \right\rangle \\
&\qquad \geq \kappa \left(\|\phi\|^2_{\mathcal{L}^2_h} - |\langle \varphi_* | \phi_0 \rangle|^2 \right) - \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0}} \left\langle \hat{\phi}_m \left| Q_m \right| \hat{\phi}_m \right\rangle \\
&\qquad \geq \kappa \|\phi_0\|^2_{L^2(\mathbb{R}^2)} + \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0}} \left(1 + \lambda_m^2 - \left|\lambda_{2m} - \lambda_m^2\right|\right) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)}.
\end{aligned}$$

It is easy to check that

$$\min_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0}} \left(1 + \lambda_m^2 - \left|\lambda_{2m} - \lambda_m^2\right|\right) \geq c,$$

where c is a positive constant. Consequently,

$$\left\langle \phi \left| P^\perp \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P^\perp \phi \right. \right\rangle \geq \frac{\kappa}{2} \|\phi\|^2_{\mathcal{L}^2_h},$$

which proves (6.23).

In our second case, we suppose that $\phi = P_\varepsilon^\perp \phi$. In this case we have $\hat{\phi}_m \perp \varphi_*$ if $h|m| \leq \varepsilon$, but this may not hold if $h|m| > \varepsilon$. This means

$$\begin{aligned}
&\left\langle \phi \left| \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) \phi \right. \right\rangle \\
&\geq \kappa \left(\|\phi\|^2_{\mathcal{L}^2_h} - |\langle \varphi_* | \phi_0 \rangle|^2 \right) - \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0, h|m| \leq \varepsilon}} \left(1 - \lambda_m^2 + \left|\lambda_{2m} - \lambda_m^2\right|\right) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)} \\
&\qquad\qquad - \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0, h|m| > \varepsilon}} (1 + |\lambda_{2m}|) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)} \\
&= \kappa \|\phi_0\|^2_{L^2(\mathbb{R}^2)} + \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0, h|m| \leq \varepsilon}} \left(1 + \lambda_m^2 - \left|\lambda_{2m} - \lambda_m^2\right|\right) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)} \\
&\qquad\qquad + \frac{\kappa}{2} \sum_{\substack{m \in (2\pi\mathbb{Z})^2 \\ m \neq 0, h|m| > \varepsilon}} (1 - |\lambda_{2m}|) \left\| \hat{\phi}_m \right\|^2_{L^2(\mathbb{R}^2)}.
\end{aligned}$$

The last step is to understand the last term in the above estimate, that is the behavior of $\max_{h|m|>\varepsilon} |\lambda_{2m}|$ with respect to h. Since $|\lambda_{2m}|$ attains its maximum at $m = 0$, we see that if h is small enough, then it suffices to consider the case m of the form μe_k, where the e_k are the standard basis of $\mathbb{R}^2$ and $\varepsilon < h\mu < 2\varepsilon$. If h is small enough, we then have that $\lambda_{2m} > 0$ and also

$$1-\lambda_{2m} \geq \frac{1}{2}\int_{|r|\leq 1/2} \mathrm{d}r\ (hm\cdot r)^2 \left|\varphi_*(r)\right|^2 \geq \frac{\varepsilon^2}{2}\min_{k=1,2}\int_{|r|\leq 1/2} \mathrm{d}r\, r_k^2 \left|\varphi_*(r)\right|^2,$$

where r_k denotes the k-th component of the vector r. We have thus shown that

$$\left\langle \phi \,\middle|\, P_\varepsilon^\perp \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp \phi \right\rangle \geq \frac{\kappa}{2} C\varepsilon^2 \left\|\phi\right\|^2_{\mathcal{L}^2_h},$$

for a positive constant C which proves (6.24). □

6.5. Proof of Theorem A

The first step in proving Theorem A is to prove estimates on $P_\varepsilon^\perp \phi_h$. For convenience we introduce $\Psi_h \in L^2(Q_h)$ to be such that $\varphi_* \Psi_h = P_\varepsilon \phi_h$. Note that due to the cut-off, Ψ_h is in fact smooth, and we also have the bound

$$\|\Delta\Psi_h\|^2_{L^2(Q_h)} = \sum_{\substack{m\in(2\pi\mathbb{Z})^2\\ h|m|\leq\varepsilon}} h^4|m|^4 \left|\hat{\Psi}(m)\right|^2 \leq \varepsilon \sum_{\substack{m\in(2\pi\mathbb{Z})^2\\ h|m|\leq\varepsilon}} h^2|m|^2 \left|\hat{\Psi}(m)\right|^2$$
$$= \varepsilon \left\|\nabla\Psi_h\right\|_{L^2(Q_h)} \lesssim \varepsilon h^2.$$

The cut-off also leads to a bound on $\|\Psi_h\|_{L^\infty}$ as follows. Write $\Psi_h = \widetilde{\Psi}_h + \hat{\Psi}_h(0)$. Note that this means

$$\hat{\widetilde{\Psi}}_h(0) = 0.$$

We also note that $|\hat{\Psi}_h(0)| \leq \|\Psi_h\|_{L^2(Q_h)} \lesssim h$. Now, for any $X \in \mathbb{R}^2$,

$$|\widetilde{\Psi}_h(X)| \leq \sum_{\substack{m\in(2\pi\mathbb{Z})^2\\ 0<h|m|\leq\varepsilon}} |\hat{\Psi}(m)| \leq \left(\sum_{\substack{m\in(2\pi\mathbb{Z})^2\\ 0<h|m|\leq\varepsilon}} h^{-2}|m|^{-2}\right)^{1/2} \left(\sum_{\substack{m\in(2\pi\mathbb{Z})^2\\ 0<h|m|\leq\varepsilon}} h^2|m|^2|\hat{\Psi}(m)|^2\right)^{1/2}$$
$$\lesssim h^{-1}\left(\log\frac{\varepsilon}{h}\right)^{1/2} \|\nabla\Psi_h\|_{L^2(Q_h)} \lesssim h\left(\log\frac{\varepsilon}{h}\right)^{1/2}.$$

We therefore see that

$$\|\Psi_h\|_{L^\infty} \lesssim h\left(\log\frac{\varepsilon}{h}\right)^{1/2}.$$

We begin with the following lemma, which directly implies the first claim in Theorem A.

Lemma 6.14. *Suppose that ϕ_h, A_h, T_h are as in Theorem A. Then*

$$\left\|P_\varepsilon^\perp \phi_h\right\|_{\mathcal{L}^2_h} \lesssim \varepsilon^{-2}h^3, \tag{6.26}$$

$$\left\|\chi_\varepsilon P^\perp \phi_h\right\|_{\mathcal{L}^2_h} \lesssim \varepsilon h^2, \tag{6.27}$$

where, as before, $\varepsilon = h^{17/48}$.

Proof. For convenience let $\mathcal{L}_{T,A} := 1 - V^{1/2} L_{T,A} V^{1/2}$. By Lemma 6.13, the operator $P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp$ is invertible on $P_\varepsilon^\perp \mathcal{L}^2_{h,\mathrm{sym}}$ for sufficiently small h, and we have that

$$\| \left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} \|_{\mathcal{L}^2_h} \lesssim \varepsilon^{-2}.$$

Now, since (ϕ_h, A_h) is a solution of the BdG equations, we have that

$$P_\varepsilon^\perp \phi_h = - \left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp \left(\mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right). \quad (6.28)$$

We first estimate the term $\mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h$. In order to do this, we write

$$\mathcal{L}_{T_h,A_h} = (1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}) - V^{1/2}(L_{T_h,A_h} - L_{\mathcal{T}_c}) V^{1/2}.$$

By Lemma 6.12, we have

$$\left\|\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h\right\|_{\mathcal{L}^2_h} \lesssim \|\Delta \Psi_h\|_{L^2(Q_h)} \lesssim \varepsilon h^2,$$

and by Lemmas 6.9, 6.10 and 6.11, we see that

$$\begin{aligned}
&\left\|V^{1/2}\left(L_{T_h,A_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon \phi_h\right\|_{\mathcal{L}^2_h} \\
&\le \left\|V^{1/2}\left(L_{T_h,A_h} - M_{T_h,A_h}\right) V^{1/2} P_\varepsilon \phi_h\right\|_{\mathcal{L}^2_h} + \left\|V^{1/2}\left(M_{T_h,A_h} - L_{T_h}\right) V^{1/2} P_\varepsilon \phi_h\right\|_{\mathcal{L}^2_h} \\
&\qquad + \left\|V^{1/2}\left(L_{T_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon \phi_h\right\|_{\mathcal{L}^2_h} \\
&\lesssim \left(\|D^2 A_h\|_{L^\infty} + \|D A_h\|^2_{L^\infty} + \|D A_h\|_{L^\infty} + \|A_h\|^2_{L^\infty} + |T_h - \mathcal{T}_c|\right) \|P_\varepsilon \phi_h\|_{\mathcal{L}^2_h} \\
&\lesssim (h^3 + h^4 + h^2) h \\
&\lesssim h^3.
\end{aligned}$$

For the non-linear term, that is, the term involving N_{T_h,A_h}, we simply calculate that

$$\begin{aligned}
&\left\|V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \Big\| (z_n - k_A)^{-1} V^{1/2} \phi_h \left(z_n + \overline{k}_A\right)^{-1} V^{1/2} \overline{\phi}_h (z_n - k_A)^{-1} V^{1/2} \phi_h \\
&\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}\phi_h\right)\right)^{-1}_{22} \Big\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \left\|(z_n - k_A)^{-1}\right\|^2_\infty \left\|\left(z_n + \overline{k}_A\right)^{-1}\right\|_\infty \left\|(z_n - H_A(\Delta))^{-1}_{22}\right\|_\infty \left\|V^{1/2}\phi_h\right\|^3_{\mathcal{L}^6}.
\end{aligned}$$

From every factor containing z_n we get an $|n|^{-1}$ and the first part of Lemma 6.2 implies

$$\left\|V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right\|_{\mathcal{L}^2_h} \lesssim \sum_{n \text{ odd}} |n|^{-4} \left(\|\phi_h\|_{\mathcal{L}^2_h} + h^{-1} \|\nabla_X \phi_h\|_{\mathcal{L}^2_h}\right)^3 \lesssim h^3.$$

We now write

$$\begin{aligned}
&\left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} \\
&= \left(P_\varepsilon^\perp \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} + \left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp V^{1/2}\left(L_{T_h,A_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon^\perp \\
&\qquad \times \left(P_\varepsilon^\perp \left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}.
\end{aligned}$$

For the second term we simply have

$$\left\|\left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp V^{1/2}\left(L_{T_h,A_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon^\perp \left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}\right\|_{\mathcal{L}_h^2}$$
$$\lesssim \varepsilon^{-2} \cdot h^2 \cdot \varepsilon^{-2} \lesssim \varepsilon^{-4} h^2,$$

and therefore

$$\Big\|\left(P_\varepsilon^\perp \mathcal{L}_{T_h,A_h} P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp V^{1/2}\left(L_{T_h,A_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon^\perp \left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}$$
$$\times P_\varepsilon^\perp\left(\mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right)\Big\|_{\mathcal{L}_h^2}$$
$$\lesssim \varepsilon^{-4} h^2 \cdot \varepsilon h^2 \lesssim \varepsilon^{-3} h^4.$$

We now turn to the first, that is, to $\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}$. To begin with, we see that

$$\Big\|\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}$$
$$\times P_\varepsilon^\perp\left(V^{1/2}\left(L_{T_h,A_h} - L_{\mathcal{T}_c}\right) V^{1/2} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right)\Big\|_{\mathcal{L}_h^2}$$
$$\lesssim \varepsilon^{-2} \cdot h^3.$$

It remains to estimate the term involving $(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2})$ and here we will us the fact that this operator commutes with the cut-off χ_ε. Indeed, we have that

$$\begin{aligned}
&P_\varepsilon^\perp\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h \\
&\quad = \chi_\varepsilon P^\perp\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} \chi_\varepsilon P^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h \\
&\quad = \chi_\varepsilon P^\perp\left(P^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P^\perp\right)^{-1} \chi_\varepsilon P^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h,
\end{aligned}$$

and therefore

$$\left\|P_\varepsilon^\perp\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h\right\|_{L^2(Q_h)}$$
$$\lesssim \left\|\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon \phi_h\right\|_{L^2(Q_h)} \lesssim \varepsilon h^2.$$

Since $\varepsilon^{-3} h^4 \lesssim \varepsilon^{-2} h^3$, and $\varepsilon h^2 \lesssim \varepsilon^{-2} h^3$, we see that we have established (6.26).

We now turn to the proof of (6.27). Note that $\chi_\varepsilon P^\perp = \chi_\varepsilon P_\varepsilon^\perp$, and therefore it is clear that we only need to show that

$$\Big\|\chi_\varepsilon P^\perp\left(P_\varepsilon^\perp\left(1 - V^{1/2} L_{\mathcal{T}_c} V^{1/2}\right) P_\varepsilon^\perp\right)^{-1}$$
$$\times P_\varepsilon^\perp\left(\mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right)\Big\|_{\mathcal{L}_h^2} \lesssim \varepsilon h^2.$$

But since χ_ε commutes with $V^{1/2}L_{\mathcal{T}_c}V^{1/2}$, we have

$$\begin{aligned}\chi_\varepsilon P^\perp \left(P_\varepsilon^\perp \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} P_\varepsilon^\perp \\ = \chi_\varepsilon P^\perp \left(P_\varepsilon^\perp \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} \chi_\varepsilon P^\perp \\ = \chi_\varepsilon P^\perp \left(P^\perp \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right) P^\perp\right)^{-1} \chi_\varepsilon P^\perp,\end{aligned}$$

which means

$$\begin{aligned}&\Big\| \chi_\varepsilon P^\perp \left(P_\varepsilon^\perp \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right) P_\varepsilon^\perp\right)^{-1} \\ &\qquad \times P_\varepsilon^\perp \left(\mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right)\Big\|_{\mathcal{L}_h^2} \\ &\lesssim \left\| \left(P^\perp \left(1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right) P^\perp\right)^{-1}\right\|_{\mathcal{L}_h^2} \left\| \mathcal{L}_{T_h,A_h} P_\varepsilon \phi_h + \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}\phi_h\right)\right\|_{\mathcal{L}_h^2} \\ &\lesssim \varepsilon h^2 + h^3 \lesssim h^3.\end{aligned}$$

That completes the proof of the lemma. □

We now wish to prove an estimate on the terms in the Bogoliubov-de Gennes equations which involve $P_\varepsilon^\perp \phi$. We have the following lemma.

LEMMA 6.15. *For all T sufficiently close to $\mathcal{T}_c$, we have the estimates*

$$\begin{aligned}&\left\| P_\varepsilon F_T^{\mathrm{BCS}}(\phi, A) - P_\varepsilon F_T^{\mathrm{BCS}}(P_\varepsilon \phi, A)\right\|_{\mathcal{L}_h^2} \\ &\lesssim \varepsilon^2 \left\|\chi_\varepsilon P^\perp \phi\right\|_{\mathcal{L}_h^2} + \left(\left\|D^2 A\right\|_{L^\infty} + \|DA\|_{L^\infty}^2 + \|DA\|_{L^\infty} + \|A\|_{L^\infty}^2 + |T - \mathcal{T}_c|\right) \left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2} \\ &\quad + h^{-1}\left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}_h^2}\right)^3 \left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2} \\ &\quad + h^{-2/3}\left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}_h^2}\right)^2 \left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2} \\ &\quad + h^{-4/3}\left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}_h^2}\right) \left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2}^2 + h^{-2}\left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2}^3, \qquad (6.29)\end{aligned}$$

and

$$\begin{aligned}&\left\| G_T^{\mathrm{BCS}}(\phi, A) - G_T^{\mathrm{BCS}}(P_\varepsilon \phi, A)\right\|_{L^2(Q_h)} \\ &\quad \lesssim \left(1 + \|A\|_{L^\infty}\right) \left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2} \\ &\qquad \times \left(h^{-2/3}\left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}_h^2}\right)^2 + \|\Psi\|_{L^\infty} + h^{-1}\left\|P_\varepsilon^\perp \phi\right\|_{\mathcal{L}_h^2}\right), \qquad (6.30)\end{aligned}$$

where Ψ is such that $\varphi_* \Psi = P_\varepsilon \phi$.

PROOF. We begin with the proof of (6.29). We note that

$$\begin{aligned}P_\varepsilon F_T^{\mathrm{BCS}}(\phi, A) - P_\varepsilon F_T^{\mathrm{BCS}}(P_\varepsilon \phi, A) = P_\varepsilon \left(1 - V^{1/2}L_{T,A}V^{1/2}\right) P_\varepsilon^\perp \phi \\ + P_\varepsilon V^{1/2} N_{T,A}\left(-2V^{1/2}\phi\right) - P_\varepsilon V^{1/2} N_{T,A}\left(-2V^{1/2}P_\varepsilon \phi\right).\end{aligned}$$

We will first estimate the linear term, and we again write

$$1 - V^{1/2}L_{T,A}V^{1/2} = 1 - V^{1/2}L_{\mathcal{T}_c}V^{1/2} - V^{1/2}(L_{T,A} - L_{\mathcal{T}_c})V^{1/2}.$$

For the second term we have the simple estimate

$$\begin{aligned}&\left\|V^{1/2}\left(L_{T,A}-L_{\mathcal{T}_c}\right)V^{1/2}P_\varepsilon^\perp\phi\right\|_{\mathcal{L}_h^2}\\&\qquad\lesssim\left(\|D^2A\|_{L^\infty}+\|DA\|_{L^\infty}^2+\|DA\|_{L^\infty}+\|A\|_{L^\infty}^2+|T-\mathcal{T}_c|\right)\left\|P_\varepsilon^\perp\phi\right\|_{\mathcal{L}_h^2}.\end{aligned}$$

For the first term, we start by noting that

$$P_\varepsilon\left(1-V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)P_\varepsilon^\perp\phi=P_\varepsilon\left(1-V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)\chi_\varepsilon P^\perp\phi.$$

We now let $\xi:=\chi_\varepsilon P^\perp\phi$ and define $\hat{\xi}_m:\mathbb{R}^2\to\mathbb{C}$ by the Fourier series

$$\xi(\zeta_X^r,\zeta_X^{-r})=\sum_{m\in(2\pi\mathbb{Z})^2}e^{ihm\cdot X}\hat{\xi}_m(r),$$

where we note that $\hat{\xi}_m=0$ if $h|m|>\varepsilon$. We then see that

$$\begin{aligned}&PV^{1/2}L_{\mathcal{T}_c}V^{1/2}\xi(\zeta_X^r,\zeta_X^{-r})\\&=\varphi_*(r)\int_{\mathbb{R}^6}\mathrm{d}s\mathrm{d}t\mathrm{d}Y\ (V\alpha_*)(s)V^{1/2}(t)F_{\mathcal{T}_c}(Y,s-t)e^{iY\cdot(-i\nabla_X)}\xi(\zeta_X^t,\zeta_X^{-t})\\&=\frac{1}{(2\pi)^4}\int_{\mathbb{R}^{10}}\mathrm{d}s\mathrm{d}t\mathrm{d}Y\mathrm{d}p\mathrm{d}q\,\varphi_*(r)\left(V\alpha_*\right)(s)V^{1/2}(t)f_{\mathcal{T}_c}\left(p+q/2,p-q/2\right)e^{ip\cdot(s-t)}e^{iq\cdot Y}\\&\qquad\qquad\times\sum_{m\in(2\pi\mathbb{Z})^2}e^{ihm\cdot X}e^{ihm\cdot Y}\hat{\xi}_m(r)\\&=\frac{1}{(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}\int_{\mathbb{R}^6}\mathrm{d}s\mathrm{d}t\mathrm{d}p\,e^{ip\cdot(s-t)}\varphi_*(r)\left(V\alpha_*\right)(s)V^{1/2}(t)\\&\qquad\qquad\times f_{\mathcal{T}_c}\left(p+hm/2,p-hm/2\right)\hat{\xi}_m(r).\end{aligned}$$

On the other hand, since $\varphi_*=V^{1/2}K_{\mathcal{T}_c}^{-1}V^{1/2}\varphi_*$, we see that

$$\begin{aligned}&P\xi\left(\zeta_X^r,\zeta_X^{-r}\right)\\&=\varphi_*(r)\int_{\mathbb{R}^2}\mathrm{d}t\,\varphi_*(t)\xi(\zeta_X^t,\zeta_X^{-t})\\&=\frac{1}{(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}e^{ihm\cdot X}\int_{\mathbb{R}^6}\mathrm{d}p\mathrm{d}s\mathrm{d}t\,e^{ip\cdot(t-s)}\varphi_*(r)V^{1/2}(r)\left(V\alpha_*\right)(s)f_{\mathcal{T}_c}(p,p)\xi(\zeta_X^t,\zeta_X^{-t})\\&=\frac{1}{(2\pi)^2}\sum_{m\in(2\pi\mathbb{Z})^2}e^{ihm\cdot X}\int_{\mathbb{R}^6}\mathrm{d}p\mathrm{d}s\mathrm{d}t\,e^{ip\cdot(s-t)}\varphi_*(r)V^{1/2}(r)\left(V\alpha_*\right)(s)f_{\mathcal{T}_c}(p,p)\xi(\zeta_X^t,\zeta_X^{-t}).\end{aligned}$$

Using the inequality $|f_{\mathcal{T}_c}(p,p)-f_{\mathcal{T}_c}\left(p+hm/2,p-hm/2\right)|\lesssim\min\{1,h^2|m|^2\}$, where the right-hand side decays exponentially in p, we then see that

$$\left\|P_\varepsilon\left(1-V^{1/2}L_{\mathcal{T}_c}V^{1/2}\right)P_\varepsilon^\perp\phi\right\|_{\mathcal{L}_h^2}^2\lesssim\int_{\mathbb{R}^2}\mathrm{d}r\sum_{m\in(2\pi\mathbb{Z})^2}h^4|m|^4|\hat{\xi}_m(r)|^2\le\varepsilon^4\left\|\chi_\varepsilon P^\perp\phi\right\|_{\mathcal{L}_h^2}^2,$$

and that completes the estimate for the linear term.

We now estimate the difference of the non-linear terms. For convenience we define $\xi = P_\varepsilon^\perp \phi$. We start by writing

$$
\begin{aligned}
&N_{T,A}\left(-2V^{1/2}\phi\right) \\
&= -\frac{16}{\beta}\sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2}\phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\phi}\,(z_n - k_A)^{-1} V^{1/2}\phi \\
&\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{22}^{-1} \qquad (6.31) \\
&\quad -\frac{16}{\beta}\sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2}\phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\phi}\,(z_n - k_A)^{-1} V^{1/2}\phi \\
&\qquad\qquad \times \left(\left(z_n - H_A\left(-2V^{1/2}\phi\right)\right)_{22}^{-1} - \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{22}^{-1}\right). \qquad (6.32)
\end{aligned}
$$

For the last factor of the second series, we note that

$$
\begin{aligned}
&\left\|\left(z - H_A\left(-2V^{1/2}\phi\right)\right)_{22}^{-1} - \left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{22}^{-1}\right\|_\infty \\
&= 2\left\|\left(z - H_A\left(-2V^{1/2}\phi\right)\right)_{21}^{-1} V^{1/2}\xi \left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{22}^{-1}\right. \\
&\qquad\qquad \left. + \left(z - H_A\left(-2V^{1/2}\phi\right)\right)_{22}^{-1} \overline{V^{1/2}\xi}\left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{12}^{-1}\right\|_\infty \\
&\lesssim \left\|V^{1/2}\xi\right\|_\infty,
\end{aligned}
$$

which means we can estimate the second part of $N_{T,A}$ as follows.

$$
\begin{aligned}
\|\,(6.32)\,\|_{\mathcal{L}_h^2} &\lesssim \sum_{n \text{ odd}} |n|^{-3}\left\|V^{1/2}\phi\right\|_{\mathcal{L}^6}^3 \left\|V^{1/2}\xi\right\|_\infty \\
&\lesssim h^{-1}\left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\|\nabla_X\phi\|_{\mathcal{L}_h^2}\right)^3 \|\xi\|_{\mathcal{L}_h^2}.
\end{aligned}
$$

We note that we have here used the fact that if T is sufficiently close to $\mathcal{T}_c$, then $1/\beta = T \lesssim 1$. We can therefore turn to the first term, that is, to (6.31), which we write in the form

$$
\begin{aligned}
&-\frac{16}{\beta}\sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2}\left(P_\varepsilon\phi + \xi\right)\left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\left(P_\varepsilon\overline{\phi} + \overline{\xi}\right) \\
&\qquad\qquad \times (z_n - k_A)^{-1} V^{1/2}\left(P_\varepsilon\phi + \xi\right)\left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{22}^{-1}.
\end{aligned}
$$

The difference between this term and $N_{T,A_h}(-2V^{1/2}P_\varepsilon\phi)$ then consists of seven terms. There are three terms containing a single ξ and these can be estimated as follows,

$$\begin{aligned}
&\Bigg\| \sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2} P_\varepsilon \phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2} P_\varepsilon \overline{\phi} \left(z_n - k_A\right)^{-1} V^{1/2} \xi \\
&\qquad\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22} \Bigg\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \left\|(z_n - k_A)^{-1}\right\|^2_\infty \left\|\left(z_n + \overline{k}_A\right)^{-1}\right\|_\infty \left\|\left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22}\right\|_\infty \\
&\qquad\qquad\qquad \times \left\|V^{1/2}P_\varepsilon\phi\right\|^2_{\mathcal{L}^6_h} \left\|V^{1/2}\xi\right\|_{\mathcal{L}^6_h} \\
&\lesssim h^{-2/3} \left(\|\phi\|_{\mathcal{L}^2_h} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}^2_h}\right)^2 \|\xi\|_{\mathcal{L}^2_h}.
\end{aligned}$$

There are also three terms containing two ξ and we estimate these terms as follows,

$$\begin{aligned}
&\Bigg\| \sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2} P_\varepsilon \phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2} \overline{\xi} \left(z_n - k_A\right)^{-1} V^{1/2} \xi \\
&\qquad\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22} \Bigg\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \left\|(z_n - k_A)^{-1}\right\|^2_\infty \left\|\left(z_n + \overline{k}_A\right)^{-1}\right\|_\infty \left\|\left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22}\right\|_\infty \\
&\qquad\qquad\qquad \times \left\|V^{1/2}P_\varepsilon\phi\right\|_{\mathcal{L}^6_h} \left\|V^{1/2}\xi\right\|^2_{\mathcal{L}^6_h} \\
&\lesssim h^{-4/3} \left(\|\phi\|_{\mathcal{L}^2_h} + h^{-1}\|\nabla_X \phi\|_{\mathcal{L}^2_h}\right) \|\xi\|^2_{\mathcal{L}^2_h}.
\end{aligned}$$

Finally there is the term with three ξ. For this term we have

$$\begin{aligned}
&\Bigg\| \sum_{n \text{ odd}} (z_n - k_A)^{-1} V^{1/2} \xi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2} \overline{\xi} \left(z_n - k_A\right)^{-1} V^{1/2} \xi \\
&\qquad\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22} \Bigg\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \left\|(z_n - k_A)^{-1}\right\|^2_\infty \left\|\left(z_n + \overline{k}_A\right)^{-1}\right\|_\infty \left\|\left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)^{-1}_{22}\right\|_\infty \left\|V^{1/2}\xi\right\|^3_{\mathcal{L}^6_h} \\
&\lesssim h^{-6/3} \|\xi\|^3_{\mathcal{L}^2_h}.
\end{aligned}$$

That establishes (6.29).

We now turn to the proof of (6.30). Recall that $J_{T,A}(\phi) = \operatorname{Re} \pi_A \gamma_A(\phi)|_{y=x}$. By expanding $(z - H_A(\phi))^{-1}_{11}$, we see that

$$J_{T,A}(\phi) = \mathcal{J}_{T,A} + \frac{1}{\beta} \sum_{n \text{ odd}} \operatorname{Re} \pi_A \left(z_n - k_A\right)^{-1} \phi \left(z_n + \overline{k}_A\right)^{-1} \overline{\phi} (z_n - H_A(\phi))^{-1}_{11} \Bigg|_{y=x},$$

where

$$\mathcal{J}_{T,A_h} = \operatorname{Re} \pi_A \left(1 + e^{\beta k_A}\right)^{-1}\Big|_{y=x} .$$

This means that

$$\begin{aligned} G_T^{\mathrm{BCS}}(\phi, A) - G_T^{\mathrm{BCS}}(P_\varepsilon \phi, A) &= \frac{4}{\beta} \sum_{n \text{ odd}} \operatorname{Re}(-i\nabla + A) \\ &\Big((z_n - k_A)^{-1} V^{1/2}\phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\phi}(z_n - H_A\left(-2V^{1/2}\phi\right))_{11}^{-1} \\ &- (z_n - k_A)^{-1} V^{1/2} P_\varepsilon\phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{P_\varepsilon\phi}(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right))_{11}^{-1}\Big)\Big|_{y=x} . \end{aligned}$$

We begin by estimating the terms involving A and we will study the terms with $(-i\nabla)$ afterwards. We again define $\xi = P_\varepsilon^\perp \phi$. We now estimate the error in replacing $(z_n - H_A\left(-2V^{1/2}\phi\right))_{11}^{-1}$ by $(z_n - H_A(-2V^{1/2}P_\varepsilon\phi))_{11}^{-1}$. To do this, we first note that

$$\begin{aligned} &\left\| \left(z - H_A\left(-2V^{1/2}\phi\right)\right)_{11}^{-1} - \left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{11}^{-1} \right\|_{\mathcal{L}_h^6} \\ &= 2 \Big\| \left(z - H_A\left(-2V^{1/2}\phi_h\right)\right)_{11}^{-1} V^{1/2}\xi_h \left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi_h\right)\right)_{21}^{-1} \\ &\qquad + \left(z - H_A\left(-2V^{1/2}\phi\right)\right)_{12}^{-1} \overline{V^{1/2}\xi} \left(z - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{11}^{-1} \Big\|_{\mathcal{L}_h^6} \\ &\lesssim \left\|V^{1/2}\xi\right\|_{\mathcal{L}_h^6} \\ &\lesssim h^{-2/3} \left\|\xi\right\|_{\mathcal{L}_h^2} . \end{aligned}$$

Applying Lemma 6.4, we see that we can estimate

$$\begin{aligned} &\Big\| \sum_{n \text{ odd}} \operatorname{Re} A \left(z_n - k_A\right)^{-1} V^{1/2}\phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\phi} \\ &\qquad \times \left(\left(z_n - H_A\left(-2V^{1/2}\phi\right)\right)_{11}^{-1} - \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{11}^{-1} \right)\Big|_{y=x} \Big\|_{L^2(Q_h)} \\ &\lesssim \sum_{n \text{ odd}} \left\|K_n^{(1)}\left(1 - \Delta\right)\right\|_{\mathcal{L}_h^2} , \end{aligned}$$

where $K_n^{(1)}$ denotes the n-th summand in the series giving the operator in question. We proceed by making use of Hölder's inequality (6.4) and see that

$$\begin{aligned} \sum_{n \text{ odd}} \left\|K_n^{(1)}\left(1 - \Delta\right)\right\|_{\mathcal{L}_h^2} &\lesssim \|A\|_{L^\infty} \cdot \|V^{1/2}\phi\|_{\mathcal{L}_h^6} \cdot h^{-2/3} \left\|\xi\right\|_{\mathcal{L}^2} \\ &\lesssim h^{-2/3}\|A\|_{L^\infty} \left(\|\phi\|_{\mathcal{L}_h^2} + h^{-1}\left\|\nabla_X\phi\right\|_{\mathcal{L}_h^2}\right)^2 \left\|\xi\right\|_{\mathcal{L}^2} . \end{aligned}$$

We now write $\phi = P_\varepsilon\phi + \xi$ and expand. This gives us three terms to estimate. There are two terms containing a single ξ and these can be estimated as follows. We have

$$\begin{aligned}
&\Bigg\| \sum_{n \text{ odd}} A\left(z_n - k_A\right)^{-1} V^{1/2} P_\varepsilon \phi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\xi} \\
&\qquad\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{11}^{-1}\Big|_{y=x}\Bigg\|_{L^2(Q_h)} \\
&\lesssim \sum_{n \text{ odd}} \left\| K_n^{(2)}\left(1-\Delta\right)\right\|_{\mathcal{L}_h^2},
\end{aligned}$$

where we again applied Lemma 6.4 and where $K_n^{(2)}$ denotes the n-th summand in the series giving the operator in question, analogously to the notation used in the previous calculations. As before we apply Hölder's inequality (6.4) and conclude that

$$\begin{aligned}
\sum_{n \text{ odd}} \left\| K_n^{(2)}\left(1-\Delta\right)\right\|_{\mathcal{L}_h^2} &\lesssim \sum_{n \text{ odd}} |n|^{-3} \|A\|_{L^\infty} \left\|V^{1/2}P_\varepsilon\phi\right\|_\infty \left\|V^{1/2}\xi\right\|_{\mathcal{L}_h^2} \\
&\lesssim \|A\|_{L^\infty}\|\Psi\|_{L^\infty}\|\xi\|_{\mathcal{L}_h^2}.
\end{aligned}$$

For the term containing two ξ, we proceed similarly and get

$$\begin{aligned}
&\Bigg\| \sum_{n \text{ odd}} A\left(z_n - k_A\right)^{-1} V^{1/2} \xi \left(z_n + \overline{k}_A\right)^{-1} V^{1/2}\overline{\xi} \\
&\qquad\qquad\qquad \times \left(z_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)\right)_{11}^{-1}\Big|_{y=x}\Bigg\|_{L^2(Q_h)} \\
&\lesssim \sum_{n \text{ odd}} \left\| K_n^{(3)}\left(1-\Delta\right)\right\|_{\mathcal{L}_h^2} \\
&\lesssim \sum_{n \text{ odd}} |n|^{-3} \|A\|_{L^\infty} \left\|V^{1/2}\xi\right\|_{\mathcal{L}_h^4}^2 \\
&\lesssim h^{-1}\|A\|_{L^\infty}\|\xi\|_{\mathcal{L}_h^2}^2.
\end{aligned}$$

This finishes the estimates for the terms involving A. For the terms involving $-i\nabla$, we can in fact use the same estimates, but, of course, without the $\|A\|_{L^\infty}$ factor. That completes the proof of the lemma. □

We can now prove the second part of Theorem A. Using Lemma 6.14 and that fact that ϕ_h and A_h are solutions, we see that

$$\begin{aligned}
&\left\|P_\varepsilon F_{T_h}^{\mathrm{BCS}}(P_\varepsilon\phi_h, A_h)\right\|_{\mathcal{L}_h^2} \\
&= \left\|P_\varepsilon F_{T_h}^{\mathrm{BCS}}(\phi_h, A_h) - P_\varepsilon F_{T_h}^{\mathrm{BCS}}(P_\varepsilon\phi_h, A_h)\right\|_{\mathcal{L}_h^2} \\
&\lesssim \varepsilon^2\cdot\varepsilon h^2 + h^2\cdot\varepsilon^{-2}h^3 + h^{-1}\cdot h^3\cdot\varepsilon^{-2}h^3 + h^{-2/3}\cdot h^2\cdot\varepsilon^{-2}h^3 + h^{-4/3}\cdot h\cdot\varepsilon^{-4}h^6 \\
&\qquad\qquad\qquad + h^{-6/3}\cdot\varepsilon^{-6}h^9 \\
&\lesssim h^3\varepsilon^3h^{-1}.
\end{aligned}$$

And also

$$
\begin{aligned}
\left\|G_{T_h}^{\mathrm{BCS}}\left(P_{\varepsilon}\phi_h, A_h\right)\right\|_{L^2(Q_h)} &= \left\|G_{T_h}^{\mathrm{BCS}}\left(\phi_h, A_h\right) - G_{T_h}^{\mathrm{BCS}}\left(P_{\varepsilon}\phi_h, A_h\right)\right\|_{L^2(Q_h)} \\
&\lesssim \varepsilon^{-2}h^3\left(h^{-2/3}\cdot h^2 + h\left(\log\frac{\varepsilon}{h}\right)^{1/2} + h^{-1}\cdot\varepsilon^{-2}h^3\right) \\
&\lesssim h^3\varepsilon^{-2}h\left(\log\frac{\varepsilon}{h}\right)^{1/2},
\end{aligned}
$$

which finishes the proof of Theorem A.

6.6. Proof of Theorem B

In order to proof Theorem B it turns out to be convenient to introduce the following notation for the non-linear maps. We will write $N_{T,A} = \widetilde{N}_{T,A} + N'_{T,A}$, where

$$
\widetilde{N}_{T,A}(\alpha) = \frac{2}{\beta}\sum_{n\text{ odd}}\left(z_n - k_A\right)^{-1}\alpha\left(z_n + \overline{k}_A\right)^{-1}\overline{\alpha}\left(z_n - k_A\right)^{-1}\alpha\left(z_n + \overline{k}_A\right)^{-1},
$$

and

$$
\begin{aligned}
N'_{T,A}(\alpha) = \frac{2}{\beta}\sum_{n\text{ odd}}&\left(z_n - k_A\right)^{-1}\alpha\left(z_n + \overline{k}_A\right)^{-1}\overline{\alpha}\left(z_n - k_A\right)^{-1}\alpha \\
&\times\left(z_n + \overline{k}_A\right)^{-1}\overline{\alpha}\left(z - H_A(\alpha)\right)^{-1}_{12}.
\end{aligned}
$$

Similarly we will write $J_{T,A} = \mathcal{J}_{T,A} + \widetilde{J}_{T,A} + J'_{T,A}$, where as in the proof of (6.30),

$$
\begin{aligned}
\mathcal{J}_{T,A} &= \pi_A\left(\frac{1}{1+e^{\beta k_A}}\right)\bigg|_{y=x}, \\
\widetilde{J}_{T,A}(\alpha) &= \operatorname{Re}\pi_A\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\left(z - k_A\right)^{-1}\alpha\left(z + \overline{k}_A\right)^{-1}\overline{\alpha}\left(z - k_A\right)^{-1}\bigg|_{y=x}, \\
J'_{T,A}(\alpha) &= \operatorname{Re}\pi_A\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\left(z - k_A\right)^{-1}\alpha\left(z + \overline{k}_A\right)^{-1}\overline{\alpha} \\
&\qquad\times\left(z - k_A\right)^{-1}\alpha\left(z - H_A(\alpha)\right)^{-1}_{21}\bigg|_{y=x},
\end{aligned}
$$

In this section we show that the leading order of the Bogoliubov-de Gennes equations is indeed given by the Ginzburg-Landau equations when the coefficients are correctly defined. We first define

$$
\psi_h = \mathcal{Q}_h P_{\varepsilon}\phi_h, \quad a_h = \vec{\mathcal{Q}}_h. \tag{6.33}
$$

Note that

$$
\|\psi_h\|_{L^2(Q)} \lesssim 1, \quad \|\psi_h\|_{H^1(Q)} \lesssim 1, \quad \text{and} \quad \|\psi_h\|_{H^2(Q)} \lesssim \varepsilon h^{-1}.
$$

We have the following lemma.

LEMMA 6.16. *We have the estimates*

$$\Big\| h^{-2}\mathcal{Q}_h P_\varepsilon \left(1 - V^{1/2} L_{T_h,A_h} V^{1/2}\right) P_\varepsilon \phi_h \\ - \left((-i\nabla + 2a_h)\cdot \mathbb{B}\,(-i\nabla + 2a_h)\,\psi_h - C_1\psi_h\right)\Big\|_{L^2(Q)} \lesssim \varepsilon^3 h^{-1}, \quad (6.34)$$

$$\left\| h^{-2}\mathcal{Q}_h P_\varepsilon \frac{1}{2} V^{1/2} N_{T_h,A_h}\left(-2V^{1/2}P_\varepsilon\phi_h\right) - C_2 \left|\psi_h\right|^2 \psi_h \right\|_{L^2(Q)} \lesssim \varepsilon, \quad (6.35)$$

$$\left\| h^{-2}\vec{\mathcal{Q}}_h \operatorname{curl}^2 A_h - \frac{1}{2}\operatorname{curl}^2 2a_h \right\|_{L^1(Q)} = 0, \quad (6.36)$$

$$\Bigg\| h^{-2}\vec{\mathcal{Q}}_h J_{T,A_h}\left(-2V^{1/2}P_\varepsilon\phi_h\right) \\ - \left(-\frac{1}{24\pi\left(1+e^{-\beta_c\mu}\right)}\operatorname{curl}^2 a_h(x) + \operatorname{Re}\overline{\psi_h(x)}2\mathbb{B}\left(-i\nabla + 2a_h(x)\right)\psi_h(x)\right)\Bigg\|_{L^1(Q)} \\ \lesssim \varepsilon^{1/2} h^{1/2}, \quad (6.37)$$

where ψ_h *and* a_h *given as in* (6.33).

PROOF OF (6.34). Since $P_\varepsilon\phi_h(x,y) = h\varphi_*(x-y)\psi_h(h(x+y)/2)$, we see that

$$\begin{aligned} &V^{1/2}L_{T,A}V^{1/2}P_\varepsilon\phi\left(\zeta_X^r, \zeta_X^{-r}\right) \\ &= -h\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\int \mathrm{d}\mathbf{w}_{12}\, V^{1/2}(r) G_A^z\left(\zeta_X^r, w_1\right)\left(V\alpha_*\right)(w_1 - w_2) \\ &\qquad\qquad \times \psi_h\left(h(w_1+w_2)/2\right) G_{-A}^{-z}\left(w_2, \zeta_X^{-r}\right) \\ &= -h\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\int_{\mathbb{R}^4}\mathrm{d}Y\mathrm{d}s\, V^{1/2}(r)\left(V\alpha_*\right)(s) G_A^z\left(\zeta_X^r,\zeta_Y^s\right) G_{-A}^{-z}\left(\zeta_Y^{-s},\zeta_X^{-r}\right)\psi_h(hY). \end{aligned}$$

This means

$$\begin{aligned} &h^{-2}\mathcal{Q}_h P V^{1/2} L_{T,A} V^{1/2} P_\varepsilon \phi(X) \\ &= -h^{-2}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\int_{\mathbb{R}^6}\mathrm{d}Y\mathrm{d}r\mathrm{d}s\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s) G_A^z\left(\zeta_{X/h}^r, \zeta_Y^s\right) \\ &\qquad\qquad \times G_{-A}^{-z}\left(\zeta_Y^{-s}, \zeta_{X/h}^{-r}\right)\psi_h(hY) \\ &= -h^{-2}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\int_{\mathbb{R}^6}\mathrm{d}Y\mathrm{d}r\mathrm{d}s\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s) G_A^z\left(\zeta_{X/h}^r, \zeta_{X/h+Y}^s\right) \\ &\qquad\qquad \times G_{-A}^{-z}\left(\zeta_{X/h+Y}^{-s}, \zeta_{X/h}^{-r}\right)\psi_h\left(X + hY\right). \end{aligned} \quad (6.38)$$

We now wish to replace G^z_A with K^z_A and G^{-z}_{-A} with K^{-z}_{-A}. This gives two terms to estimate. We bound the first term as follows,

$$\begin{aligned}
&\int_Q \mathrm{d}X \, \bigg| h^{-2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, (V\alpha_*)(r)(V\alpha_*)(s) K^{-z}_{-A}\left(\zeta^{-s}_{X/h+Y}, \zeta^{-r}_{X/h}\right) \\
&\qquad\qquad \times \left(G^z_A\left(\zeta^r_{X/h}, \zeta^s_{X/h+Y}\right) - K^z_A\left(\zeta^r_{X/h}, \zeta^s_{X/h+Y}\right)\right) \psi_h(X+hY) \bigg|^2 \\
&\lesssim h^{-2} \int_Q \mathrm{d}X \left(\int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, |(V\alpha_*)(r)(V\alpha_*)(s)| \right. \\
&\qquad\qquad \left. \times \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \left|G^{-z}\left(\zeta^{r-s}_Y\right)\right| H^z_A\left(\zeta^{s-r}_Y\right) |\psi_h(X+hY)| \right)^2 \\
&\lesssim h^{-2} \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, |(V\alpha_*)(r)(V\alpha_*)(s)| \\
&\qquad\qquad \times \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \left|G^{-z}\left(\zeta^{r-s}_Y\right)\right| H^z_A\left(\zeta^{s-r}_Y\right) \|\psi_h\|_{L^2(Q)} \\
&\lesssim h,
\end{aligned}$$

where we estimated the integral in the second-to-last line by M_A. The other term can be estimated similarly and we omit these technical details here. This allows us now to go back and consider (6.38) with G^z_A replaced by K^z_A. We get

$$\begin{aligned}
&- h^{-2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, (V\alpha_*)(r)(V\alpha_*)(s) K^z_A\left(\zeta^r_{X/h}, \zeta^s_{X/h+Y}\right) \\
&\qquad\qquad \times K^{-z}_{-A}\left(\zeta^{-s}_{X/h+Y}, \zeta^{-r}_{X/h}\right) \psi_h(X+hY) \\
&= -h^{-2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta^{s-r}_Y\right) G^{-z}\left(\zeta^{r-s}_Y\right) \\
&\qquad\qquad \times e^{i\Phi_A\left(\zeta^r_{X/h}, \zeta^s_{X/h+Y}\right)} e^{-i\Phi_A\left(\zeta^{-s}_{X/h+Y}, \zeta^{-r}_{X/h}\right)} e^{ihY\cdot(-i\nabla_X)} \psi_h(X) \\
&= -h^{-2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta^{s-r}_Y\right) G^{-z}\left(\zeta^{r-s}_Y\right) \\
&\quad \times e^{i\Phi_a\left(\zeta^{hr}_X, \zeta^{hs}_{X+hY}\right)} e^{-i\Phi_a\left(\zeta^{-hs}_{X+hY}, \zeta^{-hr}_X\right)} e^{-i\Phi_{2a}(X, X+hY)} e^{ihY\cdot(-i\nabla_X + 2a(X))} \psi_h(X),
\end{aligned} \tag{6.39}$$

where in the last step we made use of the operator identity (6.18). Next, we apply Lemma 6.7, which implies

$$\begin{aligned}
&\exp\left(i\Phi_a\left(\zeta^{hr}_X, \zeta^{hs}_{X+hY}\right)\right) \exp\left(-i\Phi_a\left(\zeta^{-hs}_{X+hY}, \zeta^{hr}_X\right)\right) \exp\left(-i\Phi_{2a}(X, X+hY)\right) \\
&\qquad\qquad = e^{ih^2(r-s)/4\cdot J_A(X)(r+s)} + O(h^2 M_A)(|s|^2 + |r|^2).
\end{aligned}$$

The error we get by going from (6.39) to

$$\begin{aligned}
&- h^{-2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\, \rho(\beta z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta^{s-r}_Y\right) G^{-z}\left(\zeta^{r-s}_Y\right) \\
&\qquad\qquad \times e^{ih^2(r-s)/4\cdot J_A(X)(r+s)} e^{ihY\cdot(-i\nabla_X + 2a(X))} \psi_h(X)
\end{aligned}$$

can be estimated to be $\lesssim h^{-2}h^2 M_A \|\psi_h\|_{L^2(Q)} \lesssim h^3$ by making use of Lemma 6.7. A symmetry argument shows that we are now dealing with

$$
\begin{aligned}
-h^{-2}\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta z)\int_{\mathbb{R}^6} \mathrm{d}Y\mathrm{d}r\mathrm{d}s\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s)G^z\left(\zeta_Y^{s-r}\right)G^{-z}\left(\zeta_Y^{r-s}\right)\\
\times e^{ih^2(r-s)/4\cdot J_A(X)(r+s)}\cos(hY\cdot(-i\nabla_X+2a(X)))\psi_h(X). \quad (6.40)
\end{aligned}
$$

Now, we write

$$
\rho(\beta z) = \rho\left(\beta_c z\right) + 2h^2\beta_c D z\rho'\left(\beta_c z\right) + O(h^4),
$$

where the error term decays exponentially in z. Moreover, note that

$$
e^{ih^2(r-s)/4\cdot J_A(X)(r+s)} = 1 + ih^2(r-s)/4\cdot J_A(X)(r+s) + O\left(h^4\left\|D^2A\right\|_{L^\infty}\right)\left(|s|^2+|r|^2\right)
$$

as well as

$$
\begin{aligned}
&\cos\left(hY\cdot(-i\nabla_X+2a(X))\right)\\
&\qquad = 1-\frac{1}{2}\left(hY\cdot(-i\nabla_X+2a(X))\right)^2 + g\left(hY\cdot(-i\nabla_X+2a(X))\right),
\end{aligned}
$$

where $g(x) = \cos x - (1-x^2/2)$. These facts will allow us to simplify the second line in (6.40) to $\psi_h(X)$. It is straightforward to estimate the errors, except for the one involving g. The fact that $g(x)\lesssim x^4$ allows us to show the error to be $\lesssim h^{-2}h^4\left\|\Delta^2\psi_h\right\|_{L^2(Q)} \lesssim h^2\varepsilon^3 h^{-3}\left\|\nabla\psi_h\right\|_{L^2(Q)} \lesssim \varepsilon^3 h^{-1}$, where we use the momentum cut-off to control $\Delta^2\psi_h$.

We now go back to (6.40), that is the leading term, and thus consider

$$
\begin{aligned}
&-h^{-2}\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^6} \mathrm{d}Y\mathrm{d}r\mathrm{d}s\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s)G^z\left(\zeta_Y^{s-r}\right)G^{-z}\left(\zeta_Y^{r-s}\right)\psi_h(X)\\
&= \frac{h^{-2}}{(2\pi)^4}\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^{10}} \mathrm{d}Y\mathrm{d}r\mathrm{d}s\mathrm{d}p\mathrm{d}q\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s)\\
&\qquad\qquad \times \frac{e^{ip\cdot\left(\zeta_Y^{s-r}\right)}e^{iq\cdot\left(\zeta_Y^{r-s}\right)}\psi_h(X)}{(z-(|p|^2-\mu))(z+(|q|^2-\mu))}. \quad (6.41)
\end{aligned}
$$

We integrate over Y and then over q, to see that (6.41) equals

$$
\begin{aligned}
&\frac{h^{-2}}{(2\pi)^2}\int_{\mathbb{R}^6} \mathrm{d}r\mathrm{d}s\mathrm{d}p\,\left(V\alpha_*\right)(r)\left(V\alpha_*\right)(s)e^{ip\cdot(s-r)}K_{T_c}^{-1}(p)\psi_h(X)\\
&\qquad = h^{-2}\int_{\mathbb{R}^2} \mathrm{d}s\,\left(V\alpha_*\right)(s)\alpha_*(s)\psi_h(X) = h^{-2}\mathcal{Q}_h P_\varepsilon\phi_h(X),
\end{aligned}
$$

and this cancels out the first term of the estimate in (6.34).

There are three next-to-leading terms. We start with

$$
\begin{aligned}
&-\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} 2\beta_c D z \rho'(\beta_c z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s\, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta_Y^{s-r}\right) G^{-z}\left(\zeta_Y^{r-s}\right) \psi_h(X) \\
&= \frac{2\beta_c D}{(2\pi)^4} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} z\rho'(\beta_c z) \int_{\mathbb{R}^{10}} \mathrm{d}Y \mathrm{d}r \mathrm{d}s \mathrm{d}p \mathrm{d}q\, (V\alpha_*)(r)(V\alpha_*)(s) \\
&\qquad \times \frac{e^{ip\cdot\left(\zeta_Y^{s-r}\right)} e^{iq\cdot\left(\zeta_Y^{r-s}\right)}}{\left(z-\left(|p|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)} \psi_h(X).
\end{aligned} \tag{6.42}
$$

We now integrate all variable except for p and get (6.43) is equal to

$$
\begin{aligned}
&\frac{2D\beta_c}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p\, |\eta(p)|^2 \\
&\qquad \times \left(\frac{\left(|p|^2-\mu\right)\rho'\left(\beta_c\left(|p|^2-\mu\right)\right)}{2\left(|p|^2-\mu\right)} + \frac{-\left(|p|^2-\mu\right)\rho'\left(-\beta_c\left(|p|^2-\mu\right)\right)}{-2\left(|p|^2-\mu\right)} \right) \psi_h(X) \\
&= \frac{2D\beta_c}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p\, |\eta(p)|^2 \rho'\left(\beta_c\left(|p|^2-\mu\right)\right) \psi_h(X) \\
&= \frac{2D\beta_c}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p\, |\eta(p)|^2 \left(\cosh^2\left(\frac{\beta_c}{2}\left(|p|^2-\mu\right)\right)\right)^{-1} \psi_h(X),
\end{aligned}
$$

which is one of the Ginzburg-Landau terms, precisely it is C_1, see Theorem 6.1. Next, we consider the second of the three next-to-leading-order terms. By exchanging s and r and replacing Y with $-Y$, we see that

$$
\begin{aligned}
&-\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_c z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s\, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta_Y^{s-r}\right) G^{-z}\left(\zeta_Y^{r-s}\right) \\
&\qquad \times \frac{i}{4}(r-s)\cdot J_A(X)(r+s)\psi_h(X) = 0.
\end{aligned}
$$

Finally, for the third and last of the next-to-leading-order terms, we have

$$
\begin{aligned}
&-\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_c z) \int_{\mathbb{R}^6} \mathrm{d}Y \mathrm{d}r \mathrm{d}s\, (V\alpha_*)(r)(V\alpha_*)(s) G^z\left(\zeta_Y^{s-r}\right) G^{-z}\left(\zeta_Y^{r-s}\right) \\
&\qquad \times \frac{1}{2}\left(Y\cdot\left(-i\nabla_X + 2a(X)\right)\right)^2 \psi_h(X)
\end{aligned} \tag{6.43}
$$

and we want to show that this equals

$$
\sum_{j,k=1}^{2} \left(-i\delta_{X_j} + 2a_j(X)\right) \mathbb{B}_{jk} \left(-i\delta_{X_k} + 2a_k(X)\right) \psi_h(X).
$$

In order to see this, note that (6.43) equals

$$
\begin{aligned}
&\frac{1}{(2\pi)^4} \sum_{j,k=1}^{2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_c z) \int_{\mathbb{R}^{10}} \mathrm{d}Y \mathrm{d}s \mathrm{d}u \mathrm{d}p \mathrm{d}q\, (V\alpha_*)(s)(V\alpha_*)(u) \\
&\quad \times \frac{e^{ip\cdot\left(\zeta_Y^{u-s}\right)} e^{iq\cdot\left(\zeta_Y^{s-u}\right)} Y_j Y_k}{\left(z-\left(|p|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)} \left(-i\delta_{X_j} + 2a_j(X)\right)\left(-i\delta_{X_k} + 2a_k(X)\right)\psi_h(X).
\end{aligned}
$$

We now integrate over u and s to see that

$$\frac{1}{(2\pi)^4}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^{10}}\mathrm{d}Y\mathrm{d}s\mathrm{d}u\mathrm{d}p\mathrm{d}q\;\left(V\alpha_*\right)(s)\left(V\alpha_*\right)(u)$$
$$\times\frac{e^{ip\cdot\left(\zeta_Y^{u-s}\right)}e^{iq\cdot\left(\zeta_Y^{s-u}\right)}Y_jY_k}{\left(z-\left(|p|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)}\psi_h(X)$$
$$=\frac{1}{2(2\pi)^4}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^6}\mathrm{d}Y\mathrm{d}p\mathrm{d}q\;\left|\left(\hat{V}*\hat{\alpha}_*\right)\left(\frac{p-q}{2}\right)\right|^2$$
$$\times\frac{e^{i(p+q)\cdot Y}Y_jY_k}{\left(z-\left(|p|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)}\psi_h(X),$$

which is equal to

$$\frac{1}{2(2\pi)^4}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^6}\mathrm{d}Y\mathrm{d}p\mathrm{d}q\;\left|\left(\hat{V}*\hat{\alpha}_*\right)(q)\right|^2$$
$$\times\frac{(-i)^2\partial^2_{p_jp_k}e^{ip\cdot Y}}{\left(z-\left(|q+p/2|^2-\mu\right)\right)\left(z+\left(|q-p/2|^2-\mu\right)\right)}\psi_h(X)$$

We calculate the derivatives for p_j and p_k, integrate over Y and p and finally see that (6.43) equals

$$-\frac{1}{2(2\pi)^2}\int_{\mathcal{C}}\frac{\mathrm{d}z}{2\pi i}\,\rho\left(\beta_c z\right)\int_{\mathbb{R}^2}\mathrm{d}q\;\left|\left(\hat{V}*\hat{\alpha}_*\right)(q)\right|^2\psi_h(X)$$
$$\times\Bigg(\frac{\delta_{jk}/2}{\left(z-\left(|q|^2-\mu\right)\right)^2\left(z+\left(|q|^2-\mu\right)\right)}-\frac{\delta_{jk}/2}{\left(z-\left(|q|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)^2}$$
$$+\frac{2q_kq_j}{\left(z-\left(|q|^2-\mu\right)\right)^3\left(z+\left(|q|^2-\mu\right)\right)}+\frac{2q_kq_j}{\left(z-\left(|q|^2-\mu\right)\right)^2\left(z+\left(|q|^2-\mu\right)\right)^2}$$
$$+\frac{2q_kq_j}{\left(z-\left(|q|^2-\mu\right)\right)\left(z+\left(|q|^2-\mu\right)\right)^3}\Bigg).$$

Next we integrate over z. The calculations for the complex integrals are presented in the appendix and show that (6.43) is equal to

$$-\frac{1}{2(2\pi)^2}\sum_{j,k=1}^{2}\int_{\mathbb{R}^2}\mathrm{d}q\;\left|\left(\hat{V}*\hat{\alpha}_*\right)(q)\right|^2\left(-i\delta_{X_j}+2a_j(X)\right)\left(-i\delta_{X_k}+2a_k(X)\right)\psi_h(X)$$
$$\times\left(\frac{\delta_{jk}}{2}\left(\frac{2\beta_c\rho'\left(\beta_c\left(|q|^2-\mu\right)\right)}{2\left(|q|^2-\mu\right)}-\frac{2\rho\left(\beta_c\left(|q|^2-\mu\right)\right)}{2\left(|q|^2-\mu\right)^2}\right)+2q_kq_j\frac{2\beta_c^2\rho''\left(\beta_c\left(|q|^2-\mu\right)\right)}{4\left(|q|^2-\mu\right)}\right)$$
$$=(-i\nabla+2a)\cdot\mathbb{B}(-i\nabla+2a)\psi_h(X),$$

which is again a Ginzburg-Landau term as stated in Theorem 6.1. That proves the estimate. □

Proof of (6.35). Recall the decomposition of $N_{T,A}$ into $N'_{T,A}$ and $\widetilde{N}_{T,A}$ at the beginning of Section 6.6. We first want to show that $N'_{T_h,A_h}\left(-2V^{1/2}\varphi_*\Psi_h\right)$ is small.

To do this we simply calculate that

$$
\begin{aligned}
&\left\|N'_{T_h,A_h}\left(-2V^{1/2}\varphi_*\Psi_h\right)\right\|_{\mathcal{L}^2_h} \\
&\lesssim \sum_{n \text{ odd}} \left\|(z_n - k_{A_h})^{-1}\right\|_\infty^2 \left\|\left(z_n + \overline{k}_{A_h}\right)^{-1}\right\|_\infty^2 \left\|\left(z_n - H_{A_h}\left(-2V^{1/2}\varphi_*\Psi_h\right)\right)^{-1}_{12}\right\|_\infty \\
&\qquad\qquad \times \left\|V^{1/2}\varphi_*\Psi_h\right\|_\infty \left\|V^{1/2}\varphi_*\Psi_h\right\|^3_{\mathcal{L}^6} \\
&\lesssim \sum_{n \text{ odd}} |n|^{-5} \cdot \|\Psi_h\|_{L^\infty}\left(\|\Psi_h\|_{L^2(Q_h)} + h^{-1}\|\nabla\Psi_h\|_{L^2(Q_h)}\right)^3 \\
&\lesssim h^4\|\psi\|_{L^\infty}\left(\|\psi\|_{L^2(Q)} + \|\nabla\psi\|_{L^2(Q)}\right)^3.
\end{aligned}
$$

This means

$$
\frac{1}{2}\left\|h^{-2}\mathcal{Q}_h P V^{1/2} N'_{T_h,A_h}\left(-2V^{1/2}\varphi_*\Psi_h\right)\right\|_{\mathcal{L}^2_h} \lesssim h\|\psi\|_{L^\infty}\left(\|\psi\|_{L^2(Q)} + \|\nabla\psi\|_{L^2(Q)}\right)^3.
$$

We now turn to the estimate for $\widetilde{N}_{T,A}$. We first calculate that

$$
\begin{aligned}
&\frac{1}{2}V^{1/2}\widetilde{N}_{T_h,A_h}\left(-2V^{1/2}\varphi_*\Psi_h\right)\left(\zeta^r_X, \zeta^{-r}_X\right) \\
&= -4h^3\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta_h z)\int \mathrm{d}\mathbf{w}_{123456}\, V^{1/2}(r) G^z_A\left(\zeta^r_X, w_1\right)(V\alpha_*)(w_1 - w_2) \\
&\qquad \times \psi_h\left(h(w_1+w_2)/2\right) G^{-z}_{-A}(w_2,w_3)(V\alpha_*)(w_3-w_4)\overline{\psi_h\left(h(w_3+w_4)/2\right)} G^z_A(w_4,w_5) \\
&\qquad \times (V\alpha_*)(w_5-w_6)\psi_h\left(h(w_5+w_6)/2\right) G^{-z}_{-A}\left(w_6, \zeta^{-r}_X\right),
\end{aligned}
$$

which by substitution can be seen to be exactly

$$
\begin{aligned}
&-4h^3\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta_h z)\int \mathrm{d}\mathbf{Y}_{123}\,\mathrm{d}\mathbf{s}_{123}\, V^{1/2}(r)(V\alpha_*)(s_1)(V\alpha_*)(s_2)(V\alpha_*)(s_3) \\
&\qquad\qquad \times G^z_A\left(\zeta^r_X, \zeta^{s_1}_{Y_1}\right) G^{-z}_{-A}\left(\zeta^{-s_1}_{Y_1}, \zeta^{s_2}_{Y_2}\right) G^z_A\left(\zeta^{-s_2}_{Y_2}, \zeta^{s_3}_{Y_3}\right) G^{-z}_{-A}\left(\zeta^{-s_3}_{Y_3}, \zeta^{-r}_X\right) \\
&\qquad\qquad\qquad\qquad \times \psi_h(hY_1)\overline{\psi_h(hY_2)}\psi_h(hY_3).
\end{aligned}
$$

This means that

$$
\begin{aligned}
&\frac{h^{-2}}{2}\mathcal{Q}_h P V^{1/2}\widetilde{N}_{T_h,A_h}\left(-2V^{1/2}P_\varepsilon\phi_h\right)(X) \\
&= -4\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta_h z)\int \mathrm{d}\mathbf{Y}_{123}\,\mathrm{d}\mathbf{s}_{0123}\,(V\alpha_*)(s_0)(V\alpha_*)(s_1)(V\alpha_*)(s_2)(V\alpha_*)(s_3) \\
&\qquad \times G^z_A\left(\zeta^{s_0}_{X/h}, \zeta^{s_1}_{Y_1}\right) G^{-z}_{-A}\left(\zeta^{-s_1}_{Y_1}, \zeta^{s_2}_{Y_2}\right) G^z_A\left(\zeta^{-s_2}_{Y_2}, \zeta^{s_3}_{Y_3}\right) G^{-z}_{-A}\left(\zeta^{-s_3}_{Y_3}, \zeta^{-s_0}_{X/h}\right) \\
&\qquad\qquad \times \psi_h(hY_1)\overline{\psi_h(hY_2)}\psi_h(hY_3),
\end{aligned}
$$

which is equal to

$$
\begin{aligned}
&-4\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta_h z)\int \mathrm{d}\mathbf{Y}_{123}\,\mathrm{d}\mathbf{s}_{0123}\,(V\alpha_*)(s_0)(V\alpha_*)(s_1)(V\alpha_*)(s_2)(V\alpha_*)(s_3) \\
&\quad \times G^z_A\left(\zeta^{s_0}_{X/h}, \zeta^{s_1}_{X/h+Y_1}\right) G^{-z}_{-A}\left(\zeta^{-s_1}_{X/h+Y_1}, \zeta^{s_2}_{X/h+Y_2}\right) G^z_A\left(\zeta^{-s_2}_{X/h+Y_2}, \zeta^{s_3}_{X/h+Y_3}\right) \\
&\quad \times G^{-z}_{-A}(\left(\zeta^{-s_3}_{X/h+Y_3}, \zeta^{-s_0}_{X/h}\right)\psi_h(X+hY_1)\overline{\psi_h(X+hY_2)}\psi_h(X+hY_3).
\end{aligned}
\tag{6.44}
$$

The first step is to show that we can replace G_A^z with K_A^z and the error is small. To do this we define

$$
\begin{aligned}
&R\left(X,Y_1,Y_2,Y_3,s_0,s_1,s_2,s_3\right)\\
&:= \frac{2}{\beta_h}\sum_{n \text{ odd}} \Big[G_A^{z_n}\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_1}\right) G_{-A}^{-z_n}\left(\zeta_{X/h+Y_1}^{-s_1},\zeta_{X/h+Y_2}^{s_2}\right)\\
&\qquad\qquad\qquad \times G_A^{z_n}\left(\zeta_{X/h+Y_2}^{-s_2},\zeta_{X/h+Y_3}^{s_3}\right) G_{-A}^{-z_n}\left(\zeta_{X/h+Y_3}^{-s_3},\zeta_{X/h}^{-s_0}\right)\\
&\quad - K_A^{z_n}\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_1}\right) K_{-A}^{-z_n}\left(\zeta_{X/h+Y_1}^{-s_1},\zeta_{X/h+Y_2}^{s_2}\right)\\
&\qquad\qquad\qquad \times K_A^{z_n}\left(\zeta_{X/h+Y_2}^{-s_2},\zeta_{X/h+Y_3}^{s_3}\right) K_{-A}^{-z_n}(\left(\zeta_{X/h+Y_3}^{-s_3},\zeta_{X/h}^{-s_0}\right)\Big].
\end{aligned}
$$

We want to estimate

$$
\int \mathrm{d}\mathbf{Y}_{123} \operatorname*{ess\,sup}_{X/h\in Q_h} \left|R\left(X,Y_1,Y_2,Y_3,s_0,s_1,s_2,s_3\right)\right|.
$$

This involves estimating terms like

$$
\begin{aligned}
&\Big|\left(G_A^{z_n}\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_1}\right) - K_A^{z_n}\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_1}\right)\right)\\
&\qquad \times K_{-A}^{-z_n}\left(\zeta_{X/h+Y_1}^{-s_1},\zeta_{X/h+Y_2}^{s_2}\right) K_A^{z_n}\left(\zeta_{X/h+Y_2}^{-s_2},\zeta_{X/h+Y_3}^{s_3}\right) K_{-A}^{-z_n}\left(\zeta_{X/h+Y_3}^{-s_3},\zeta_{X/h}^{-s_0}\right)\Big|\\
&\le H_A^{z_n}\left(\zeta_{Y_1}^{s_1-s_0}\right)\left|G^{-z_n}\left(\zeta_{Y_1-Y_2}^{-(s_1+s_2)}\right) G^{z_n}\left(\zeta_{Y_2-Y_3}^{-(s_2+s_3)}\right) G^{-z_n}\left(\zeta_{Y_3}^{s_0-s_3}\right)\right|.
\end{aligned}
$$

Since all the terms are very similar to estimate, we restrict the presentation here to the example given above. Altogether, one can show that

$$
\int \mathrm{d}\mathbf{Y}_{123} \operatorname*{ess\,sup}_{X/h\in Q_h} \left|R\left(X,Y_1,Y_2,Y_3,s_0,s_1,s_2,s_3\right)\right| \lesssim M_A.
$$

This then allows us to show that the error is

$$
\left(\int_Q \mathrm{d}X\ |(6.44)|^2\right)^{1/2} \lesssim M_A \left(\int_Q \mathrm{d}X\,|\psi_h(X)|^6\right)^{3/6} \lesssim M_A \left\|\psi_h\right\|_{H^1(Q)} \lesssim h^3.
$$

We now go back to (6.44) with G_A^z replaced by K_A^z. That means we are now dealing with

$$
\begin{aligned}
&-4\int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i}\,\rho(\beta_h z)\int \mathrm{d}\mathbf{Y}_{123}\,\mathrm{d}\mathbf{s}_{0123}\ \left(V\alpha_*\right)(s_0)\left(V\alpha_*\right)(s_1)\left(V\alpha_*\right)(s_2)\left(V\alpha_*\right)(s_3)\\
&\quad\times G^z\left(\zeta_{Y_1}^{s_1-s_0}\right) G^{-z}\left(\zeta_{Y_1-Y_2}^{-(s_1+s_2)}\right) G^z\left(\zeta_{Y_2-Y_3}^{-(s_2+s_3)}\right) G^{-z}\left(\zeta_{Y_3}^{s_0-s_3}\right)\\
&\quad\times e^{i\Phi_A\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_0}\right)} e^{-i\Phi_A\left(\zeta_{X/h+Y_1}^{-s_0},\zeta_{X/h+Y_2}^{s_2}\right)} e^{i\Phi_A\left(\zeta_{X/h+Y_2}^{-s_2},\zeta_{X/h+Y_3}^{s_3}\right)} e^{-i\Phi_A\left(\zeta_{X/h+Y_3}^{-s_3},\zeta_{X/h}^{-s_0}\right)}\\
&\quad\times \psi_h(X+hY_1)\overline{\psi_h\left(X+hY_2\right)}\psi_h\left(X+hY_3\right),
\end{aligned}
\tag{6.45}
$$

where we already inserted the definition of K_z^A given in (6.13). Now, let us state the fact that, for example,

$$
\left|\exp\left(i\Phi_A\left(\zeta_{X/h}^{s_0},\zeta_{X/h+Y_1}^{s_0}\right)\right)-1\right| \le \left\|DA\right\|_{L^\infty}\left|\zeta_{Y_1}^{s_1-s_0}\right|
$$

and $\|DA\|_{L^\infty} \lesssim h^2$. Thus, we see that we obtain an error of order h^2 if we replace all the exponential factors in (6.45) by 1 and only consider

$$
\begin{aligned}
& -4 \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_h z) \int \mathrm{d}\mathbf{Y}_{123}\, \mathrm{d}\mathbf{s}_{0123}\ (V\alpha_*)(s_0)\,(V\alpha_*)(s_1)\,(V\alpha_*)(s_2)\,(V\alpha_*)(s_3) \\
& \quad \times G^z\left(\zeta_{Y_1}^{s_1-s_0}\right) G^{-z}\left(\zeta_{Y_1-Y_2}^{-(s_1+s_2)}\right) G^z\left(\zeta_{Y_2-Y_3}^{-(s_2+s_3)}\right) G^{-z}\left(\zeta_{Y_3}^{s_0-s_3}\right) \\
& \quad \times e^{ihY_1\cdot(-i\nabla_X)}\psi_h(X)\overline{e^{ihY_2\cdot(-i\nabla_X)}\psi_h(X)}e^{ihY_3\cdot(-i\nabla_X)}\psi_h(X).
\end{aligned}
$$

Next, let us mention that one can show that

$$
\left\|\left(1-e^{ihY_1\cdot(-i\nabla_X)}\right)\psi_h(X)\right\|_{H^1(Q)} \lesssim h\,|Y_1|\,\|\psi_h\|_{H^2(Q)} \lesssim \varepsilon\,|Y_1|\,.
$$

This means we obtain an error of ε if we consider only

$$
\begin{aligned}
& -4 \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_h z) \int \mathrm{d}\mathbf{Y}_{123}\, \mathrm{d}\mathbf{s}_{0123}\ (V\alpha_*)(s_0)\,(V\alpha_*)(s_1)\,(V\alpha_*)(s_2)\,(V\alpha_*)(s_3) \\
& \quad \times G^z\left(\zeta_{Y_1}^{s_1-s_0}\right) G^{-z}\left(\zeta_{Y_1-Y_2}^{-(s_1+s_2)}\right) G^z\left(\zeta_{Y_2-Y_3}^{-(s_2+s_3)}\right) G^{-z}\left(\zeta_{Y_3}^{s_0-s_3}\right) |\psi_h(X)|^2\,\psi_h(X) \\
& = \frac{-4}{(2\pi)^8} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_h z) \int \mathrm{d}\mathbf{p}_{1234}\, \mathrm{d}\mathbf{Y}_{123}\, \mathrm{d}\mathbf{s}_{0123}\ (V\alpha_*)(s_0)\,(V\alpha_*)(s_1)\,(V\alpha_*)(s_2) \\
& \times (V\alpha_*)(s_3) \frac{e^{ip_1\cdot\left(\zeta_{Y_1}^{s_1-s_0}\right)}}{z-\left(|p_1|^2-\mu\right)} \frac{e^{ip_2\cdot\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)}}{z+\left(|p_2|^2-\mu\right)} \frac{e^{ip_3\cdot\left(\zeta_{Y_2-Y_3}^{-s_2-s_3}\right)}}{z-\left(|p_3|^2-\mu\right)} \frac{e^{ip_4\cdot\left(\zeta_{Y_3}^{s_0-s_3}\right)}}{z+\left(|p_4|^2-\mu\right)} \\
& \quad \times |\psi_h(X)|^2\,\psi_h(X).
\end{aligned}
\tag{6.46}
$$

Now we integrate over p_2, p_3 and then p_4 and get that

$$
\begin{aligned}
& \frac{4}{(2\pi)^2} \int_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi i} \rho(\beta_h z) \int_{\mathbb{R}^2} \mathrm{d}p_1 \mathrm{d}\mathbf{s}_{0123}\ (V\alpha_*)(s_0)\,(V\alpha_*)(s_1)\,(V\alpha_*)(s_2)\,(V\alpha_*)(s_3) \\
& \times \frac{e^{ip_1\cdot(-(s_0-s_1)/2)}}{z-\left(|p_1|^2-\mu\right)} \frac{e^{-ip_1\cdot(-(s_1+s_2)/2)}}{z+\left(|p_1|^2-\mu\right)} \frac{e^{-ip_1\cdot(-(s_2+s_3)/2)}}{z-\left(|p_1|^2-\mu\right)} \frac{e^{-ip_1\cdot(-(s_3-s_0)/2)}}{z+\left(|p_1|^2-\mu\right)} |\psi_h(X)|^2\,\psi_h(X) \\
& = -\frac{4}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p\, |\eta(p)|^4 \left(\frac{\beta_h \rho'\left(\beta_h\left(|p|^2-\mu\right)\right)}{2\left(|p|^2-\mu\right)^2} - \frac{\rho\left(\beta_h\left(|p|^2-\mu\right)\right)}{2\left(|p|^2-\mu\right)^3} \right) |\psi_h(X)|^2\,\psi_h(X)
\end{aligned}
$$

is an equivalent expression for (6.46). Since $\beta_h = \upbeta_c + O(h^2)$, it is straightforward to verify that we obtain an error of h^2 if we replace β_h with $\upbeta_c$. Hence, we are left with

$$
\begin{aligned}
& -4(2\pi)^{-2} \int_{\mathbb{R}^2} \mathrm{d}p\, |\eta(p)|^4 \\
& \times \left(\frac{\upbeta_c}{4\left(|p|^2-\mu\right)^2 \cosh^2\left(\upbeta_c\left(|p|^2-\mu\right)/2\right)} - \frac{\tanh\left(\upbeta_c\left(|p|^2-\mu\right)/2\right)}{2\left(|p|^2-\mu\right)^3} \right) |\psi_h(X)|^2\,\psi_h(X) \\
& = C_2\, |\psi_h(X)|^2\,\psi_h(X).
\end{aligned}
$$

That proves the estimate. □

Proof of (6.36). By the definition of a_h, we have $A_h(x) = h a_h(hx)$, which implies $\operatorname{curl}^2 A_h(x) = h^3 \operatorname{curl}^2 a_h(hx)$, and therefore $\vec{\mathcal{Q}}_h \operatorname{curl}^2 A_h(x) = h^2 \operatorname{curl}^2 a_h(x)$. The estimate now follows trivially. □

Proof of (6.37). We will first show that

$$\left\| h^{-2}\vec{\mathcal{Q}}_h J'_{T,A_h}\left(-2V^{1/2}P_\varepsilon\phi\right)\right\|_{L^2(Q)} \lesssim \varepsilon^{1/2}h^{1/2}$$

and therefore is negligiable. To do this we simply apply Lemma 6.4 and calculate that

$$\left\| h^{-2}\vec{\mathcal{Q}}_h J'_{T,A_h}\left(-2V^{1/2}P_\varepsilon\phi\right)\right\|_{L^2(Q)} \lesssim h^{-3}\sum_{n\text{ odd}}\left\|K_n^{(5)}\left(1-\Delta\right)\right\|_{\mathcal{L}_h^2},$$

where

$$\begin{aligned}K_n^{(5)} = {}&(\overline{z}_n - H_A\left(-2V^{1/2}P_\varepsilon\phi\right)^*)_{11}^{-1}V^{1/2}P_\varepsilon\phi\left(\overline{z}_n+\overline{k}_A\right)^{-1}\overline{V^{1/2}P_\varepsilon\phi}(\overline{z}_n-k_A)^{-1}\\ &\times V^{1/2}P_\varepsilon\phi\left(\overline{z}_n+\overline{k}_A\right)^{-1}\overline{V^{1/2}P_\varepsilon\phi}(\overline{z}_n-k_A)^{-1}\pi_A.\end{aligned}$$

Thus, we get the following estimate

$$\begin{aligned}\left\| h^{-2}\vec{\mathcal{Q}}_h J'_{T,A_h}\left(-2V^{1/2}P_\varepsilon\phi\right)\right\|_{L^2(Q)} &\lesssim h^{-3}\left(\sum_{n\text{ odd}}|n|^{-10}\left\|V^{1/2}P_\varepsilon\phi\right\|_\infty^2\left\|V^{1/2}P_\varepsilon\phi\right\|_{\mathcal{L}_h^6}^6\right)^{1/2}\\ &\lesssim \varepsilon^{1/2}h^{1/2}.\end{aligned}$$

We now turn to the actual derivation of the coefficients that appear in the Ginzburg-Landau equation. Our first goal is to show that

$$\left\| h^{-2}\vec{\mathcal{Q}}_h\mathcal{J}_{T,A_h} + \frac{1}{24\pi\left(1+e^{-\beta_c\mu}\right)}\operatorname{curl}^2 a(x)\right\|_{L^2(Q)} \lesssim h^2. \tag{6.47}$$

We begin by writing

$$\mathcal{J}_{T,A_h} = \operatorname{Re}\pi_A\left(\frac{1}{1+e^{\beta k_A}}\right)\bigg|_{y=x} = \operatorname{Re}\left(-i\nabla + A\right)\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\left(z-k_A\right)^{-1}\bigg|_{y=x},$$

where $\widetilde{\mathcal{C}}$ is $\{r \pm i\pi/(2\beta_c)\,|\,r\in[-1,\infty)\}\cup\{-1+is\,|\,s\in[-i\pi//(2\beta_c), i\pi/(2\beta_c)]\}$.

We now wish to replace $(z-k_A)^{-1}$ with $K_A^z - K_A^z T_A^z$. We recall that $(z-k_A)^{-1} = K_A^z(1+T_A^z)^{-1}$ and therefore we let

$$R_A^z = \left(-i\nabla + A\right)\left((z-k_A)^{-1} - K_A^z + K_A^z T_A^z\right) = S_A^z\widetilde{R}_z^A,$$

where

$$S_A^z(x,y) = \left(-i\nabla + A\right)K_A^z(x,y) = e^{i\Phi_A(x,y)}\left(-i\nabla G^z(x,y) + \widetilde{A}(x,y)G^z(x,y)\right)$$

satisfies $|(1-\Delta)\,S_A^z\|_{\mathcal{L}_h^2} \lesssim 1$, and $\left\|\widetilde{R}_A^z\right\|_\infty \lesssim \|T_A^z\|_\infty^2 \lesssim M_A^2$. By applying Lemma 6.4, we then get the estimate

$$\begin{aligned}\left\|\operatorname{Re}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\sigma(\beta z)R_A^z\bigg|_{x=y}\right\|_{L^2(Q_h)} &\leq \int_{\widetilde{\mathcal{C}}}|\mathrm{d}z|\,|\sigma(\beta z)|\,\left\|R_A^z|_{x=y}\right\|_{L^2(Q_h)}\\ &\lesssim \|(1-\Delta)\,R_A^z\|_{\mathcal{L}_h^2} \leq \|(1-\Delta)\,S_A^z\|_{\mathcal{L}_h^2}\,\|T_A^z\|_\infty^2 \lesssim M_A^2 \lesssim h^6.\end{aligned}$$

We proceed term by term. For the first term, notice that

$$\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\sigma(\beta z)K_A^z(x,y) = \int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\sigma(\beta z)e^{i\Phi_A(x,y)}G^z(x,y).$$

It is easy to see that

$$\int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z)\nabla G^z(0) = 0$$

using Fourier transform. Therefore,

$$\begin{aligned}
&\mathrm{Re}\,(-i\nabla + A) \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i} \sigma(\beta z) K_A^z(x,y) \Big|_{y=x} \\
&= \mathrm{Re} \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z) e^{i\Phi_A(x,y)} \left(-i\nabla G^z\,(x-y) + \widetilde{A}(x,y) G^z\,(x-y) \right) \Big|_{y=x} = 0
\end{aligned}$$

and the first term does not contribute. For the second term we start by noting that

$$\begin{aligned}
&\int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z) (K_A^z T_A^z)(x,y) \\
&\qquad = \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i} \sigma(\beta z)\, G^z(x-w) G^z(w-y) e^{i\Phi_A(x,w)} e^{i\Phi_A(w,y)} \\
&\qquad\qquad \times \left(i\mathrm{div}\widetilde{A}(w,y) - |\widetilde{A}(w,y)|^2 \right).
\end{aligned}$$

This means that we have to consider two terms, namely,

$$\begin{aligned}
&\mathrm{Re}\,(-i\nabla + A) \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z) K_A^z(x,y) T_A^z(x,y) \Big|_{y=x} \\
&= \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z)(-i)\nabla G^z(x-w) G^z(w-x) \left(i\mathrm{div}\widetilde{A}(w,x) - |\widetilde{A}(w,x)|^2 \right) \\
&+ \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z) G^z(x-w) G^z(w-x) \widetilde{A}(x,w) \left(i\mathrm{div}\widetilde{A}(w,x) - |\widetilde{A}(w,x)|^2 \right).
\end{aligned}$$

Remember, that we want to estimate the L_2-norm. We start with the second term. Taking into account $\vec{\mathcal{Q}}_h$ we see that we need to estimate

$$\begin{aligned}
&h^{-3} \Bigg(\int_Q \mathrm{d}x \,\Bigg| \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i} \sigma(\beta z) G^z\,(x/h - w)\, G^z\,(w - x/h)\, \widetilde{A}\,(x/h, w) \\
&\qquad\qquad \times \left(i\mathrm{div}\widetilde{A}\,(w, x/h) - \left| \widetilde{A}\,(w, x/h) \right|^2 \right) \Bigg|^2 \Bigg)^{1/2} \\
&\lesssim h^{-3} \left\| \mathrm{curl} A \right\|_{L^\infty(\mathbb{R}^2)} \left(\left\| \mathrm{curl}^2 A \right\|_{L^\infty(\mathbb{R}^2)} + \left\| \mathrm{curl} A \right\|_{L^\infty(\mathbb{R}^2)}^2 \right) \\
&\lesssim h^2.
\end{aligned}$$

In order to estimate the first term, we split it into two parts. So, we first look at

$$\begin{aligned}
&\int_Q \mathrm{d}x \left| \mathrm{Re} \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z)(-i)\nabla G^z\,(x/h - w)\, G^z\,(w - x/h) \left| \widetilde{A}\,(w, x/h) \right|^2 \right| \\
&= \int_Q \mathrm{d}x \left| \mathrm{Re} \int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\beta z)(-i)\nabla G^z(w) G^z(w) |\widetilde{A}\,(w + x/h, x/h)\,|^2 \right| \\
&\lesssim \left\| \mathrm{curl} A \right\|_{L^\infty(\mathbb{R}^2)}^2.
\end{aligned}$$

There is now one term left to consider and that is

$$
\begin{aligned}
&\operatorname{Re}\int_{\mathbb{R}^2}\mathrm{d}w\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)(-i)\nabla G^z\left(x/h-w\right)G^z\left(w-x/h\right)i\mathrm{div}\widetilde{A}\left(w,x/h\right)\\
&=\frac{1}{2}\mathrm{Re}\int_{\mathbb{R}^2}\mathrm{d}w\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\nabla_w\left(G^z(w)\right)^2\mathrm{div}\widetilde{A}\left(w+x/h,x/h\right)\\
&=\frac{1}{2}\,\mathrm{Re}\int_{\mathbb{R}^2}\mathrm{d}w\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\left(G^z(w)\right)^2\nabla_w\mathrm{div}\widetilde{A}\left(w+x/h,x/h\right).
\end{aligned}
$$

Let us first calculate

$$
\begin{aligned}
&\nabla_w\mathrm{div}\widetilde{A}\left(w+x/h,x/h\right)\\
&=\left(\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2A\left(x/h+tw\right)\cdot\nabla_w\right)w+\left(w\cdot\nabla_w\right)\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2A\left(x/h+tw\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad+w\wedge\int_0^1\mathrm{d}t\,t^3\mathrm{curl}^3A\left(x/h+tw\right)\\
&=\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2A\left(x/h+tw\right)+w\cdot\nabla_w\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2A\left(x/h+tw\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad+w\wedge\int_0^1\mathrm{d}t\,t^3\mathrm{curl}^3A\left(x/h+tw\right)\\
&=h^3\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2a\left(x+thw\right)+h^3w\cdot\nabla_w\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2a\left(x+thw\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad+w\wedge\int_0^1\mathrm{d}t\,t^3\mathrm{curl}^3A\left(x/h+tw\right)
\end{aligned}
$$

The latter two terms can be estimated as before, whereas for the first we first write

$$
\begin{aligned}
&\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2a\left(x+thw\right)\\
&\qquad\qquad=\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2a(x)+\int_0^1\mathrm{d}t\,t^2\int_0^1\mathrm{d}s\,thw\cdot\nabla\mathrm{curl}^2a\left(x+sthw\right),
\end{aligned}
$$

but the second term can also be estimated as before. We are now left with the contributing term and we have all together, after writing $\sigma(\beta z)=\sigma(\beta_c z)+O(h^2)$, where the remainder decays exponentially in z,

$$
\begin{aligned}
&\mathrm{Re}\,\frac{1}{2}\int_{\mathbb{R}^2}\mathrm{d}w\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta_c z)(G^z(w))^2\int_0^1\mathrm{d}t\,t^2\mathrm{curl}^2a(x)\\
&=\frac{1}{6}\,\mathrm{Re}\int_{\mathbb{R}^2}\mathrm{d}w\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta_c z)(G^z(w))^2\mathrm{curl}^2a(x)\\
&=\frac{\beta_c}{6(2\pi)^2}\int_{\mathbb{R}^2}\mathrm{d}p\,\sigma'(\beta_c(|p|^2-\mu))\mathrm{curl}^2a(x),
\end{aligned}
$$

where we used the Fourier transform to calculate that

$$\begin{aligned}
&\int_{\mathbb{R}^2} \mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\upbeta_c z)(G^z(w))^2 \\
&= \frac{1}{(2\pi)^4} \int_{\mathbb{R}^6} \mathrm{d}p\mathrm{d}q\mathrm{d}w \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\upbeta_c z) \frac{e^{i(p+q)\cdot w}}{(z-(|p|^2-\mu))\,(z-(|q|^2-\mu))} \\
&= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p \int_{\widetilde{\mathcal{C}}} \frac{\mathrm{d}z}{2\pi i}\, \sigma(\upbeta_c z) \frac{1}{(z-(|p|^2-\mu))^2} \\
&= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}p\, \upbeta_c \sigma'\left(\upbeta_c\left(|p|^2-\mu\right)\right).
\end{aligned}$$

We can, however, explicitly calculate that

$$\begin{aligned}
\int_{\mathbb{R}^2} \mathrm{d}p\, \sigma'\left(\upbeta_c\left(|p|^2-\mu\right)\right) &= -\int_{\mathbb{R}^2} \mathrm{d}p\, \frac{e^{\upbeta_c\left(|p|^2-\mu\right)}}{\left(1+e^{\upbeta_c(|p|^2-\mu)}\right)^2} \\
&= -2\pi \int_0^\infty dr\, \frac{re^{\upbeta_c\left(r^2-\mu\right)}}{\left(1+e^{\upbeta_c(r^2-\mu)}\right)^2} = \frac{2\pi}{2\upbeta_c} \frac{1}{1+e^{\upbeta_c(r^2-\mu)}}\bigg|_{r=0}^{\infty} = -\frac{2\pi}{2\upbeta_c} \frac{1}{1+e^{-\upbeta_c\mu}}.
\end{aligned}$$

That completes the proof of (6.47).

We now consider $\widetilde{J}_{T_h,A_h}(-2V^{1/2}P_\varepsilon\phi_h)$. Let us recall that $P_\varepsilon\phi_h(x,y) = h\varphi_*(x-y)\psi_h(h(x+y)/2)$. This means

$$\begin{aligned}
&(z-k_A)^{-1}\left(-2V^{1/2}P_\varepsilon\phi_h\right)\left(z+\overline{k}_A\right)^{-1}\overline{\left(-2V^{1/2}P_\varepsilon\phi_h\right)}\,(z-k_A)^{-1}(x,y) \\
&= -4h^2 \int \mathrm{d}\mathbf{w}_{1234}\, G_A^z(x,w_1)\,(V\alpha_*)(w_1-w_2)\,\psi_h\left(h(w_1+w_2)/2\right) G_{-A}^{-z}(w_2,w_3) \\
&\qquad\qquad \times (V\alpha_*)(w_3-w_4)\,\overline{\psi_h\left(h(w_3+w_4)/2\right)} G_A^z(w_4,y) \\
&= -4h^2 \int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\, G_A^z\left(x,\zeta_{Y_1}^{s_1}\right)(V\alpha_*)(s_1)\psi_h(hY_1) G_{-A}^{-z}\left(\zeta_{Y_1}^{-s_1},\zeta_{Y_2}^{s_2}\right) \\
&\qquad\qquad \times (V\alpha_*)(s_2)\overline{\psi_h(hY_2)} G_A^z\left(\zeta_{Y_2}^{-s_2},y\right).
\end{aligned}$$

As before, we can show that the error arising from replacing G_A^z with K_A^z is negligiable. Therefore we consider the term

$$\begin{aligned}
&-4h^2 \int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\;\; (V\alpha_*)(s_1)\,(V\alpha_*)(s_2) G^z\left(x-\zeta_{Y_1}^{-s_1}\right) G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right) G^z\left(\zeta_{Y_2}^{-s_2}-y\right) \\
&\qquad\qquad \times e^{i\Phi_A\left(x,\zeta_{Y_1}^{s_1}\right)} e^{-i\Phi_A\left(\zeta_{Y_1}^{-s_1},\zeta_{Y_2}^{s_2}\right)} e^{i\Phi_A\left(\zeta_{Y_2}^{-s_2},y\right)} \psi_h(hY_1)\overline{\psi_h(hY_2)}.
\end{aligned}$$

Applying $(-i\nabla + A)$ and taking the diagonal, we have

$$\begin{aligned}
&-4h^2\,\mathrm{Re} \int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\;\; (V\alpha_*)(s_1)\,(V\alpha_*)(s_2) \\
&\times \left(-i\nabla G^z\left(x-\zeta_{Y_1}^{-s_1}\right) + \widetilde{A}\left(x,\zeta_{Y_1}^{s_1}\right) G^z\left(x-\zeta_{Y_1}^{-s_1}\right)\right) G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right) G^z\left(\zeta_{Y_2}^{-s_2}-x\right) \\
&\qquad\qquad \times e^{i\Phi_A\left(x,\zeta_{Y_1}^{s_1}\right)} e^{-i\Phi_A\left(\zeta_{Y_1}^{-s_1},\zeta_{Y_2}^{s_2}\right)} e^{i\Phi_A\left(\zeta_{Y_2}^{-s_2},x\right)} \psi_h(hY_1)\overline{\psi_h(hY_2)}
\end{aligned}$$

And finally, applying $h^{-2}\vec{\mathcal{Q}}_h$ we obtain

$$
\begin{aligned}
-\frac{4}{h}\,\mathrm{Re}\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\ \ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2) \\
\times \Big(-i\nabla G^z\left(-\zeta_{Y_1}^{-s_1}\right)+\widetilde{A}\left(x/h, x/h+\zeta_{Y_1}^{s_1}\right)G^z\left(-\zeta_{Y_1}^{-s_1}\right)\Big)\,G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)G^z\left(\zeta_{Y_2}^{-s_2}\right) \\
\times e^{i\Phi_A\left(x/h,x/h+\zeta_{Y_1}^{s_1}\right)}e^{-i\Phi_A\left(x/h+\zeta_{Y_1}^{-s_1},x/h+\zeta_{Y_2}^{s_2}\right)}e^{i\Phi_A\left(x/h+\zeta_{Y_2}^{-s_2},x/h\right)} \\
\times \psi_h\left(x+hY_1\right)\overline{\psi_h\left(x+hY_2\right)}.
\end{aligned}
$$

As before we can show that the term involving $\widetilde{A}$ is of smaller order, and therefore we consider

$$
\begin{aligned}
&\frac{4}{h}\,\mathrm{Re}\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\ \ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2)\,(i\nabla)\,G^z\left(-\zeta_{Y_1}^{-s_1}\right)G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)G^z\left(\zeta_{Y_2}^{-s_2}\right) \\
&\quad\times e^{i\Phi_a\left(x,x+\zeta_{hY_1}^{hs_1}\right)}e^{-i\Phi_a\left(x+\zeta_{hY_1}^{-hs_1},x+\zeta_{hY_2}^{hs_2}\right)}e^{i\Phi_a(x+\zeta_{hY_2}^{-hs_2},x)}e^{ihY_1\cdot(-i\nabla_x)}\psi_h(x)\overline{e^{ihY_2\cdot(-i\nabla_x)}\psi_h(x)} \\
&=\frac{4}{h}\,\mathrm{Re}\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\ \ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2)\,(i\nabla)\,G^z\left(-\zeta_{Y_1}^{-s_1}\right)G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)G^z\left(\zeta_{Y_2}^{-s_2}\right) \\
&\quad\times e^{i\Phi_a\left(x,x+\zeta_{hY_1}^{hs_1}\right)}e^{-i\Phi_{2a}(x,x+hY_1)}e^{-i\Phi_a\left(x+\zeta_{hY_1}^{-hs_1},x+\zeta_{hY_2}^{hs_2}\right)} \\
&\quad\times e^{\Phi_a(x+\zeta_{hY_2}^{-hs_2},x)}e^{i\Phi_{2a}(x,x+hY_2)}e^{ihY_1\cdot(-i\nabla_x+2a(x))}\psi_h(x)\overline{e^{ihY_2\cdot(-i\nabla_x+2a(x))}\psi_h(x)}
\end{aligned}
$$

The leading term in h can be shown to be

$$
\begin{aligned}
&-4\,\mathrm{Re}\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\ \ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2)\,(\nabla G^z)\left(-\zeta_{Y_1}^{-s_1}\right)G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right) \\
&\quad\times G^z\left(\zeta_{Y_2}^{-s_2}\right)\Big(\overline{\psi_h(x)}Y_1\cdot(-i\nabla_x+2a(x))\psi_h(x)-\psi_h(x)\overline{Y_2\cdot(-i\nabla_x+2a(x))\psi_h(x)}\Big)
\end{aligned}
\tag{6.48}
$$

We now wish to calculate the Ginzburg-Landau coefficient. For the term involving Y_1, we see that

$$
\begin{aligned}
&\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\ \ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2) \\
&\qquad\times \partial_k G^z(-\zeta_{Y_1}^{-s_1})G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)G^z(\zeta_{Y_2}^{-s_2})(Y_1)_j \\
&=-\frac{1}{(2\pi)^6}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\,\mathrm{d}\mathbf{p}_{123}\ (V\alpha_*)\,(s_1)\,(V\alpha_*)\,(s_2) \\
&\qquad\times\frac{i(p_1)_k e^{ip_1\cdot(-\zeta_{Y_1}^{-s_1})}}{z-\left(|p_1|^2-\mu\right)}\frac{e^{ip_2\cdot(\zeta_{Y_1-Y_2}^{-(s_1+s_2)})}}{z+\left(|p_2|^2-\mu\right)}\frac{e^{ip_3\cdot(\zeta_{Y_2}^{-s_2})}}{z-(|p_3|^2-\mu)}(Y_1)_j,
\end{aligned}
$$

which is equal to

$$
\begin{aligned}
&-\frac{1}{(2\pi)^6}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int \mathrm{d}\mathbf{Y}_{12}\,\mathrm{d}\mathbf{p}_{123}\,\eta(-(p_1+p_2)/2)\eta(-(p_2+p_3)/2)\\
&\qquad\qquad\times\frac{i(p_1)_k(Y_2)_j e^{i(-p_1+p_2)\cdot Y_1}e^{i(-p_2+p_3)\cdot Y_2}}{(z-(|p_1|^2-\mu))(z+(|p_2|^2-\mu))(z-(|p_3|^2-\mu))}. \qquad (6.49)
\end{aligned}
$$

Integrating over p_3 and then p_2 we get

$$
\begin{aligned}
&-\frac{1}{(2\pi)^4}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^6}\mathrm{d}Y\,\mathrm{d}p\mathrm{d}q\,\eta(p)\eta(p-q/2)\\
&\quad\times\frac{i\,(p_k+q_k/2)\,i\partial_{q_j}e^{-iq\cdot Y}}{(z-(|p+q/2|^2-\mu))(z+(|p-q/2|^2-\mu))(z-(|p-q/2|^2-\mu))}\\
&=-\frac{1}{(2\pi)^4}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\\
&\quad\times\partial_{q_j}\left(\frac{\eta(p-q/2)\eta(p)\,(p_k+q_k/2)}{(z-(|p+q/2|^2-\mu))(z+(|p-q/2|^2-\mu))(z-(|p-q/2|^2-\mu))}\right)\Bigg|_{q=0}
\end{aligned}
$$

and finally see that (6.49) equals

$$
\begin{aligned}
&\frac{1}{(2\pi)^2}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\,\left(\frac{p_k\partial_j|\eta(p)|^2}{4\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))}\right.\\
&\qquad\left.-\frac{|\eta(p)|^2\delta_{kj}}{2\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))}-\frac{p_kp_j|\eta(p)|^2}{(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))^2}\right)\\
&=\frac{1}{(2\pi)^2}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\,|\eta(p)|^2\left(-\frac{3\delta_{kj}}{4\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))}\right.\\
&\qquad\left.-\frac{p_kp_j}{(z-(|p|^2-\mu))^3\,(z+(|p|^2-\mu))}-\frac{p_kp_j}{2\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))^2}\right).
\end{aligned}
$$

The analogous calculation for the term involving Y_2 shows that

$$
\begin{aligned}
&-\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int \mathrm{d}\mathbf{s}_{12}\,\mathrm{d}\mathbf{Y}_{12}\,\left(V\alpha_*\right)(s_1)\left(V\alpha_*\right)(s_2)\\
&\qquad\times\partial_k G^z(-\zeta_{Y_1}^{-s_1})G^{-z}\left(\zeta_{Y_1-Y_2}^{-s_1-s_2}\right)G^z(\zeta_{Y_2}^{-s_2})(Y_2)_j\\
&=\frac{1}{(2\pi)^2}\int_{\widetilde{\mathcal{C}}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\,|\eta(p)|^2\left(\frac{\delta_{kj}}{4\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))}\right.\\
&\qquad\left.-\frac{p_kp_j}{(z-(|p|^2-\mu))^3\,(z+(|p|^2-\mu))}-\frac{p_kp_j}{2\,(z-(|p|^2-\mu))^2\,(z+(|p|^2-\mu))^2}\right).
\end{aligned}
$$

This means that (6.48) is equal to

$$\frac{-4}{(2\pi)^2}\sum_{k,j=1}^{2}\int_{\widetilde{C}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\,|\eta(p)|^2\Bigg(-\frac{\delta_{kj}}{2\left(z-(|p|^2-\mu)\right)^2(z+(|p|^2-\mu))}$$
$$-\frac{2p_kp_j}{\left(z-(|p|^2-\mu)\right)^3(z+(|p|^2-\mu))}-\frac{p_kp_j}{\left(z-(|p|^2-\mu)\right)^2(z+(|p|^2-\mu))^2}\Bigg)$$
$$\times\,\overline{\psi_h(x)}\cdot(-i\nabla+2a(x))\,\psi_h(x)$$
$$=\frac{4}{(2\pi)^2}\sum_{k,j=1}^{2}\int_{\widetilde{C}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta z)\int_{\mathbb{R}^2}\mathrm{d}p\,|\eta(p)|^2\left[\frac{1}{2}\left(\frac{\rho(\beta(|p|^2-\mu))}{4(|p|^2-\mu)^2}-\frac{\beta\rho'(\beta(|p|^2-\mu))}{4(|p|^2-\mu)}\right)\delta_{kj}\right.$$
$$\left.-\frac{\beta^2\rho''\left(\beta\left(|p|^2-\mu\right)\right)}{8\left(|p|^2-\mu\right)}\right]\overline{\psi_h(x)}\cdot(-i\nabla+2a(x))\,\psi_h(x).$$

In order to determine the correct Ginzburg-Landau coefficient, we once again write $\sigma(\beta z)=\sigma(\beta_c z)+O(h^2)$, and see that we have

$$\frac{4}{(2\pi)^2}\sum_{k,j=1}^{2}\int_{\widetilde{C}}\frac{\mathrm{d}z}{2\pi i}\,\sigma(\beta_c z)\int_{\mathbb{R}^2}\mathrm{d}p\;|\eta(p)|^2\left[\frac{1}{2}\left(\frac{\rho\left(\beta\left(|p|^2-\mu\right)\right)}{4\left(|p|^2-\mu\right)^2}-\frac{\beta\rho'\left(\beta\left(|p|^2-\mu\right)\right)}{4\left(|p|^2-\mu\right)}\right)\delta_{kj}\right.$$
$$\left.-\frac{\beta^2\rho''\left(\beta\left(|p|^2-\mu\right)\right)}{8\left(|p|^2-\mu\right)}\right]\overline{\psi_h(x)}\cdot(-i\nabla+2a(x))\,\psi_h(x)$$
$$=\operatorname{Re}\overline{\psi_h(x)}2\mathbb{B}\left(-i\nabla+2a_h(x)\right)\psi_h(x).$$

That completes the proof. □

6.7. Proof of Theorem C

We finally show that weak limits of (ψ_h,a_h) are solutions of the Ginzburg-Landau equations.

THEOREM 6.17. *Suppose that (ψ_*,a_*) is a weak limit point of $\{(\psi_h,a_h)\}$ in $H^1(Q)\times H^1(Q)$, i.e., suppose that $(\psi_{h_n},a_{h_n})\rightharpoonup(\psi_*,a_*)$ for some sequence $h_n\to 0$. Then $(\psi_*,2a_*)$ is a weak solution of the Ginzburg-Landau equations.*

PROOF. For the first Ginzburg-Landau equation, we want to show that for any $u\in H^1(Q)$,

$$\langle(-i\nabla+2a_*)\,u\,|\mathbb{B}\,(-i\nabla+2a_*)\,\psi_*\rangle_{L^2(Q)}+\left\langle u\,\middle|-C_1\psi_*+C_2|\psi_*|^2\psi_*\right\rangle_{L^2(Q)}=0.$$

Now we know that for all h we have

$$0=\left\langle u\,\middle|\,h^{-2}\mathcal{Q}_hP_\varepsilon\left(\left(1-V^{1/2}L_{T_h,A_h}V^{1/2}\right)\phi_h+\frac{1}{2}V^{1/2}N_{T_h,A_h}\left(-2V^{1/2}\phi\right)\right)\right\rangle_{L^2(Q)}$$
$$=\left\langle u\,\middle|(-i\nabla+2a_h)\cdot\mathbb{B}\,(-i\nabla+2a_h)\,\psi_h-C_1\psi_h+C_2|\psi_h|^2\psi_h\right\rangle_{L^2(Q)}+\langle u\,|R_h\rangle_{L^2(Q)},$$

where $\|R_h\|_{L^2(Q)}\to 0$ as $h\to 0$. This means we need to show that

$$\lim_{n\to\infty}\left\langle u\,\middle|(-i\nabla+2a_h)\cdot\mathbb{B}\,(-i\nabla+2a_h)\,\psi_h-C_1\psi_h+C_2|\psi_h|^2\psi_h\right\rangle_{L^2(Q)}$$
$$=\langle(-i\nabla+2a_*)\,u\,|\mathbb{B}\,(-i\nabla+2a_*)\,\psi_*\rangle_{L^2(Q)}+\left\langle u\,\middle|-C_1\psi_*+C_2|\psi_*|^2\psi_*\right\rangle_{L^2(Q)}.$$

We do this term by term. For the first term we have

$$\begin{aligned}\langle u\,|(-i\nabla)\cdot\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)} &= \langle(-i\nabla)\,u\,|\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)}\\ &\to \langle(-i\nabla)\,u\,|\mathbb{B}\,(-i\nabla\psi_*)\rangle_{L^2(Q)}\,,\end{aligned}$$

since $\psi_{h_n} \rightharpoonup \psi_*$ in $H^1(Q)$. For the second term, we note that since Q is compact and $a_{h_n} \to a_*$ weakly in $H^1(Q)$, $a_{h_n} \to a_*$ strongly in $L^p(Q)$ for $1 \leq p \leq 6$. This means

$$\begin{aligned}&\langle u\,|a_{h_n}\cdot\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)}\\ &\quad= \langle(a_{h_n}-a_*)\,u\,|\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)} + \langle a_*u\,|\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)}\\ &\qquad\to \langle a_*u\,|\mathbb{B}\,(-i\nabla\psi_*)\rangle_{L^2(Q)}\,,\end{aligned}$$

where we used that fact that $\|\nabla\psi_h\|_{L^2(Q)} \lesssim 1$ to see that

$$\begin{aligned}\left|\langle(a_{h_n}-a_*)\,u\,|\mathbb{B}\,(-i\nabla\psi_{h_n})\rangle_{L^2(Q)}\right| &\lesssim \|(a_{h_n}-a_*)\,u\|_{L^2(Q)}\,\|\nabla\psi_{h_n}\|_{L^2(Q)}\\ &\lesssim \|a_{h_n}-a_*\|_{L^4(Q)}\,\|u\|_{L^4(Q)}\\ &\lesssim \|a_{h_n}-a_*\|_{L^4(Q)}\,\|u\|_{H^1(Q)} \to 0.\end{aligned}$$

We similarly have

$$\begin{aligned}&\langle u\,|(-i\nabla)\cdot\mathbb{B}a_{h_n}\psi_{h_n}\rangle_{L^2(Q)}\\ &\quad= \langle(-i\nabla)\,u\,|\mathbb{B}\,(a_{h_n}-a_*)\,\psi_{h_n}\rangle_{L^2(Q)} + \langle(-i\nabla)\,u\,|\mathbb{B}a_*\psi_{h_n}\rangle_{L^2(Q)}\\ &\qquad\to \langle(-i\nabla)\,u\,|\mathbb{B}a_*\psi_*\rangle_{L^2(Q)}\,,\end{aligned}$$

and

$$\begin{aligned}&\langle u\,|a_{h_n}\cdot\mathbb{B}a_{h_n}\psi_{h_n}\rangle_{L^2(Q)}\\ &= \langle(a_{h_n}-a_*)\,u\,|\mathbb{B}\,(a_{h_n}-a_*)\,\psi_{h_n}\rangle_{L^2(Q)} + \langle a_*u\,|\mathbb{B}\,(a_{h_n}-a_*)\,\psi_{h_n}\rangle_{L^2(Q)}\\ &\qquad + \langle(a_{h_n}-a_*)\,u\,|\mathbb{B}a_*\psi_{h_n}\rangle_{L^2(Q)} + \langle a_*u\,|\mathbb{B}a_*\psi_{h_n}\rangle_{L^2(Q)}\\ &\qquad\qquad\to \langle a_*u\,|\mathbb{B}a_*\psi_*\rangle_{L^2(Q)}\,.\end{aligned}$$

The term with $\langle u|\psi_{h_n}\rangle_{L^2(Q)}$ is trivial and for the term with $\langle u||\psi_{h_n}|^2\psi_{h_n}\rangle_{L^2(Q)}$ we proceed similarly as above.

We now turn to the second Ginzburg-Landau equation, where we wish to show that for any $u \in H^1(Q)$,

$$\langle\operatorname{curl} u\,|C_3\operatorname{curl} a_*\rangle_{L^2(Q)} - \left\langle u\,\middle|\operatorname{Re}\overline{\psi}_*\mathbb{B}\,(-i\nabla+2a_*)\,\psi_*\right\rangle_{L^2(Q)} = 0.$$

The strategy and the arguments are very similar to the case of the first Ginzburg-Landau equation. Therefore, we omit the details here.

□

6.8. Appendix: Proof of the lemmas in Section 6.2.1

Proof of Lemma 6.2. The first estimate follows essentially from the Sobolev inclusion $H^1(Q_h) \subseteq L^p(Q_h)$,

$$\|u\|_{L^p(Q_h)} \lesssim \|u\|_{L^2(Q_h)} + h^{-1}\,\|\nabla u\|_{L^2(Q_h)}\,.$$

Let $\phi_{(k)}$ denote ϕ when k is even and $\overline{\phi}$ when k is odd. We now calculate that

$$\begin{aligned}\left\|V^{1/2}\phi\right\|_{\mathcal{L}_p}^p &= \operatorname{Tr}\chi_{Q_h}\left(\left(V^{1/2}\phi\right)^*\left(V^{1/2}\phi\right)\right)^{p/2}\chi_{Q_h}\\ &= \prod_{k=1}^{p}\int_{\mathbb{R}^2}\mathrm{d}x_k\,\chi_{Q_h}\left(x_1\right)V^{1/2}\left(x_k-x_{k+1}\right)\phi_{(k)}\left(x_k,x_{k+1}\right),\end{aligned}$$

where $x_{p+1}=x_1$. We now use the change of variables

$$X=x_1, r_1=x_1-x_2, r_2=x_2-x_3,\ldots,r_{p-1}=x_{p-1}-x_p,$$

and we define $r_p:=-r_1-\cdots-r_{p-1}=x_p-x_1$. We also define S_i to be the linear combination of r_i satisfying the condition that $(x_i+x_{i+1})/2=X-S_i$. We then have

$$\begin{aligned}&\left\|V^{1/2}\phi\right\|_{\mathcal{L}_p}^p\\ &\leq\prod_{m=1}^{p-1}\int_{\mathbb{R}^2}\mathrm{d}r_m\prod_{k=1}^{p}V^{1/2}(r_k)\left(\int_{Q_h}\mathrm{d}X\ \left|\phi(X-S_k+r_k/2,X-S_k-r_k/2)\right|^p\right)^{1/p}\\ &\leq\prod_{m=1}^{p-1}\int_{\mathbb{R}^2}\mathrm{d}r_m\prod_{k=1}^{p}V^{1/2}(r_k)\left\|\phi\left(\cdot+r_k/2,\cdot-r_k/2\right)\right\|_{L^p(Q_h)}\\ &=\int_{\mathbb{R}^2}\mathrm{d}r_1\,t(r_1)(t*\cdots*t)(-r_1),\end{aligned}$$

where

$$t(r)=V^{1/2}(r)\left(\left\|\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}+h^{-1}\left\|\nabla_X\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}\right).$$

We therefore have that

$$\left\|V^{1/2}\phi\right\|_{\mathcal{L}_p}^p\lesssim\|t\|_1\|t*\cdots*t\|_\infty\leq\|t\|_1^{p-2}\|t*t\|_\infty\leq\|t\|_1^{p-2}\|t\|_2^2.$$

Now we calculate that

$$\begin{aligned}\|t\|_1&\lesssim\left(\int_{\mathbb{R}^2}\mathrm{d}r\,V(r)\right)^{1/2}\\ &\times\left(\int_{\mathbb{R}^2}\mathrm{d}r\ \left\|\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}^2+h^{-2}\left\|\nabla_X\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}^2\right)^{1/2}\\ &\lesssim\|V\|_1\left(\|\phi\|_{\mathcal{L}_h^2}+h^{-1}\|\nabla_X\phi\|_{\mathcal{L}_h^2}\right),\end{aligned}$$

and

$$\begin{aligned}\|t\|_2^2&\lesssim\int_{\mathbb{R}^2}\mathrm{d}r\,V(r)\left(\left\|\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}^2+h^{-2}\left\|\nabla_X\phi\left(\cdot+r/2,\cdot-r/2\right)\right\|_{L^2(Q_h)}^2\right)\\ &\leq\|V\|_\infty\left(\|\phi\|_{\mathcal{L}_h^2}+h^{-1}\|\nabla_X\phi\|_{\mathcal{L}_h^2}\right)^2.\end{aligned}$$

This finishes the proof of the first estimate. The second and the third estimate are easy to see and we do not carry out the proofs here.

In order to prove the last estimate, we define $\Psi \in L^2(Q_h)$ by the condition that $P_\varepsilon \phi(x,y) = \varphi_*(x-y)\Psi((x+y)/2)$. We now write $\Psi = \Psi_0 + \Psi_1$, where $\Psi_0 \in \mathbb{C}$ is given by

$$\Psi_0 = h^2 \int_{Q_h} \mathrm{d}X\, \Psi(X).$$

One can easily check that $|\Psi_0| \leq \|\Psi\|_{L^2(Q_h)} = \|P_\varepsilon \phi\|_{\mathcal{L}^2_h}$. Note that Ψ_1 is then such that its Fourier transform satisfies $\hat{\Psi}_1(0) = 0$. This allows us to estimate

$$\begin{aligned} |\Psi_1(X)| \leq \sum_{\substack{m\in(2\pi\mathbb{Z})^2 \\ 0<h|m|\leq\varepsilon}} \left|\hat{\Psi}(m)\right| &\leq \left(\sum_{0<h|m|\leq\varepsilon} |m|^{-2}\right)^{1/2} \left(\sum_{0<h|m|\leq\varepsilon} |m|^2 \left|\hat{\Psi}(m)\right|\right)^{1/2} \\ &\leq h^{-1} \|\nabla \Psi_1\|_{L^2(Q_h)} \left(\log \frac{\varepsilon}{h}\right)^{1/2}. \end{aligned}$$

We now turn to $V^{1/2} P_\varepsilon \phi$ as an operator on $L^2(\mathbb{R}^2)$. Note that $V^{1/2}\varphi_* \Psi_0$ is simply a convolution operator and we have

$$\left\|V^{1/2}\varphi_*\Psi_0\right\|_\infty = |\Psi_0| \, \|V\alpha_*\|_{L^1} \lesssim \|P_\varepsilon\phi\|_{\mathcal{L}^2_h}.$$

For $V^{1/2}\varphi_*\Psi_1$, we see that for any $u, v \in L^2(\mathbb{R}^2)$, we have

$$\begin{aligned} \left|\left\langle u \left| V^{1/2}\phi v \right.\right\rangle\right| &= \left| \int_{\mathbb{R}^2} \mathrm{d}x\, \overline{u(x)} \int_{\mathbb{R}^2} \mathrm{d}y\, V^{1/2}(x-y)\varphi_*(x-y)\Psi_1((x+y)/2)v(y) \right| \\ &\lesssim \|\Psi_1\|_{L^\infty} \int_{\mathbb{R}^2} \mathrm{d}r\mathrm{d}Y \left| (V\alpha_*)(r) \overline{u\,(Y + r/2)} v\,(Y - r/2) \right| \\ &\lesssim \|\Psi_1\|_{L^\infty} \|V\alpha_*\|_{L^1} \|u\|_{L^2} \|v\|_{L^2}, \end{aligned}$$

and that establishes the result.

□

Proof of Lemma 6.3. Assume $u \in L^2(\mathbb{R}^2)$ is smooth and compactly supported. We start by calculating that

$$\begin{aligned} \left| \int_{\mathbb{R}^2} \mathrm{d}y\, K(x,y)u(y) \right| &\leq \int_{\mathbb{R}^2} \mathrm{d}y\, g(x-y)\, |u(y)| \\ &\leq \left(\int_{\mathbb{R}^2} \mathrm{d}y\, g(x-y)\right)^{1/2} \left(\int_{\mathbb{R}^2} \mathrm{d}y\, g(x-y)\, |u(y)|^2\right)^{1/2}, \end{aligned}$$

and this gives

$$\begin{aligned} \left(\int_{R^2} \mathrm{d}x \left| \int_{\mathbb{R}^2} \mathrm{d}y\, K(x,y)u(y) \right|^2\right)^{1/2} &\leq \|g\|_{L^1}^{1/2} \left(\int_{\mathbb{R}^2} \mathrm{d}x\mathrm{d}y\, g(x-y)|u(y)|^2\right)^{1/2} \\ &= \|g\|_{L^1}^{1/2} \left(\int_{\mathbb{R}^2} \mathrm{d}x\, g(x) \int_{\mathbb{R}^2} \mathrm{d}y\, |u(y)|^2\right)^{1/2} \end{aligned}$$

and the lemma now follows. For the second claim, we simply calculate that

$$\left|(K_1K_2)\left(\zeta_X^r,\zeta_X^{-r}\right)\right| \le \int_{\mathbb{R}^2} \mathrm{d}w\, g_1(\zeta_X^r - w) g_2(w-\zeta_X^r) = (g_1 * g_2)(r),$$

which proves the lemma. □

Proof of Lemma 6.4. Let $u \in L^2(Q_h)$ be arbitrary. We let M_u denote the multiplication operator on $L^2(\mathbb{R}^2)$ defined by multiplication by $u\chi_{Q_h}$. We then have

$$\begin{aligned}
&\left|h^2 \int_{Q_h} \mathrm{d}x\, K(x,x)u(x)\right| \\
&\quad = \left|h^2 \operatorname{Tr}_{Q_h} KM_u\right| \\
&\quad = \left|h^2 \operatorname{Tr} \chi_{Q_h} K\,(1-\Delta)\,(1-\Delta)^{-1}\, M_u\chi_{Q_h}\right| \\
&\quad = \left(h^2 \operatorname{Tr} \chi_{Q_h} K\,(1-\Delta)^2\, K^*\chi_{Q_h}\right)^{1/2} \left(h^2 \operatorname{Tr} \chi_{Q_h} M_u^*\,(1-\Delta)^{-2}\, M_u\chi_{Q_h}\right)^{1/2}.
\end{aligned}$$

Now the second square root is given by

$$\begin{aligned}
&\left(h^2 \operatorname{Tr} \chi_{Q_h} M_u^*\,(1-\Delta)^{-2}\, M_u\chi_{Q_h}\right)^{1/2} \\
&\qquad = \left(\int_{\mathbb{R}^2} \frac{\mathrm{d}p}{\left(1+|p|^2\right)^2}\right)^{1/2} \left(h^2 \int_{Q_h} \mathrm{d}x\, \left|u(x)\right|^2\right)^{1/2} \lesssim \left\|u\right\|_{L^2(Q_h)}.
\end{aligned}$$

The lemma now follows from the cyclicity of the trace per unit volume. □

CHAPTER 7

Entropy decay for the Kac evolution

F. Bonetto, A. Geisinger, M. Loss, T. Ried

We consider solutions to the Kac master equation for initial conditions where N particles are in a thermal equilibrium and $M \leq N$ particles are out of equilibrium. We show that such solutions have exponential decay in entropy relative to the thermal state. More precisely, the decay is exponential in time with an explicit rate that is essentially independent on the particle number. This is in marked contrast to previous results which show that the entropy production for arbitrary initial conditions is inversely proportional to the particle number. The proof relies on Nelson's hypercontractive estimate and the geometric form of the Brascamp-Lieb inequalities due to Franck Barthe. Similar results hold for the Kac-Boltzmann equation with uniform scattering cross sections.

7.1. Introduction

Among the models describing a gas of interacting particles, the Kac master equation [**61**], due to its simplicity, occupies a special place. It is useful in illuminating various issues in kinetic theory, e.g., providing a reasonably satisfactory derivation of the spatially homogeneous Boltzmann equation and giving a mathematical framework for investigating the approach to equilibrium. These issues were, in fact, the motivation for Kac's original work [**61**]. Although it does not have a foundation in Hamiltonian mechanics, the Kac master equation is based on simple probabilistic principles and yields a linear evolution equation for the velocity distribution for N particles undergoing collisions. It is in this context that Kac invented the notion of propagation of chaos and he used this notion to derive the spatially homogeneous, non-linear Kac-Boltzmann equation. The approach through master equations led Kac to formulate the notion of approach to equilibrium and suggested various avenues to investigate this problem as the number of particles, N, becomes large. He emphasized that this could be done in a quantitative way if one could show, e.g., that the gap of the generator is bounded below uniformly in N. This, known as Kac's conjecture [**61**], was proved by Élise Janvresse in [**60**] and, as a further sign of the simplicity of the model, the gap was computed explicitly in [**19, 20**], see also [**72**]. One of the problems in using the gap is that the approach to equilibrium is measured in terms of an L^2 distance. While this does seem to be a natural way to look at this problem, the size of the L^2 norm of approximately independent probability distributions increases exponentially with the size of the system. Thus, the half life of the L^2 norm is of order N.

A natural measure is, of course, given by the entropy, which is extensive, i.e, proportional to N. There has not been much success in proving exponential decay of the entropy with good rates. In **[82]** Cedric Villani showed that the entropy decays exponentially, albeit with a rate that is bounded below by a quantity that is inversely proportional to N. This estimate was complemented by Amit Einav **[29]**, who gave an example of a state that has entropy production essentially of order $1/N$. His example is the initial state in which most of the energy is concentrated in a few particles while most of the others have very little energy. One might surmise, based on physical intuition, that this state is physically very improbable and still has low entropy production because most of the particles are in some sort of equilibrium. This intuition can be made rigorous, see **[29]**, although by a quite difficult computation. One should add that low entropy production does not preclude exponential decay in entropy, i.e., large entropy production for the initial state might not be necessary for an exponential decay rate for the entropy.

A breakthrough was achieved by Mischler and Mouhot in **[74, 73]**. They undertook a general investigation of the Kac program for gases of hard spheres and true Maxwellian molecules in three dimensions. Among the results of Mischler and Mouhot is a proof that these systems relax towards equilibrium in relative entropy as well as in Wasserstein distance with a rate that is independent of the particle number. As expected, they achieve this not for any initial condition, but rather for a natural class of chaotic states. The rate of relaxation is, however, polynomial in time.

To summarize, there is so far no mathematical evidence that the entropy in the Kac model in general decays exponentially with a rate that is independent of N and physical intuition suggests that for highly "improbable" states, such as the one used by Einav, this cannot be expected. One can restrict the class of initial conditions by considering chaotic states as done by Mischler and Mouhot, which shifts the problem of finding suitable initial conditions for proving exponential decay to the level of the non-linear Boltzmann equation.

In this paper we take a different approach, one which is based on the idea of coupling a system of particles to a reservoir. Recall from **[15]** the master equation of M particles with velocities $\mathbf{v} = (v_1, v_2, \ldots, v_M)$ interacting with a thermostat at temperature $1/\beta$,

$$\frac{\partial f}{\partial t} = \mathcal{L}_T f \ , \ f(\mathbf{v}, 0) = f_0(\mathbf{v}) \ . \tag{7.1}$$

The operator $\mathcal{L}_T$ is given by

$$\mathcal{L}_T f = \mu \sum_{j=1}^{M} (B_j - I) f \ ,$$

where

$$B_j[f](\mathbf{v}) := \int_{\mathbb{R}} \mathrm{d}w\,\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}\theta\,\sqrt{\frac{\beta}{2\pi}}\,e^{-\beta w_j^*(\theta)^2/2} f(\mathbf{v}_j(\theta,w))\ ,$$

$$\mathbf{v}_j(\theta,w) = (v_1,...,v_j\cos(\theta)+w\sin(\theta),...,v_M) \text{ and } w_j^*(\theta) = -v_j\sin(\theta)+w\cos(\theta)\ .$$

Thus, $B_j[f](\mathbf{v})$ describes the effect of a collision between particle j in the system and a particle in the reservoir. After the collision, the particle from the thermostat is discarded, which ensures that the thermostat stays in equilibrium. The interaction times with the thermostat are given by a Poisson process whose intensity μ is chosen so that the average time between two successive interactions of a given particle with the thermostat is independent of the number of particles in the system. For the case where $\rho(\theta) = (2\pi)^{-1}$, the entropy decays exponentially fast. In fact, abbreviating $\sqrt{\beta/(2\pi)}e^{-\beta/2\mathbf{v}^2} = \Gamma_\beta(\mathbf{v})$, we know from **[15]**, that

$$S(f(\cdot,t)) := \int_{\mathbb{R}^M} f(\mathbf{v},t)\log\left(\frac{f(\mathbf{v},t)}{\Gamma_\beta(\mathbf{v})}\right)d\mathbf{v} \le e^{-\mu t/2}S(f_0)\ .$$

Thus, one might guess that if a "small" *system* of M particles out of equilibrium interacts with a *reservoir*, that is a large system of $N \ge M$ particles in thermal equilibrium, then the entropy decays exponentially fast in time. This intuition is also supported by the results in **[14]**. There it was shown that if the thermostat is replaced by a large but finite reservoir *initially* in thermal equilibrium, this evolution is close to the evolution given by the thermostat. This results holds in various norms and, in particular, it is *uniform* in time. We would like to emphasize that the reservoir will not stay in thermal equilibrium as time progresses, nevertheless it will not veer far from it.

Since this is the model that we consider in this work, we will now describe it in detail. We consider probability distributions $F : \mathbb{R}^{M+N} \to \mathbb{R}_+$ and write $F(\mathbf{v},\mathbf{w})$ where $\mathbf{v} = (v_1,\dots,v_M)$ describes the particles in the small system, whereas $\mathbf{w} = (w_{M+1},\dots,w_{N+M})$ describes the particles in the large system. The Kac master equation is given by

$$\frac{\partial F}{\partial t} = \mathcal{L}F\ ,\ F(\mathbf{v},\mathbf{w},0) = F_0(\mathbf{v},\mathbf{w}) = f_0(\mathbf{v})e^{-\pi|\mathbf{w}|^2}\ , \tag{7.2}$$

where

$$\mathcal{L} = \frac{\lambda_S}{M-1}\sum_{1\le i<j\le M}(R_{ij}-I) + \frac{\lambda_R}{N-1}\sum_{M<i<j\le N+M}(R_{ij}-I) + \frac{\mu}{N}\sum_{i=1}^{M}\sum_{j=M+1}^{M+N}(R_{ij}-I)\ , \tag{7.3}$$

and R_{ij} is given as follows. For $1 \le i < j \le M$ we have

$$(R_{ij}F)(\mathbf{v},\mathbf{w}) = \int_{-\pi}^{\pi}\rho(\theta)\,\mathrm{d}\theta\,F(r_{ij}(\theta)^{-1}(\mathbf{v},\mathbf{w}))\ ,$$

where

$$r_{ij}(\theta)^{-1}(\mathbf{v},\mathbf{w}) = (v_1,\dots,v_i\cos\theta - v_j\sin\theta,\dots,v_i\sin\theta+v_j\cos\theta,\dots,v_M,\mathbf{w})\ . \tag{7.4}$$

The other R_{ij}s are defined analogously. We assume that the probability measure ρ is smooth and satisfies

$$\int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\, \sin\theta \cos\theta = 0\,. \tag{7.5}$$

In particular, we do *not* require $\mathcal{L}$ to be self-adjoint on $L^2(\mathbb{R}^{N+M})$, a condition called *microscopic reversibility.* The initial state of the reservoir is assumed to be a thermal equilibrium state and we have chosen units in which the inverse temperature $\beta = 2\pi$. Note that λ_S is the rate at which one particle from the system will scatter with any other particle in the system and similarly for λ_R. Likewise, μ is the rate at which a single particle of the system will scatter with any particle in the reservoir. The rate at which a particular particle from the reservoir will scatter with a particle in the system is given by $\mu M/N$. Hence, when N is large compared to M this process is suppressed and one expects that the reservoir does not move far from its equilibrium. Indeed, it is shown in [**14**] that the solution of the master equation (7.3) stays close to the solution of a thermostated system in the Gabetta-Toscani-Wennberg metric,

$$d_{GTW}(F,G) := \sup_{k\neq 0} \frac{|\widehat{F}(k) - \widehat{G}(k)|}{|k|^2}\ ,$$

see [**44**]. Here, $\widehat{F}$ denotes the Fourier transform of F. More precisely, with the initial conditions (7.1) and (7.2), it was shown that

$$d_{GTW}(f(\mathbf{v},t)e^{-\pi|\mathbf{w}|^2}, F(\mathbf{v},\mathbf{w},t)) \le C(f_0)\frac{M}{N}\ ,$$

where $C(f_0)$ is a constant that depends on the initial condition but is of order one. The distance varies inversely as N, the size of the reservoir and, moreover, this estimate holds *uniformly* in time. For a detailed description of the results we refer the reader to [**14**]. From this result and the fact that the entropy of the system interacting with a thermostat decays exponentially in time, one might surmise that the entropy of the system interacting with a finite reservoir also decays exponentially fast in time. In fact we shall show this to be true if we consider the entropy relative to the thermal state.

7.2. Results

For the solution of the master equation (7.2) we use use interchangeably the notation

$$F(\mathbf{v},\mathbf{w},t) = (e^{\mathcal{L}t}F_0)(\mathbf{v},\mathbf{w}). \tag{7.6}$$

The energy is preserved under this evolution and hence it suffices to consider it on $L^1(\mathbb{S}^{N+M}(\sqrt{N+M}))$ with the normalized surface measure. Likewise, it is easy to see that the evolution is ergodic on $\mathbb{S}^{N+M}(\sqrt{N+M})$ in the sense that $e^{\mathcal{L}t}F_0 \to 1$ as $t\to\infty$ and 1 is the only normalized equilibrium state.

For our purposes it is convenient to consider the evolution in $L^1(\mathbb{R}^{M+N})$ with Lebesgue measure. Then $e^{\mathcal{L}t}F_0$ converges to the spherical average of F_0 taken over spheres in $\mathbb{R}^{M+N}$. In this space we choose the initial condition

$$F_0(\mathbf{v},\mathbf{w}) = f_0(\mathbf{v})e^{-\pi|\mathbf{w}|^2}\ . \tag{7.7}$$

Moreover, we introduce the function f,

$$f(\mathbf{v},t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}t}F_0\right](\mathbf{v},\mathbf{w})\,\mathrm{d}\mathbf{w} \tag{7.8}$$

and we call

$$S(f(\cdot,t)) := \int_{\mathbb{R}^M} f(\mathbf{v},t)\log\left(\frac{f(\mathbf{v},t)}{e^{-\pi|\mathbf{v}|^2}}\right)\mathrm{d}\mathbf{v}\ ,$$

the entropy of f *relative* to the thermal state $e^{-\pi|\mathbf{v}|^2}$. Our main result is the following theorem.

THEOREM 7.1. *Let $N \geq M$ and let ρ be a probability distribution with an absolutely convergent Fourier series such that* (7.5) *holds. The entropy of f relative of to the thermal state $e^{-\pi|\mathbf{v}|^2}$ then satisfies*

$$S(f(\cdot,t)) \leq \left[\frac{M}{N+M} + \frac{N}{N+M}e^{-t\mu_\rho(N+M)/N}\right] S(f_0)\ ,$$

where

$$\mu_\rho = \mu\int_{-\pi}^{\pi}\rho(\theta)\,\mathrm{d}\theta\,\sin^2(\theta)\ ,$$

and f_0 is as introduced in (7.7).

REMARK.
1. Note that the theorem deals with the entropy relative to the thermal state and not with respect to the equilibrium state. The entropy relative to the equilibrium state tends to zero as $t \to \infty$. We do not know how to adapt our proof to this situation nor do we have any evidence that it does indeed tend to zero at an exponential rate. If this were the case, the rate would most likely depend on the initial condition.
2. The decay rate is universal in the sense that it only depends on μ and the distribution ρ. The intra-particle interactions in the system and in the reservoir do not seem to matter.
3. The statement of the theorem becomes particularly simple as $N \to \infty$. This corresponds to the thermostat problem treated in [**15**] with the exact same decay rate. It is known that for the thermostat the decay rate is optimal, see [**80**], and hence the decay rate here is optimal as well.
4. Although we assume that ρ is smooth, our result also holds for the case where ρ is a finite sum of Dirac measures. In particular Theorem 7.1 also holds if ρ is a delta measure that has its mass at the angles $\theta = \pm\pi/2$, that is, our result does not depend on ergodicity of the evolution.

As a consequence of Remark 7.2(2), one obtains a result for the standard Kac model. Recall that the generator of the standard Kac model is given by

$$\mathcal{L}_{\mathrm{cl}} = \frac{2}{N+M-1}\sum_{1\leq i<j\leq N+M}(R_{ij}-I)\ .$$

We may arbitrarily split the variables into two groups, that is $(v_1, \ldots, v_M)$ and $(w_{M+1}, \ldots, w_{M+N})$. Splitting the generator accordingly,

$$\mathcal{L}_{\rm cl} = \frac{2}{N+M-1} \sum_{1\le i<j\le M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{M+1\le i<j\le N+M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{i=1}^{M} \sum_{j=M+1}^{N+M} (R_{ij} - I) ,$$

we see that the standard Kac model can be cast in the from (7.3) by setting

$$\lambda_S = \frac{2(M-1)}{N+M-1} \quad , \quad \lambda_R = \frac{2(N-1)}{N+M-1} \quad \text{and} \quad \mu = \frac{2N}{N+M-1} \ .$$

Hence, we obtain the following Corollary.

COROLLARY 7.2. *Let $N \geq M$ and consider the time evolution defined by $\mathcal{L}_{\rm cl}$ with initial condition* (7.7)*. Assume that the function f_0 in the initial condition has finite entropy. The entropy of the function*

$$f(\mathbf{v}, t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}_{\rm cl} t} F_0\right] (\mathbf{v}, \mathbf{w}) \, {\rm d}\mathbf{w}$$

relative to the thermal state $e^{-\pi|\mathbf{v}|^2}$, satisfies

$$S(f(\cdot, t)) \leq \left[\frac{M}{N+M} + \frac{N}{N+M} e^{-t\mu_\rho 2(N+M)/(N+M-1)}\right] S(f_0) \ ,$$

where

$$\mu_\rho = \int_{-\pi}^{\pi} \rho(\theta) \, {\rm d}\theta \, \sin^2(\theta)$$

and ρ is a probability distribution such that (7.5) *holds.*

On a mathematical level, an efficient way of proving approach to equilibrium is through a logarithmic Sobolev inequality, which presupposes that the generator of the time evolution is given by a Dirichlet form. This kind of structure is notably absent in the Kac master equation. We shall see however, that the logarithmic Sobolev inequality in the form of Nelson's hypercontractive estimate is an important tool for the proof of Theorem 7.1. We will use an iterated version of it, which expresses the result in terms of marginals of the functions involved. This, coupled with an auxiliary computation and a sharp version of the Brascamp-Lieb inequalities **[18]** (see also **[70]**) will lead to the result.

In our opinion, the main result of this paper is the description of a simple mechanism for obtaining exponential relaxation towards equilibrium. One can extend the results to three dimensional momentum preserving collisions, however, so far only for a caricature of Maxwellian molecules. To carry this method over to the case of hard spheres and for true Maxwellian molecules is an open problem.

The plan of the paper is as follows: In Section 7.3 we derive a representation formula for the Kac evolution ${\rm e}^{\mathcal{L}t}$ which is reminiscent of the Ornstein-Uhlenbeck process. This allows us to prove an entropy inequality based upon Nelson's hypercontractive estimate in Section 7.4. In Section 7.5 we show how the sharp version of the geometric

Brascamp-Lieb inequality leads to a correlation inequality for the entropy involving marginals, which in turn proves our main entropy inequality. The fact that our Brascamp-Lieb datum is geometric relies on a sum rule which will be proved in Section 7.6. A short proof of the geometric form of the Brascamp-Lieb inequalities is deferred to Appendix 7.8.1, as well as some technical details to ensure its applicability in Appendix 7.8.2. In Section 7.7 we show how our method can be applied to three-dimensional Maxwellian collisions with a very simple angular dependence.

Acknowledgements: A.G. and T.R. would like to thank Georgia Tech for its hospitality. The work of M.L. and A.G. was supported in part by NSF grant DMS- 1600560 and the Humboldt Foundation. A.G. and T.R. gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) through GRK 1838 (A.G.) and CRC 1173 (T.R.). T.R. thanks the Karlsruhe House of Young Scientists (KHYS) for a Research Travel Grant supporting the stay at Georgia Tech.

7.3. The representation formula

The aim of this section is to rewrite (7.6), that is $e^{\mathcal{L}t}F_0$, in a way which is reminiscent of the Ornstein-Uhlenbeck process. This representation will naturally lead to the next step in the proof of Theorem 7.1, namely the entropy inequality that will be presented in Theorem 7.5.

It is convenient to write

$$\mathcal{L} = \Lambda(Q - I)\ , \ \text{where } \Lambda = \lambda_S \frac{M}{2} + \lambda_R \frac{N}{2} + \mu M\ ,$$

and the operator Q is a convex combination of R_{ij}s, given by

$$Q = \frac{\lambda_S}{\Lambda(M-1)} \sum_{1\le i<j\le M} R_{ij} + \frac{\lambda_R}{\Lambda(N-1)} \sum_{M<i<j\le N+M} R_{ij} + \frac{\mu}{\Lambda N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} R_{ij}\ ,$$

i.e., Q is an average over rotation operators. The right hand side of (7.6) can be written as

$$(e^{\mathcal{L}t}F_0)(\mathbf{v},\mathbf{w}) = e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k\Lambda^k}{k!} Q^k F_0(\mathbf{v},\mathbf{w})\ , \tag{7.9}$$

where

$$Q^k F_0(\mathbf{v},\mathbf{w}) \\ = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\mathrm{d}\theta_k\, F_0\left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]^{-1}(\mathbf{v},\mathbf{w})\right)\ . \tag{7.10}$$

Here, α labels pairs of particles, that is, $\alpha = (i, j)$, $1 \le i < j \le M + N$, $r_\alpha(\theta)$ is defined in (7.4) and λ_α is given by the rotation corresponding to the index α, that is,

$$\lambda_{(i,j)} = \frac{\lambda_S}{\Lambda(M-1)} \quad \text{if} \quad 1 \le i < j \le M \ ,$$

$$\lambda_{(i,j)} = \frac{\lambda_R}{\Lambda(N-1)} \quad \text{if} \quad M+1 \le i < j \le M+N \ ,$$

$$\lambda_{(i,j)} = \frac{\mu}{\Lambda N} \quad \text{if} \quad 1 \le i \le M \ , M+1 \le j \le M+N \ .$$

Note that the sum over *all* pairs $\sum_\alpha \lambda_\alpha \ = \ 1$.

For our purpose, it is convenient to write the function f_0, introduced in (7.7), as $f_0(\mathbf{v}) = h_0(\mathbf{v}) e^{-\pi|\mathbf{v}|^2}$. Since the Gaussian function is invariant under rotations, (7.9) takes the form

$$(e^{\mathcal{L}t} F_0)(\mathbf{v}, \mathbf{w}) = e^{-\pi\left(|\mathbf{v}|^2 + |\mathbf{w}|^2\right)} e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} Q^k \left(h_0 \circ P\right)(\mathbf{v}, \mathbf{w}) \ .$$

We introduce the projection $P : \mathbb{R}^{N+M} \to \mathbb{R}^M$ by $P(\mathbf{v}, \mathbf{w}) = \mathbf{v}$, as a reminder that the semigroup $e^{\mathcal{L}t}$ acts on functions that depend on $\mathbf{v}$ as well as $\mathbf{w}$. If we write

$$f(\mathbf{v}, t) = e^{-\pi|\mathbf{v}|^2} h(\mathbf{v}, t) \ ,$$

then (7.8) can be written as

$$h(\mathbf{v}, t) = e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} h_k(\mathbf{v}) \ ,$$

where the functions h_k are given by

$$h_k(\mathbf{v}) := \int_{\mathbb{R}^N} Q^k \left(h_0 \circ P\right)(\mathbf{v}, \mathbf{w}) e^{-\pi|\mathbf{w}|^2} \, \mathrm{d}\mathbf{w} \ .$$

Likewise, the entropy of f is expressed as

$$S(f(\cdot, t)) = \int_{\mathbb{R}^M} h(\mathbf{v}, t) \log h(\mathbf{v}, t) e^{-\pi|\mathbf{v}|^2} \, \mathrm{d}\mathbf{v} =: \mathcal{S}(h(\cdot, t)) \ .$$

Expanding the function $Q^k(h_0 \circ P)(\mathbf{v}, \mathbf{w})$, we find that

$$h_k(\mathbf{v}) = \sum_{\alpha_1, \ldots, \alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1) \, \mathrm{d}\theta_1 \cdots \rho(\theta_k) \mathrm{d}\theta_k \times$$
$$\times \int_{\mathbb{R}^N} (h_0 \circ P) \left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l) \right]^{-1} (\mathbf{v}, \mathbf{w}) \right) e^{-\pi|\mathbf{w}|^2} \, \mathrm{d}\mathbf{w} \ , \quad (7.11)$$

where, as before, see (7.10), $r_\alpha(\theta)$ rotates the plane given by the index pair α by an angle θ while keeping the other directions fixed. Since $P(\mathbf{v}, \mathbf{w}) = \mathbf{v}$, it is natural to write

$$\left[\prod_{j=1}^{k} r_{\alpha_j}(\theta_j) \right]^{-1} = \begin{pmatrix} A_k(\underline{\alpha}, \underline{\theta}) & B_k(\underline{\alpha}, \underline{\theta}) \\ C_k(\underline{\alpha}, \underline{\theta}) & D_k(\underline{\alpha}, \underline{\theta}) \end{pmatrix} \ ,$$

where $A_k \in \mathbb{R}^{M\times M}$ is an $M \times M$ matrix, $B_k \in \mathbb{R}^{M\times N}$, $C_k \in \mathbb{R}^{N\times M}$ and $D_k \in \mathbb{R}^{N\times N}$. Further, $\underline{\alpha} = (\alpha_1, \ldots, \alpha_k)$ and $\underline{\theta} = (\theta_1, \ldots, \theta_k)$. This notation allows us to rewrite (7.11) as

$$h_k(\mathbf{v}) = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1 \cdots \rho(\theta_k)\,\mathrm{d}\theta_k \times$$
$$\times \int_{\mathbb{R}^N} h_0\left(A_k(\underline{\alpha},\underline{\theta})\mathbf{v} + B_k(\underline{\alpha},\underline{\theta})\mathbf{w}\right) e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w}\ .$$

Note that, by the definition of rotations,

$$A_k(\underline{\alpha},\underline{\theta})A_k^T(\underline{\alpha},\underline{\theta}) + B_k(\underline{\alpha},\underline{\theta})B_k^T(\underline{\alpha},\underline{\theta}) = I_M\ . \tag{7.12}$$

LEMMA 7.3. *Let $A \in \mathbb{R}^{M\times M}$ and $B \in \mathbb{R}^{M\times N}$ be matrices that satisfy $AA^T + BB^T = I_M$. Then*

$$\int_{\mathbb{R}^N} h(A\mathbf{v} + B\mathbf{w})e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w} = \int_{\mathbb{R}^M} h\left(A\mathbf{v} + (I_M - AA^T)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}$$

for any integrable function h.

PROOF. Denote the range of B by $H \subset \mathbb{R}^M$ and its kernel by $K \subset \mathbb{R}^N$. We may write

$$\int_{\mathbb{R}^N} h(A\mathbf{v} + B\mathbf{w})e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w} = \int_K \int_{K^\perp} h(A\mathbf{v} + B\mathbf{u})e^{-\pi|\mathbf{u}|^2} e^{-\pi|\mathbf{u}'|^2}\,\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{u}'$$
$$= \int_{K^\perp} h(A\mathbf{v} + B\mathbf{u})e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\ .$$

The symmetric map $BB^T : \mathbb{R}^M \to \mathbb{R}^M$ has H as its range and $H^\perp$, that is the orthogonal complement of H in $\mathbb{R}^M$, as its kernel. Indeed, suppose that there exists $x \in \mathbb{R}^M$ with $BB^T x = 0$, then $B^T x = 0$, i.e., $x \in \mathrm{Ker}B^T$ or x is perpendicular to H. Hence, the map $BB^T : H \to H$ is invertible. Define the linear map $R : \mathbb{R}^N \to H$ by

$$R = \left(BB^T\right)^{-1/2} B$$

and note that $RR^T = I_H$ while $R^T R$ projects the space $K^\perp$ orthogonally onto H. Since $K^\perp$ and H have the same dimension, it follows that R^T restricted to H defines an isometry between H and $K^\perp$. Hence,

$$\int_{K^\perp} h\left(A\mathbf{v} + B\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u} = \int_{K^\perp} h\left(A\mathbf{v} + \left(BB^T\right)^{1/2} R\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}$$
$$= \int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2} RR^T\mathbf{u}\right) e^{-\pi|R^T\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}$$
$$= \int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\ .$$

The assumption $AA^T + BB^T = I_M$, together with the fact that

$$\int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}$$
$$= \int_{H^\perp} \int_H h\left(A\mathbf{v} + (BB^T)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\, e^{-\pi|\mathbf{u}'|^2}\,\mathrm{d}\mathbf{u}'$$

now implies the lemma. $\square$

The matrix $A_k(\underline{\alpha},\underline{\theta})$ has an orthogonal singular value decomposition,

$$A_k(\underline{\alpha},\underline{\theta}) = U_k(\underline{\alpha},\underline{\theta})\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta}) \ , \tag{7.13}$$

where $\Gamma_k(\underline{\alpha},\underline{\theta}) = \mathrm{diag}[\gamma_{k,1}(\underline{\alpha},\underline{\theta}),\ldots,\gamma_{k,M}(\underline{\alpha},\underline{\theta})]$ is the diagonal matrix whose entries $\gamma_{k,j}(\underline{\alpha},\underline{\theta})$, $j=1,\ldots,M$, are the singular values of $A_k(\underline{\alpha},\underline{\theta})$, and $U_k(\underline{\alpha},\underline{\theta})$ and $V_k(\underline{\alpha},\underline{\theta})$ are rotations in $\mathbb{R}^M$. Note that (7.12) implies $\gamma_{k,j}(\underline{\alpha},\underline{\theta}) \in [0,1]$ for $j=1,\ldots,M$. We shall use the abbreviation

$$h_0(U_k(\underline{\alpha},\underline{\theta})\mathbf{v}) = h_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{v}) \ .$$

These considerations can be summarized by the representation formula presented in the following theorem.

THEOREM 7.4 (Representation formula). *The function h_k can be written as*

$$h_k(\mathbf{v}) = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k\times$$
$$\times\int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta})\mathbf{v} + \left(I_M - \Gamma_k^2(\underline{\alpha},\underline{\theta})\right)^{1/2}\mathbf{w}\right)e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w} \ , \tag{7.14}$$

where $h_{0,U_k(\underline{\alpha},\underline{\theta})}$, $\Gamma_k(\underline{\alpha},\underline{\theta})$ and V_k are as defined above.

7.4. The hypercontractive estimate

Starting from (7.14) and using convexity of the entropy and Jensen's inequality together with

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k = 1 \ ,$$

we get

$$\mathcal{S}(h_k) \le \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k\,\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta})),$$

where we set

$$g_k(\mathbf{v},\underline{\alpha},\underline{\theta}) = \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\gamma_k(\underline{\alpha},\underline{\theta})\mathbf{v} + \left(I_M - \gamma_k^2(\underline{\alpha},\underline{\theta})\right)^{1/2}\mathbf{w}\right)e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w} \ , \tag{7.15}$$

and we removed the rotation $V_k^T(\underline{\alpha},\underline{\theta})$ by a change of variables.

To explain the main observation in this section we look at (7.15) when $M=1$. Since $0 \le \gamma_k(\underline{\alpha},\underline{\theta}) \le 1$, we can write $\gamma_k(\underline{\alpha},\underline{\theta}) = e^{-t}$ and we get $g_k(v,\underline{\alpha},\underline{\theta}) = N_t(h_{0,U_k(\underline{\alpha},\underline{\theta})})$ where N_t is the Ornstein-Uhlenbeck semigroup, that is

$$N_t h(x) = \int_{\mathbb{R}} h\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right)e^{-\pi y^2}\,\mathrm{d}y \ .$$

Thus Theorem 7.4 renders the function h_k as a convex combination of terms reminiscent of the Ornstein-Uhlenbeck process, albeit in matrix form. We make use of this observation to find a bound for $\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta}))$. This bound together with a suitable correlation inequality proved in the next section will lead to a bound for $\mathcal{S}(h_k)$.

In addition to the notation developed in the previous section, we need various marginals of the function $h_{0,U_k(\underline{\alpha},\underline{\theta})}$. Quite generally, if h is a function of M variables and $\sigma \subset \{1,\ldots,M\}$, we shall denote by h^σ the marginals of h with respect to the variables $v_j, j \in \sigma$, for instance,

$$h^{\{1,2\}}(v_3,\ldots,v_M) = \int_{\mathbb{R}^2} h(v_1,v_2,v_3,\ldots,v_M)e^{-\pi\left(v_1^2+v_2^2\right)}\,\mathrm{d}v_1\mathrm{d}v_2\ .$$

It will be convenient to use the matrix $P_\sigma : \mathbb{R}^M \to \mathbb{R}^{|\sigma|}$ that projects $\mathbb{R}^M$ orthogonally onto $\mathbb{R}^{|\sigma|}$ which we will identify with subspace of $\mathbb{R}^M$. To give an example, let $\mathbf{v} = (v_1,...,v_M)$. Then $P_{\{1,2\}}\mathbf{v} = (v_1,v_2)$. The following theorem is the main result of this section.

THEOREM 7.5 (Partial entropy bound). *Let $h_0 \in L^1(\mathbb{R}^M, e^{-\pi|\mathbf{v}|^2}d\mathbf{v})$ be nonnegative and assume that $\mathcal{S}(h_0) < \infty$. Then*

$$\begin{aligned}&\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta}))\\ &\le \sum_{\sigma\subset\{1,\ldots,M\}}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}^2\right)\int_{\mathbb{R}^M} h_0(\mathbf{v})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}\left(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v}\right)e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\ ,\end{aligned} \tag{7.16}$$

where σ^{c} is the complement of the set σ in $\{1,...,M\}$.

A key role in the proof of Theorem 7.5 is played by Nelson's hypercontractive estimate.

THEOREM 7.6 (Nelson's hypercontractive estimate). *The Ornstein-Uhlenbeck semigroup,*

$$N_th(x) = \int_{\mathbb{R}} h\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right)e^{-\pi y^2}\,\mathrm{d}y\ ,$$

for $t \ge 0$, is bounded from $L^p(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ to $L^q(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ if and only if

$$(p-1) \ge e^{-2t}(q-1)\ .$$

For such values of p and q,

$$\|N_th\|_q \le \|h\|_p$$

with equality if and only if h is constant.

PROOF. For a proof we refer the reader to [**75**]. For other proofs see [**48, 49, 30, 23**]. □

Nelson's hypercontractive estimate, that is Theorem 7.6, implies the following Corollary, which will be useful in the proof of Theorem 7.5.

COROLLARY 7.7 (Entropic version of Nelson's hypercontractive estimate). *Let $h : \mathbb{R} \to \mathbb{R}_+$ be a function in $L^1(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ with finite entropy, i.e.,*

$$\mathcal{S}(h) = \int_{\mathbb{R}} h(x)\log h(x)\,e^{-\pi x^2}dx < \infty\ .$$

Then

$$\mathcal{S}(N_t h) \le e^{-2t}\mathcal{S}(h) + (1-e^{-2t})\|h\|_1 \log \|h\|_1$$

for all $t \ge 0$.

PROOF. Let $h \in L^p(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$, for $p \ge 1$ small, be a nonnegative function. As $\|N_t h\|_1 = \|h\|_1$, we can apply Nelson's hypercontractive estimate, which implies that for p, q that satisfy $(p-1) = e^{-2t}(q-1)$,

$$\frac{\|N_t h\|_q - \|N_t h\|_1}{q-1} \le \frac{\|h\|_p - \|h\|_1}{q-1} = e^{-2t}\frac{\|h\|_p - \|h\|_1}{p-1} \ .$$

Sending $p \to 1$ and hence $q \to 1$, we get the claimed estimate for such functions h. If h just has finite entropy one cuts off h at large values, uses the above estimate and removes the cutoff using the monotone convergence theorem. □

We are now ready to prove Theorem 7.5.

PROOF OF THEOREM 7.5. Remember that $0 \le \gamma_{k,j}(\underline{\alpha}, \underline{\theta}) \le 1$ for $j = 1, ..., M$. Thus, by inductively applying Corollary 7.7 to

$$\int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\gamma_{k,1}v_1 + \sqrt{1-\gamma_{k,1}^2}\,u_1, \ldots, \gamma_{k,M}v_M + \sqrt{1-\gamma_{k,M}^2}\,u_M\right) \times e^{-\pi\sum_{j=1}^M u_j^2}\,\mathrm{d}u_1\cdots\mathrm{d}u_M \ ,$$

we obtain

$$\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta})) \le \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}^2 \prod_{j\in\sigma}(1-\gamma_{k,j}^2) \int_{\mathbb{R}^{|\sigma^{\mathrm{c}}|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\, e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u} \ .$$

Inserting the definition of the marginal $h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}$, we see that

$$\begin{aligned}
&\int_{\mathbb{R}^{|\sigma^{\mathrm{c}}|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\, e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u} \\
&= \int_{\mathbb{R}^M} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}\mathbf{v}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v} \\
&= \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{v}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v} \\
&= \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v},
\end{aligned}$$

which finishes the proof of Theorem 7.5. □

7.5. The key entropy bound

Collecting the results of the previous sections we get the following bound

$$\mathcal{S}(h_k) \leq \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k$$
$$\times \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}^2 \prod_{j\in\sigma}\left(1-\gamma_{k,j}^2\right) \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}\left(P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v}\right)\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}. \tag{7.17}$$

The right-hand side of (7.17) contains a large sum over the entropy of marginals of h_0. In order to bound such a sum in terms of the entropy of h_0 one may try to apply some version of the Loomis-Whitney inequality [**71**] or, more precisely, of an inequality by Han [**59**]. This is essentially correct, but will require a substantial generalization of this inequality. Let us first formulate the main theorem of this section.

THEOREM 7.8 (Entropy bound). *The estimate*

$$\mathcal{S}(h_k) \leq \left[\frac{M}{N+M} + \frac{N}{N+M}\left(1-\mu_\rho\frac{N+M}{N\Lambda}\right)^k\right]\mathcal{S}(h_0) \tag{7.18}$$

holds.

As mentioned before, to prove Theorem 7.8, we need a generalized version of an inequality by Han. This generalization was proven by Carlen-Cordero-Erausquin in [**21**]. It is based on the geometric Brascamp-Lieb inequality due to Ball [**4**], see also [**5**], in the rank one case, and due to Barthe [**7**] in the general case.

THEOREM 7.9 (Correlation inequality). *For $i = 1,\ldots K$, let $H_i \subset \mathbb{R}^M$ be subspaces of dimension d_i and $B_i : \mathbb{R}^M \to H_i$ be linear maps with the property that $B_iB_i^T = I_{H_i}$, the identity map on H_i. Assume further that there are non-negative constants $c_i, i = 1,\ldots,K$ such that*

$$\sum_{i=1}^K c_i B_i^T B_i = I_M\ . \tag{7.19}$$

Then, for nonnegative functions $f_i : H_i \to \mathbb{R}$,

$$\int_{\mathbb{R}^M} \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v} \leq \prod_{i=1}^K\left(\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\,\mathrm{d}u\right)^{c_i}\ . \tag{7.20}$$

Moreover,

$$\int_{\mathbb{R}^M} h(\mathbf{v})\log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}$$
$$\geq \sum_{i=1}^K c_i\left[\int_{\mathbb{R}^M} h(\mathbf{v})\log f_i(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v} - \log\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\,\mathrm{d}u\right], \tag{7.21}$$

for any nonnegative function $h \in L^1(\mathbb{R}^M, e^{-\pi|\mathbf{v}|^2}d\mathbf{v})$.

Since Theorem 7.9 is very useful in a number of applications, and for the readers convenience, we will give an elementary proof in Appendix 7.8.1.

Remark. By taking the trace in (7.19) one sees that

$$\sum_{i=1}^{K} c_i d_i = M \ .$$

We would like to apply (7.21) to the right hand side of (7.17). An immediate problem is that (7.17) is in terms of integrals and not sums. While there are some results available for continuous indices (see, e.g., [**8**]), they do not apply to our situation and hence we will take a more direct approach and approximate the measure $\rho(\theta)\mathrm{d}\theta$ by a discrete measure. It is important that the approximation also satisfies the constraint (7.5). The following lemma establishes such an approximation. Its proof is given in Appendix 7.8.2.

Lemma 7.10. *Let ρ be a probability density on $[-\pi,\pi]$ whose Fourier series converges absolutely and assume that* (7.5) *is satisfied. There exists a sequence of discrete probability measures ν_K, $K = 1, 2, \ldots$, such that for every continuous function f on $[-\pi,\pi]$*

$$\lim_{K\to\infty} \int_{-\pi}^{\pi} f(\theta)\, \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta)\rho(\theta)\, \mathrm{d}\theta \ .$$

Moreover,

$$\int_{-\pi}^{\pi} \cos\theta \sin\theta \, \nu_K(\mathrm{d}\theta) = 0 \ ,$$

for all $K \in \mathbb{N}$. More precisely,

$$\nu_K(\mathrm{d}\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) \delta\left(\theta - \frac{2\pi\ell}{4K+1}\right) \mathrm{d}\theta \ ,$$

where

$$\rho_K(\theta) = \int_{-\pi}^{\pi} \rho(\theta-\phi)\ p_K(\theta)\, \mathrm{d}\phi \quad \textit{and} \quad p_K(\theta) := \frac{1}{2K+1}\left(\sum_{k=-K}^{K} e^{ik\theta}\right)^2 \ .$$

At this point we can prepare the ground for the application of Theorem 7.9 to inequality (7.17). We first replace $\rho(\theta)\mathrm{d}\theta$ in (7.17) with $\nu_K(\mathrm{d}\theta)$. Setting

$$\omega_{\ell_j} = \rho_K(\theta_j) \ , \ \theta_{\ell_j} = \frac{2\pi\ell_j}{4K+1} \ , \text{ and } \underline{\theta} = (\theta_{\ell_1}, \ldots, \theta_{\ell_k}) \ ,$$

we obtain

$$\sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu_K(\mathrm{d}\theta_1)\cdots\nu_K(\mathrm{d}\theta_k) \sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2$$
$$\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}$$
$$= \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \sum_{-K\le\ell_1,\dots,\ell_k\le K} \prod_{j=1}^{k}\omega_{\ell_j} \sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2$$
$$\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\ . \tag{7.22}$$

In order to apply Theorem 7.9 to (7.22) we have to replace the sum over the index i with a sum over the indices $\alpha_1,\dots,\alpha_k,\ell_1,\dots\ell_k$ and all subsets $\sigma\subset\{1,\dots,M\}$. Moreover, we substitute

the constants c_i	by $\frac{1}{C_{k,M}}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\prod_{j=1}^{k}\omega_{\ell_j}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2)$,
the functions $f_i(\mathbf{w})$	by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{w})$,
the linear maps B_i	by $P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T$,
the functions $f_i(B_i\mathbf{v})$	by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})$,
and the subspaces H_i	by $\mathbb{R}^{\lvert\sigma^c\rvert}$.

Note that, for any given index i the condition $B_iB_i^T = I_{H_i}$ corresponds to the fact that $P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^TU_k(\underline{\alpha},\underline{\theta})P_{\sigma^{\mathrm{c}}} = P_{\sigma^{\mathrm{c}}}$ which is the identity on $\mathbb{R}^{|\sigma^{\mathrm{c}}|}$.

The next theorem establishes the sum rule (7.19) in our setting and hence ensures the applicability of Theorem 7.9 to (7.22).

THEOREM 7.11 (The sum rule). *If $\nu(\mathrm{d}\theta)$ is a probability measure satisfying* (7.5), *then*

$$\sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\,\cdots\nu(\mathrm{d}\theta_k)\times$$
$$\times \sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) U_k(\underline{\alpha},\underline{\theta}) P_{\sigma^{\mathrm{c}}}^T P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta})^T = C_{k,M} I_M\ , \tag{7.23}$$

where

$$C_{k,M} = \left[\frac{M}{N+M} + \frac{N}{N+M}\left(1-\mu_\nu\frac{N+M}{N\Lambda}\right)^k\right]$$

with

$$\mu_\nu = \mu \int \nu(\mathrm{d}\theta)\sin^2\theta\ .$$

The proof will be given in Section 7.6. We observe here that it follows from Theorem 7.10 that $\mu_\rho = \lim_{K\to\infty}\mu_{\nu_K}$.

PROOF OF THEOREM 7.8 . First we consider the case where ρ is repaced by ν_K and use Theorem 7.9 together with Theorem 7.11 and the identification rules described above. The entropy inequality (7.21) now says that

$$\begin{aligned}
&\int_{\mathbb{R}^M} h_0(\mathbf{v})\log h_0(\mathbf{v})e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\\
&\geq \frac{1}{C_{k,M}}\sum_{\alpha_1,\dots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\sum_{-K\leq\ell_1,\dots,\ell_k\leq K}\prod_{j=1}^{k}\omega_{\ell_j}\sum_{\sigma\subset\{1,\dots,M\}}\prod_{i\in\sigma^c}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\\
&\quad\times\prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right)\left[\int_{\mathbb{R}^M} h_0(\mathbf{v})\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\right.\\
&\qquad\qquad\left.-\log\int_{\mathbb{R}^{|\sigma^c|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(u)\,e^{-\pi|u|^2}\,\mathrm{d}u\right]\ .
\end{aligned}$$

However, since h_0 is normalized and $U_k(\underline{\alpha},\underline{\theta})$ is orthogonal, we find that

$$\begin{aligned}
\int_{\mathbb{R}^{|\sigma^c|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(u)\,e^{-\pi|u|^2}\,\mathrm{d}\ u &= \int_{\mathbb{R}^{|\sigma^c|}}\int_{\mathbb{R}^{|\sigma|}} h_{0,U_k(\underline{\alpha},\underline{\theta})}(v,u)\,e^{-\pi|v|^2}\,\mathrm{d}v\,e^{-\pi|u|^2}\,\mathrm{d}u\\
&= \int_{\mathbb{R}^M} h_0(U_k(\underline{\alpha},\underline{\theta})\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\\
&= 1\ .
\end{aligned}$$

Thus we find that

$$\begin{aligned}
&\sum_{\alpha_1,\dots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\sum_{-K\leq\ell_1,\dots,\ell_k\leq K}\prod_{j=1}^{k}\omega_{\ell_j}\sum_{\sigma\subset\{1,\dots,M\}}\prod_{i\in\sigma^c}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right)\\
&\quad\times\int_{\mathbb{R}^M} h_0(\mathbf{v})\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\ \leq\ C_{k,M}\mathcal{S}(h_0)\ .
\end{aligned}\tag{7.24}$$

As $K\to\infty$, the left-hand side of (7.24) converges to the right-hand side of (7.17). □

We now have all ingredients to give the proof of Theorem 7.1.

PROOF OF THEOREM 7.1. Recall from Section 7.3, that

$$f(\mathbf{v},t) = e^{-\pi|\mathbf{v}|^2}e^{-\Lambda t}\sum_{k=0}^{\infty}\frac{t^k\Lambda^k}{k!}h_k(\mathbf{v})\ ,$$

and that $S(f(\cdot,t)) = \mathcal{S}(h(\cdot,t))$. Combining Theorem 7.5 and Theorem 7.8, we obtain

$$\mathcal{S}(h_k)\leq C_{k,M}\mathcal{S}(h_0)\ ,$$

and computing

$$e^{-\Lambda t}\sum_{k=0}^{\infty}\frac{\Lambda^k t^k}{k!}C_{k,M}$$

yields Theorem 7.1. □

7.6. The sum rule. Proof of Theorem 7.11

We have to compute the matrix

$$Z := \sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\times$$
$$\times\sum_{\sigma\subset\{1,\ldots,M\}}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right)U_k(\underline{\alpha},\underline{\theta})P_{\sigma^{\mathrm{c}}}^T P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T\ .$$

Obviously $P_{\sigma^{\mathrm{c}}}^T P_{\sigma^{\mathrm{c}}} = P_{\sigma^{\mathrm{c}}}$ and hence

$$Z = \sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\times$$
$$\times\, U_k(\underline{\alpha},\underline{\theta})\left[\sum_{\sigma\subset\{1,\ldots,M\}}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right)P_{\sigma^{\mathrm{c}}}\right]U_k(\underline{\alpha},\underline{\theta})^T\ .$$

The sum on σ is easily evaluated and yields the matrix $\Gamma_k^2(\underline{\alpha},\underline{\theta})$. Hence, recalling the orthogonal singular value decomposition (7.13) of $A_k(\underline{\alpha},\underline{\theta})$, that is, $A_k(\underline{\alpha},\underline{\theta}) = U_k(\underline{\alpha},\underline{\theta})\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta})$, we find that

$$Z = \sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\,A_k(\underline{\alpha},\underline{\theta})A_k^T(\underline{\alpha},\underline{\theta})\ . \tag{7.25}$$

One can think about this expression in the following fashion. Recall that

$$\left[\prod_{l=1}^{k}r_{\alpha_l}(\theta_l)\right]^{-1} = \begin{pmatrix}A_k(\underline{\alpha},\underline{\theta}) & B_k(\underline{\alpha},\underline{\theta})\\ C_k(\underline{\alpha},\underline{\theta}) & D_k(\underline{\alpha},\underline{\theta})\end{pmatrix}\ .$$

With this notation, the matrix Z equals the top left entry of the matrix

$$\sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\left[\prod_{l=1}^{k}r_{\alpha_l}(\theta_l)\right]^{-1}\begin{pmatrix}I_M & 0\\ 0 & 0\end{pmatrix}\left[\prod_{l=1}^{k}r_{\alpha_l}(\theta_l)\right]\ .$$

The computation hinges on a repeated application of the elementary identity

$$\int_{-\pi}^{\pi}\nu(\mathrm{d}\theta)\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}m_1 & 0\\ 0 & m_2\end{pmatrix}\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix}$$
$$= \begin{pmatrix}(1-\widetilde{\nu})m_1+\widetilde{\nu}m_2 & 0\\ 0 & (1-\widetilde{\nu})m_2+\widetilde{\nu}m_1\end{pmatrix}\ ,$$

where $\widetilde{\nu} = \int \nu(\mathrm{d}\theta)\, \sin^2(\theta)$. For this to be true we just need (7.5). We easily check that for the rotations $r_\alpha(\theta)$

$$\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} r_\alpha(\theta)$$
$$= \frac{1}{\Lambda}\left(\frac{M\lambda_S}{2} + \frac{N\lambda_R}{2}\right) \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} + \frac{\mu}{\Lambda N}$$
$$\times \begin{pmatrix} N(M-1) + N((1-\widetilde{\nu})m_1 + \widetilde{\nu} m_2) I_M & 0 \\ 0 & (N-1)M + M(\widetilde{\nu} m_1 + (1-\widetilde{\nu}) m_2) I_N \end{pmatrix}$$
$$= \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} + \frac{\mu_\nu}{\Lambda N} \begin{pmatrix} N(m_2 - m_1) I_M & 0 \\ 0 & M(m_1 - m_2) I_N \end{pmatrix} . \tag{7.26}$$

where $\mu_\nu = \widetilde{\nu}\mu$. Denote by $L(\nu_1, \nu_2)$ the $(N+M)\times(N+M)$ matrix

$$L(m_1, m_2) = \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} ,$$

and set

$$\mathcal{P} = I_2 - \frac{\mu_\nu}{\Lambda N} \begin{pmatrix} N & -N \\ -M & M \end{pmatrix} .$$

Then (7.26) is recast as

$$\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} L(m_1, m_2) r_\alpha(\theta) = L(m_1', m_2') , \tag{7.27}$$

where

$$\begin{pmatrix} m_1' \\ m_2' \end{pmatrix} = \mathcal{P} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} .$$

By a repeated application of (7.27) we obtain

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right]^T L(\underline{m}) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right]$$
$$= L(\mathcal{P}^k \underline{m}) .$$

Thus,

$$Z = \left(\mathcal{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 I_M .$$

It is easy to see that $\mathcal{P}$ has eigenvalues $\ell_1 = 1$ and $\ell_2 = 1 - \mu_\nu(M+N)/(\Lambda N)$ with eigenvectors $\underline{m}_1 = (1,1)$ and $\underline{m}_2 = (N, -M)^T/(M+N)$. Consequently,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{M}{N+M} \underline{m}_1 + \underline{m}_2 ,$$

which yields

$$\left(\mathcal{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 = \frac{M}{N+M} + \frac{N}{M+N}\left(1 - \mu_\nu \frac{M+N}{\Lambda N}\right)^k ,$$

and hence completes the proof of Theorem 7.11. □

7.7. Boltzmann-Kac collisions

In this section we show that the above results can also be extended, at least in a particular case, to three-dimensional Boltzmann-Kac collisions.

Again we consider a system of M particles coupled to a reservoir consisting of N particles, but now with velocities $v_1, \ldots, v_M, w_1, \ldots, w_N \in \mathbb{R}^3$. The collisions between a pair of particles have to conserve energy and momentum,

$$\begin{aligned} z_i^2 + z_j^2 &= (z_i^*)^2 + (z_j^*)^2 \\ z_i + z_j &= z_i^* + z_j^* \ , \end{aligned}$$

where z can be either the velocity of a system particle v or of a reservoir particle w. A convenient parametrization of the post-collisional velocities in terms of the velocities before the collision is given by

$$\begin{aligned} z_i^*(\omega) &= z_i - \omega \cdot (z_i - z_j)\, \omega \\ z_j^*(\omega) &= z_j + \omega \cdot (z_i - z_j)\, \omega, \quad \text{where } \omega \in \mathbb{S}^2 \ . \end{aligned}$$

This is the so-called ω-representation. This representation is particularly useful, because the velocities are related to each other by a *linear* transformation, and the strategy used to proof the results for the one-dimensional Kac system carries over rather directly. The direction ω will be chosen according to the uniform probability distribution on the unit sphere $\mathbb{S}^2$.

Introduce the operators

$$(R_{ij} f)(\boldsymbol{z}) = \int_{\mathbb{S}^2} f(r_{ij}(\omega)^{-1} \boldsymbol{z})\, \mathrm{d}\omega \ ,$$

where $\mathrm{d}\omega$ denotes the uniform probability measure on the sphere and the matrices $r_{ij}(\omega)$ are symmetric involutions acting as

$$\begin{pmatrix} z_i^* \\ z_j^* \end{pmatrix} = \begin{pmatrix} I - \omega\omega^T & \omega\omega^T \\ \omega\omega^T & I - \omega\omega^T \end{pmatrix} \begin{pmatrix} z_i \\ z_j \end{pmatrix}$$

on the velocities of the particles i and j, and as identities otherwise. They will replace the one-dimensional Kac collision operators in (7.3) in the otherwise unchanged generator of the time evolution. Notice that the matrices $r_{ij}(\omega)$ are orthogonal, so that the expansion formula (7.10) still holds with the obvious changes in the dimension of the single-particle spaces.

We prove an analog of Theorem 7.1 for the case of three-dimensional Boltzmann-Kac collisions and pseudo-Maxwellian molecules.

THEOREM 7.12. *Let $N \geq M$ and $F_0(\boldsymbol{v}, \boldsymbol{w}) = f_0(\boldsymbol{v})\, \mathrm{e}^{-\pi|\boldsymbol{w}|^2}$ for some probability distribution f_0 on $\mathbb{R}^{3M}$. Then the entropy of the marginal*

$$f(\boldsymbol{v}, t) := \int_{\mathbb{R}^{3N}} \left(\mathrm{e}^{\mathcal{L}t} F_0\right)(\boldsymbol{v}, \boldsymbol{w})\, \mathrm{d}\boldsymbol{w}$$

with respect to the thermal state $\mathrm{e}^{-\pi|\boldsymbol{v}|^2}$ *is bounded by*

$$S(f(\cdot,t)) \leq \left[\frac{N}{N+M} + \frac{N}{N+M}\mathrm{e}^{-\frac{\mu}{3}\frac{N+M}{N}t}\right] S(f_0)\ .$$

REMARK. The result in three dimensions is very similar to the case of one-dimensional Kac collisions, with the difference that the rate of exponential decay is $\mu/3$ instead of μ_ρ. The factor 1/3 comes from the fact that $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = I_3/3$. It would be interesting to cover the true Maxwellian molecules interaction

$$(R_{ij}f)(z) = \int_{\mathbb{S}^2} b\left(\frac{v_i - v_j}{|v_i - v_j|}\cdot\omega\right) f(r_{ij}(\omega)^{-1}z)\,\mathrm{d}\omega\ .$$

However, the dependence of the scattering rate b on the velocities doesn't seem to be treatable with the above methods.

The proof of Theorem 7.12 essentially deviates from the one-dimensional case in only two places: the sum rule and the discrete approximation of the integrals. We begin by proving an analogue of Theorem 7.11. Most of the steps for the computation of the matrix Z in (7.25) are the same. What remains is to compute

$$Z := \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{\mathbb{S}^2\times\cdots\times\mathbb{S}^2} \mathrm{d}\omega_1\cdots\mathrm{d}\omega_k\, A_k(\underline{\alpha},\underline{\omega})A_k(\underline{\alpha},\underline{\omega})^T\ ,$$

which is somewhat different for the case of Boltzmann-Kac collisions. Recall that $A_k(\underline{\alpha},\underline{\omega})$ is the upper left $3M\times 3M$ block of $[\prod_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1}$, i.e.,

$$A_k(\underline{\alpha},\underline{\omega}) = P_{3M}[\Pi_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1}P_{3M}^T$$

with the projection $P_{3M} = \begin{pmatrix} I_{3M} & 0\end{pmatrix}$ from $\mathbb{R}^{3M+3N}\to\mathbb{R}^{3M}$. In particular, by linearity,

$$Z = P_{3M}\left(\sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{(\mathbb{S}^2)^k}\mathrm{d}\underline{\omega}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]^{-1}\begin{pmatrix} I_{3M} & 0\\ 0 & 0\end{pmatrix}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]\right)P_{3M}^T.$$

As in the proof of Theorem 7.11 we have

LEMMA 7.13. *Let* $\alpha,\beta\geq 0$. *Then*

$$\sum_{1\leq i<j\leq M+N}\lambda_{ij}\int_{\mathbb{S}^2}\mathrm{d}\omega\, r_{ij}(\omega)^{-1}\begin{pmatrix}\alpha I_{3M} & 0\\ 0 & \beta I_{3N}\end{pmatrix} r_{ij}(\omega) = \begin{pmatrix}\alpha' I_{3M} & 0\\ 0 & \beta' I_{3N}\end{pmatrix}\ ,$$

where α',β' *are related to* α,β *by*

$$\begin{pmatrix}\alpha'\\ \beta'\end{pmatrix} = \mathcal{P}\begin{pmatrix}\alpha\\ \beta\end{pmatrix},\quad \mathcal{P} = I_2 - \frac{\mu}{3\Lambda}\begin{pmatrix}1 & -1\\ -\frac{M}{N} & \frac{M}{N}\end{pmatrix}\ .$$

Notice that the matrix $\mathcal{P}$, which appears in Lemma 7.13, has eigenvalues 1 and $1-\mu/(3\Lambda)\,(1+M/N)$ with corresponding eigenvectors $\begin{pmatrix}1 & 1\end{pmatrix}^T$ and $\begin{pmatrix}-N/M & 1\end{pmatrix}^T$. Repeated application of Lemma 7.13 then implies, see also the argument in the one-dimensional case,

$$\sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{(\mathbb{S}^2)^k} \mathrm{d}\underline{\omega} \left[\prod_{j=1}^{k} r_{\alpha_j}(\omega_j)\right]^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} \left[\prod_{j=1}^{k} r_{\alpha_j}(\omega_j)\right] = \begin{pmatrix} \alpha^{(k)} I_{3M} & 0 \\ 0 & \beta^{(k)} I_{3N} \end{pmatrix},$$

where

$$\begin{pmatrix} \alpha^{(k)} \\ \beta^{(k)} \end{pmatrix} = \mathcal{P}^k \begin{pmatrix} \alpha \\ \beta \end{pmatrix} .$$

Before we prove Lemma 7.13, let us make an easy observation.

COROLLARY 7.14. *In the particular case $\alpha = 1$, $\beta = 0$, we get*

$$Z = \left[\frac{M}{M+N} + \frac{N}{M+N}\left(1 - \frac{\mu}{3\Lambda}\left(1 + \frac{M}{N}\right)\right)^k\right] I_{3M} .$$

PROOF OF LEMMA 7.13. For $1 \le i < j \le M$ (respectively for $M+1 \le i < j \le M+N$) the operators $r_{ij}(\omega)$ only act non-trivially in the first $3M$ (last $3N$) variables. Taking into account that $r_{ij}(\omega)^{-1}\, I\, r_{ij}(\omega) = I$, we obtain

$$\frac{\lambda_S}{M-1} \sum_{1 \le i < j \le M} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \frac{M\lambda_S}{2} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} ,$$

and

$$\frac{\lambda_R}{N-1} \sum_{M+1 \le i < j \le M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \frac{N\lambda_R}{2} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} .$$

It remains to look at the interaction terms $i = 1, \dots, M$ and $j = M+1, \dots, M+N$. Notice that

$$r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \left(\begin{array}{ccc|ccc} 0 & & & & & \\ & (\beta-\alpha)\omega\omega^T & & & 0 & \\ & & 0 & & & \\ \hline & & & 0 & & \\ & 0 & & & (\beta-\alpha)\omega\omega^T & \\ & & & & & 0 \end{array}\right),$$

where the non-zero entries in the second summand on the right-hand side correspond to the i^{th}, respectively j^{th}, 3×3 block on the diagonal. Since $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3\, I_3$, we obtain

$$\frac{\mu}{N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \mu M \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \frac{\mu}{3}(\alpha - \beta) \begin{pmatrix} -I_{3M} & 0 \\ 0 & \frac{M}{N} I_{3N} \end{pmatrix} .$$

Recall the definition of $\Lambda = M\lambda_S/2 + N\lambda_R/2 + \mu M$. Hence summation of all the three contributions yields the statement of the Lemma. $\square$

As in the one-dimensional case, in order to apply the geometric Brascamp-Lieb inequality Theorem 7.9, we need to approximate the uniform probability measure $\mathrm{d}\omega$ on the sphere by a suitable sequence of discrete measures as in the one-dimensional case (see Lemma 7.10). Additionally, in each step of the discretization, the constraint $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3I$, has to hold. This is important, because it guarantees that the geometric Brascamp-Lieb condition, i.e., the sum rule (7.19), holds in each step.

In order to find such an approximation, we parametrize the sphere in the usual way by spherical coordinates

$$\omega = \omega(\theta, \varphi) = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$$

for $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$. For $K, L \in \mathbb{N}$ we introduce the measures

$$\Phi_K := \frac{\pi}{K}\sum_{j=0}^{2K-1} \delta_{\frac{\pi}{K}j} \quad \text{on } [0, 2\pi], \quad \text{and}$$

$$\Theta_L := \sum_{i=1}^{L} \frac{2}{(1-u_i^2)^{3/2}(P_L'(u_i))^2}\delta_{\arccos u_i} \quad \text{on } [0, \pi],$$

where P_L is the Legendre polynomial of order L on $[-1, 1]$, and u_i, $i = 1, \ldots, L$, are its zeros. Then, if $f \in \mathcal{C}[0, 2\pi]$ and $g \in \mathcal{C}[-1, 1]$,

$$\int_0^{2\pi} f(\varphi)\, \Phi_k(\mathrm{d}\varphi) = \frac{\pi}{K}\sum_{j=0}^{2K-1} f\left(\frac{\pi}{K}j\right) \to \int_0^{2\pi} f(\varphi)\, \mathrm{d}\varphi$$

as $K \to \infty$ as Riemann sum. Furthermore,

$$\int_0^{\pi} g(\cos\theta)\, \sin\theta\, \Theta_L(\mathrm{d}\theta) = \sum_{i=1}^{L} \frac{2}{(1-u_i^2)(P_L'(u_i))^2} g(u_i)$$
$$\to \int_{-1}^{1} g(u)\, \mathrm{d}u = \int_0^{\pi} g(\cos\theta)\, \sin\theta\, \mathrm{d}\theta$$

as $L \to \infty$ by Gauss-Legendre quadrature . The latter approximation is exact for polynomials of order less or equal to $2L - 1$. In particular, we have

$$\int_0^{\pi} \cos^2\theta\, \sin\theta\, \Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \cos^2\theta\, \sin\theta\, \mathrm{d}\theta = \frac{2}{3}\ , \ \text{and}$$
$$\int_0^{\pi} \sin^3\theta\, \Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \sin^3\theta\, \mathrm{d}\theta = \frac{4}{3},$$

for all $L \geq 2$. It is easy to check that

$$\int_0^{2\pi} \sin\varphi \cos\varphi \, \Phi_k(\mathrm{d}\varphi) = 0,$$

$$\int_0^{2\pi} \sin\varphi \, \Phi_k(\mathrm{d}\varphi) = \int_0^{2\pi} \cos\varphi \Phi_k(\mathrm{d}\varphi) = 0,$$

$$\int_0^{2\pi} \sin^2\varphi \Phi_k(\mathrm{d}\varphi) = \int_0^{2\pi} \cos^2\varphi \, \Phi_k(\mathrm{d}\varphi) = \pi,$$

for all $K \geq 2$. Consequently,

$$\frac{1}{4\pi}\int_0^{2\pi} \omega(\theta,\varphi)\omega(\theta,\varphi)^T \, \Theta_L(\mathrm{d}\theta)\Phi_k(\mathrm{d}\varphi)$$

$$= \frac{1}{2K}\sum_{j=0}^{2K-1}\sum_{i=0}^{L} \frac{\omega\left(\arccos u_i, \pi j/K\right)\omega\left(\arccos u_i, \pi j/K\right)^T}{(1-u_i^2)(P_L'(u_i))^2} = \frac{1}{3}I_3$$

for all $K, L \geq 2$. It follows that Z is not changed by replacing the uniform measure on $\mathbb{S}^2$ by the above discrete approximation, in particular, Z is still proportional to the identity matrix, which guarantees the applicability of the geometric Brascamp-Lieb inequality.

This concludes the proof of Theorem 7.12. □

7.8. Appendix

7.8.1. The Geometric Brascamp-Lieb inequality and the entropy inequality. In this section we prove Theorem 7.9. We use the same strategy as in [**22**] and [**9**] which consists of transporting the functions f_i with the heat kernel in such a way that the right-hand side of (7.20) remains fixed while the left-hand side of that inequality increases. The results in [**9**] are quite general but for the special case in which the sum rule (7.19) holds, the proof is quite simple and this is one of the reasons why we include it here.

Proof of Theorem 7.9. The inequality (7.20) is equivalent to

$$\int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i^{c_i}(B_i\mathbf{v})\,\mathrm{d}\mathbf{v} \leq \prod_{i=1}^{K}\left(\int_{H_i} f_i(u)\,\mathrm{d}u\right)^{c_i} . \tag{7.28}$$

This follows from the identity

$$\prod_{i=1}^{K}\left(e^{-\pi|B_i\mathbf{v}|^2}\right)^{c_i} = \exp\left(-\pi\sum_{i=1}^{K}(\mathbf{v}\ c_iB_i^TB_i\ \mathbf{v})\right) = e^{-\pi|\mathbf{v}|^2} .$$

We transport the functions f_i by the heat flow, that is we define

$$f_i(B_i\mathbf{v},t) := \frac{1}{(4\pi t)^{M/2}}\int_{\mathbb{R}^M} e^{-|\mathbf{v}-\mathbf{w}|^2/(4t)} f_i(B_i\mathbf{w})\mathrm{d}\mathbf{w} . \tag{7.29}$$

For the above definition to make sense, we have to show that the right-hand side is a function of $B_i\mathbf{v}$ alone. The condition $B_iB_i^T = I_{H_i}$ means that the matrix $P_i = B_i^TB_i$ is an orthogonal projection onto a d_i dimensional subspace of $\mathbb{R}^M$. Moreover,

$B_i P_i = I_{H_i} B_i = B_i$. We rewrite the integral (7.29) by splitting it in an integral over $\mathbf{w}' \in \operatorname{Ran} P_i$ and one over integration over $\mathbf{w}'' \in \operatorname{Ran} P_i^\perp$. Carrying out the integration over $\mathbf{w}''$ we obtain

$$
\begin{aligned}
& f_i(B_i\mathbf{v}, t) \\
&= \frac{1}{(4\pi t)^{M/2}} \int_{\operatorname{Ran} P_i} \int_{\operatorname{Ran} P_i^\perp} e^{-|(P_i\mathbf{v} - P_i\mathbf{w}')|^2/(4t)} e^{-|(P_i^\perp \mathbf{v} - \mathbf{w}'')|^2/(4t)} f_i(B_i P_i \mathbf{w}) \, \mathrm{d}\mathbf{w}' \mathrm{d}\mathbf{w}'' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{\operatorname{Ran} P_i} e^{-|(P_i\mathbf{v} - P_i\mathbf{w}')|^2/(4t)} f_i(B_i P_i \mathbf{w}') \, \mathrm{d}\mathbf{w}' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{\operatorname{Ran} P_i} e^{-|(B_i\mathbf{v} - B_i\mathbf{w}')|^2/(4t)} f_i(B_i \mathbf{w}') \, \mathrm{d}\mathbf{w}' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{H_i} e^{-|(B_i\mathbf{v} - u)|^2/(4t)} f_i(u) \, \mathrm{d}u \ .
\end{aligned}
$$

where, in the last equality, we have used that B_i maps the range of P_i isometrically onto H_i. This justifies (7.29). Moreover, the above computation also shows that

$$
\int_{H_i} f_i(u,t) \mathrm{d}u = \int_{H_i} f_i(u) \mathrm{d}u
$$

so that the right-hand side of the inequality (7.28) does not change under the heat flow.

We now show that the left-hand side of (7.28) is an increasing function of t. It is convenient to set $\phi_i(u,t) = \log f_i(u,t)$. Differentiating the function $\phi_i(B_i\mathbf{v}, t)$ with respect to t yields

$$
\frac{\mathrm{d}}{\mathrm{d}t} \phi_i(B_i\mathbf{v}, t) = \Delta_{\mathbf{v}} \phi_i(B_i\mathbf{v}, t) + |\nabla_{\mathbf{v}} \phi_i(B_i\mathbf{v}, t)|^2 \ .
$$

Moreover,

$$
\begin{aligned}
& \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^M} \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v}, t) \, \mathrm{d}\mathbf{v} \\
& \qquad = \sum_{m=1}^K c_m \int_{\mathbb{R}^M} [\Delta_{\mathbf{v}} \phi_m(B_m\mathbf{v}, t) + |\nabla_{\mathbf{v}} \phi_m(B_m\mathbf{v}, t)|^2] \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v}, t) \, \mathrm{d}\mathbf{v} \ .
\end{aligned}
$$

Integrating by parts the term containing the Laplacian yields

$$
\begin{aligned}
& \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^M} \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v}, t) \, \mathrm{d}\mathbf{v} \\
& \quad = \sum_{m=1}^K c_m \int_{\mathbb{R}^M} |\nabla_{\mathbf{v}} \phi_m(B_m\mathbf{v}, t)|^2 \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v}, t) \mathrm{d}\mathbf{v} \\
& \qquad - \sum_{m,\ell=1}^K c_m c_\ell \int_{\mathbb{R}^M} \nabla_{\mathbf{v}} \phi_m(B_m\mathbf{v}, t) \cdot \nabla_{\mathbf{v}} \phi_\ell(B_\ell\mathbf{v}, t) \prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v}, t) \mathrm{d}\mathbf{v} \ .
\end{aligned}
$$

Finally, using that

$$
\nabla_{\mathbf{v}} \phi_m(B_m\mathbf{v}, t) = B_i^T (\nabla \phi_m)(B_m\mathbf{v})
$$

we get

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^M}\prod_{i=1}^{K} f_i^{c_i}(B_i\mathbf{v},t)\,\mathrm{d}\mathbf{v}\\
&\quad=\sum_{m=1}^{K} c_m\int_{\mathbb{R}^M}|B_m^T(\nabla\phi_m)(B_m\mathbf{v},t)|^2\prod_{i=1}^{K} f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}\\
&\qquad-\sum_{m,\ell=1}^{K} c_m c_\ell\int_{\mathbb{R}^M} B_m^T(\nabla\phi_m)(B_m\mathbf{v},t)\cdot B_\ell^T(\nabla\phi_\ell)(B_\ell\mathbf{v},t)\prod_{i=1}^{K} f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}\ .
\end{aligned}
$$

We claim that this expression is non-negative. The vectors $\nabla\phi_m \in H_m$ are arbitrary and hence the problem is reduced to proving that for any set of vectors $V_m \in H_m$, $m = 1, \ldots, K$, it holds

$$
\sum_{m=1}^{K} c_m|B_m^T V_m|^2 - \sum_{m,\ell=1}^{K} c_m c_\ell B_m^T V_m\cdot B_\ell^T V_\ell \ \geq\ 0\ .
$$

Recalling that $B_m B_m^T = I_{H_m}$ and setting $Y = \sum_\ell c_\ell B_\ell^T V_\ell$ we conclude that it is enough to show that

$$
|Y|^2 \leq \sum_{m=1}^{K} c_m |V_m|^2\ .
$$

This follows easily, since, by applying Schwarz's inequality, we find that

$$
|Y|^2 = \sum_{\ell=1}^{K} c_\ell Y\cdot B_\ell^T V_\ell = \sum_{\ell=1}^{K} c_\ell B_\ell Y\cdot V_\ell \leq \left(\sum_{\ell=1}^{K} c_\ell |B_\ell Y|^2\right)^{1/2}\left(\sum_{\ell=1}^{K} c_\ell |V_\ell|^2\right)^{1/2}\ .
$$

Combining this with (7.19), we learn that

$$
|Y|^2 \leq \left(Y\cdot\sum_{\ell=1}^{K} c_\ell B_\ell^T B_\ell Y\right)^{1/2}\left(\sum_{\ell=1}^{K} c_\ell|V_\ell|^2\right)^{1/2} = |Y|\left(\sum_{\ell=1}^{K} c_\ell |V_\ell|^2\right)^{1/2}\ .
$$

Thus we have that, when applying (7.28) to the functions $f_i(u,t)$, the left hand side is an increasing function of t while the right hand side does not depends on t. It is thus enough to show that the inequality holds for large t. Using once more the sum-rule (7.19), we see that

$$
\begin{aligned}
&\int_{\mathbb{R}^M}\prod_{i=1}^{K}\frac{1}{(4\pi t)^{c_i d_i/2}}\left[\int_{H_i} e^{-|B_i\mathbf{v}-u|^2/(4t)} f_i(u)\,\mathrm{d}u\right]^{c_i}\mathrm{d}\mathbf{v}\\
&=\frac{1}{(4\pi)^{M/2}}\int_{\mathbb{R}^M}\prod_{i=1}^{K}\left[\int_{H_i} e^{-|B_i\mathbf{v}-t^{-1/2}u|^2/4} f_i(u)\,\mathrm{d}u\right]^{c_i}\mathrm{d}\mathbf{v}\\
&\qquad\stackrel{t\to\infty}{\longrightarrow}\frac{1}{(4\pi)^{M/2}}\int_{\mathbb{R}^M} e^{-|\mathbf{v}|^2/4}\prod_{i=1}^{K}\left[\int_{H_i} f_i(u)\,\mathrm{d}u\right]^{c_i}\mathrm{d}\mathbf{v} = \prod_{i=1}^{K}\left[\int_{H_i} f_i(u)\,\mathrm{d}u\right]^{c_i},
\end{aligned}
$$

which proves the first part of Theorem 7.9.

To prove the entropy inequality (7.21) we follow **[21]**. Let h be a non-negative function whose L^1 norm is one and whose entropy is finite. An elementary computation then shows that

$$\int_{\mathbb{R}^M} h(\mathbf{v}) \log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2} \mathrm{d}\mathbf{v}$$
$$= \sup_{\Phi} \left\{ \int_{\mathbb{R}^M} h(\mathbf{v})\Phi(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \int_{\mathbb{R}^M} e^{\Phi(\mathbf{v})} e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \right\} .$$

Now, we set

$$\Phi(\mathbf{v}) = \sum_{i=1}^{K} c_i \log f_i(B_i \mathbf{v}) \ .$$

This leads to the lower bound

$$\int_{\mathbb{R}^M} h(\mathbf{v}) \log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v}$$
$$\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(\mathbf{v}) \log f_i(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i(B_i\mathbf{v})^{c_i}\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v}$$
$$\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(\mathbf{v}) \log f_i(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \left[\prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\, \mathrm{d}u \right)^{c_i} \right],$$

where the second step is a consequence of the Brascamp-Lieb inequality (7.20). □

7.8.2. Proof of Lemma 7.10.

Proof. For K any positive integer we convolve $\rho(\theta)$ with the non-negative trigonometric polynomial

$$p_K(\theta) := \frac{1}{2K+1} \left(\sum_{k=-K}^{K} e^{ik\theta} \right)^2 = \sum_{m=-2K}^{2K} \left(1 - \frac{|m|}{2K+1} \right) e^{im\theta} \ ,$$

and obtain a probability density $\rho_K(\theta)$. The Fourier coefficients of $\rho_K(\theta)$ are given by

$$\widehat{\rho}_K(m) = \widehat{\rho}(m) \left(1 - \frac{|m|}{2K+1} \right)$$

for $|m| \leq 2K$ and are zero otherwise. In particular,

$$\widehat{\rho}_K(2) - \widehat{\rho}_K(-2) = 4i \int_{-\pi}^{\pi} \rho_K(\theta) \sin\theta \cos\theta\, \mathrm{d}\theta = 0 \ .$$

With ρ_K we construct the measure

$$\nu_K(d\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K \left(\frac{2\pi\ell}{4K+1} \right) \delta \left(\theta - \frac{2\pi\ell}{4K+1} \right) \mathrm{d}\theta \ .$$

The measure ν_K is positive since $\rho_K((2\pi\ell)/(4K+1)) \geq 0$. Moreover, for all $m \in \mathbb{Z}$ with $|m| \leq 2K$ the Fourier coefficients $\widehat{\nu}_K(m)$ and $\widehat{\rho}_K(m)$ coincide. In particular, we have

$$\int_{-\pi}^{\pi} \nu_K(\mathrm{d}\theta) \sin\theta \cos\theta = 0 \ .$$

To see this, we compute

$$\widehat{\nu}_K(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \nu_K(\theta) e^{-im\theta} \, \mathrm{d}\theta = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) e^{-2\pi i m\ell/(4K+1)}$$

for $|m| \leq 2K$. Observe that

$$\rho_K\left(\frac{2\pi\ell}{4K+1}\right) = \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) e^{2\pi i k\ell/(4K+1)} \ ,$$

and, as a consequence,

$$\widehat{\nu}_K(m) = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) e^{2\pi i \ell(k-m)/(4K+1)} \ .$$

But

$$\sum_{\ell=-2K}^{2K} e^{2\pi i \ell(k-m)/(4K+1)} = \begin{cases} 4K+1 & \text{if } k = m \\ 0 & \text{if } k \neq m, \end{cases}$$

and hence we conclude that

$$\widehat{\nu}_K(m) = \widehat{\rho}_K(m) \tag{7.30}$$

for $|m| \leq 2K$. It is easy to see that for any continuous function f on $[-\pi, \pi]$,

$$\lim_{K\to\infty} \int_{-\pi}^{\pi} f(\theta) \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta) \rho(\theta) \, \mathrm{d}\theta \ .$$

This finishes the proof.

$\square$

Bibliography

[1] S. Albeverio, F. Gesztesy, R. Hoegh-Krohn, and H. Holden, *Solvable models in quantum mechanics*, Texts and monographs in physics, Springer-Verlag, 1988.

[2] V. Bach, E. Lieb, J. Solovej, *Generalized Hartree-Fock theory and the Hubbard model*, J. Stat. Phys. **76**, 3–89, 1994.

[3] A. Baernstein, *A unified approach to symmetrization*, Partial differential equations of elliptic type: Cortona, 1992, 47–91. Sympos. Math. XXXV. Cambridge Univ. Press, 1994.

[4] K. Ball, *Volumes of sections of cubes and related problems*, Geometric aspects of functional analysis, 1987–88, vol. 1376 of Lecture Notes in Math., 251–260. Springer, Berlin, 1989.

[5] K. Ball, *Volume ratios and a reverse isoperimetric inequality*, J. London Math. Soc. **44**(2), 351–359, 1991.

[6] J. Bardeen, L. Cooper, J. Schrieffer, *Theory of superconductivity*, Phys. Rev. **108**, 1175–1204, 1957.

[7] F. Barthe, *On a reverse form of the Brascamp-Lieb inequality* Invent. Math. **134**(2), 335–361, 1998.

[8] F. Barthe, *A continuous version of the Brascamp-Lieb inequalities* Geometric aspects of functional analysis, vol. 1850 of Lecture Notes in Math., 53–63. Springer, Berlin, 2004.

[9] J. Bennett, A. Carbery, M. Christ, and T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, Geom. Funct. Anal. **17**(5),1343–1415, 2008.

[10] P. Billard and G. Fano, *An existence proof for the gap equation in the superconductivity theory*, Comm. Math. Phys. **10**, no. 4, 274–279, 1968.

[11] F. Bonetto, E. A. Carlen, R. Esposito, J. L. Lebowitz, and R. Marra, *Propagation of chaos for a thermostated kinetic model*, J. Stat. Phys. **154**(1-2), 265–285, 2014.

[12] F. Bonetto, N. Chernov, A. Korepanov, and J. L. Lebowitz *Nonequilibrium stationary state of a current-carrying thermostated system* EPL **102**(1), 2013.

[13] F. Bonetto, A. Geisinger, M. Loss, and T. Ried, *Entropy decay for the Kac evolution*, *preprint:* arXiv:1707.09584.

[14] F. Bonetto, M. Loss, H. Tossounian, and R. Vaidyanathan. *Uniform approximation of a Maxwellian thermostat by finite reservoirs*, Comm. Math. Phys. **351**(1), 311–339, 2017.

[15] F. Bonetto, M. Loss, and R. Vaidyanathan. *The Kac model coupled to a thermostat*, J. Stat. Phys. **156**(4), 647–667, 2014.

[16] G. Bräunlich, C. Hainzl, and R. Seiringer, *Translation-invariant quasi-free states for fermionic systems and the BCS approximation*, Rev. Math. Phys. **26** (2014), no. 7.

[17] G. Bräunlich, C. Hainzl, and R. Seiringer, *Bogolubov–Hartree–Fock Theory for Strongly Interacting Fermions in the Low Density Limit*, Math. Phys. Anal. Geom. **19**, no. 2, 2016.

[18] H. J. Brascamp and E. H. Lieb, *Best constants in Young's inequality, its converse, and its generalization to more than three functions*, Advances in Math. **20**(2), 151–173, 1976.

[19] E. A. Carlen, M. C. Carvalho, and M. Loss, *Many-body aspects of approach to equilibrium*, Journées "Équations aux Dérivées Partielles", La Chapelle sur Erdre, Exp. No. XI, **12**, Univ. Nantes, Nantes, 2000.

[20] E. A. Carlen, M. C. Carvalho, and M. Loss, *Determination of the spectral gap for Kac's master equation and related stochastic evolution*, Acta Math. **191**(1), 1–54, 2003.

[21] E. A. Carlen and D. Cordero-Erausquin, *Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities*, Geom. Funct. Anal. **19**(2), 373–405, 2009.

[22] E. A. Carlen, E. H. Lieb, and M. Loss, *A sharp analog of Young's inequality on S^N and related entropy inequalities*, J. Geom. Anal. **14**(3), 487–520, 2004.

[23] E. A. Carlen and M. Loss, *Extremals of functionals with competing symmetries*, J. Funct. Anal. **88**(2), 437–456, 1990.

[24] L. Cooper, *Bound electron pairs in a degenerate Fermi gas*, Phys. Rev. **104**, 1189–1190, 1956.

[25] A. Deuchert, *A lower bound for the BCS functional with boundary conditions at infinity*, J. Math. Phys. **58**, 2017.

[26] A. Deuchert, A. Geisinger, C. Hainzl, and M. Loss, *Persistence of translational symmetry in the BCS model with a radial pair interaction*, *accepted for publication in* Ann. Henri Poincaré, arXiv:1612.03303.

[27] P. Diaconis and L. Saloff-Coste, *Bounds for Kac's master equation* Comm. Math. Phys. **209**(3), 729–755, 2000.

[28] G. Eilenberger, *Ableitung verallgemeinerter Ginzburg–Landau-Gleichungen für reine Supraleiter aus einem Variationsprinzip*, Z. f. Physik **182**(4), 427–438, 1965.

[29] A. Einav, *On Villani's conjecture concerning entropy production for the Kac master equation*, Kinet. Relat. Models **4**(2), 479–497, 2011.

[30] P. Federbush, *Partially alternate derivation of a result of Nelson*, J. Math. Phys. **10**, 50–52, 1969.

[31] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems*, Dover, New York, 2003.

[32] R. P. Feynman, *An operator calculus having applications in quantum electrodynamics*, Phys. Rev. **84**, 1951.

[33] S. Fournais and B. Helffer, *Spectral Methods in Surface Superconductivity*, Birkhäuser, 2010.

[34] R. L. Frank, C. Hainzl, and E. Langmann, *The BCS critical temperature in a weak homogeneous magnetic field*, arXiv:1706.05686, 2017.

[35] R. L. Frank, C. Hainzl, S. Naboko, and R. Seiringer, *The critical temperature for the BCS equation at weak coupling*, J. Geom. Anal. **17**, no. 4, 559–567, 2007.

[36] R. L. Frank, C. Hainzl, B. Schlein, and R. Seiringer, *Incompatibility of Time-Dependent Bogoliubov–de-Gennes and Ginzburg–Landau Equations*, Lett. Math. Phys. **106**, no. 7, 913–923, 2016.

[37] R. L. Frank, C. Hainzl, R. Seiringer, and J. P. Solovej, *Microscopic derivation of Ginzburg-Landau theory*, J. Amer. Math. Soc. **25**, no. 3, 667–713, 2012.

[38] R. L. Frank, C. Hainzl, R. Seiringer, and J. P. Solovej, *Microscopic derivation of the Ginzburg-Landau model*, XVIIth International Congress on Mathematical Physics, 575–583, World Scientific, 2013.

[39] R. L. Frank, C. Hainzl, R. Seiringer, and J. P. Solovej, *Derivation of Ginzburg-Landau theory for a one-dimensional system with contact interaction*, Operator Methods in Mathematical Physics, 57–88, Oper. Theory Adv. Appl. **227**, Birkhäuser, Basel, 2013.

[40] R. L. Frank, C. Hainzl, R. Seiringer, and J. P. Solovej, *The external field dependence of the BCS critical temperature*, Comm. Math. Phys. **342**, no. 1, 189–216, 2016.

[41] R. L. Frank and M. Lemm, *Multi-component Ginzburg-Landau theory: Microscopic derivation and examples*, Annales Henri Poincaré, 1–56, 2016.

[42] A. Freiji, C. Hainzl, and R. Seiringer, *The gap equation for spin-polarized fermions*, J. Math. Phys. **53**, no. 1, 2012.

[43] J. Fröhlich and Z. Gang, *Exponential convergence to the Maxwell distribution for some class of Boltzmann equations*, Comm. Math. Phys. **314**(2), 525–554, 2012.

[44] E. Gabetta, G. Toscani, and B. Wennberg, *Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation*, J. Stat. Phys. **81**, 901–934, 1995.

[45] P. G. de Gennes, *Superconductivity of Metals and Alloys*, Westview Press, 1966.

[46] V. L. Ginzburg and L. D. Landau, *On the Theory of superconductivity* J. Exptl. Theoret. Phys. (U.S.S.R) **20**, 1064–1082, 1950.

[47] L. P. Gorkov, *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.

[48] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97**(4), 1061–1083, 1975.

[49] L. Gross,*Logarithmic Sobolev inequalities and contractivity properties of semigroups*, Lecture Notes in Math. **1563**, 54–88. Springer, Berlin, 1993.

[50] C. Hainzl, E. Hamza, R. Seiringer, and J. P. Solovej, *The BCS functional for general pair interactions*, Comm. Math. Phys. **281**, no. 2, 349–367, 2008.

[51] C. Hainzl, M. Lewin, and R. Seiringer, *A nonlinear model for relativistic electrons at positive temperature*, Rev. Math. Phys. **20**, no. 10, 1283–1307, 2008.

[52] C. Hainzl and B. Schlein, *Dynamics of Bose-Einstein condensates of fermion pairs in the low density limit of BCS theory*, J. Funct. Anal. **265**, no. 3, 399–423, 2013.

[53] C. Hainzl and R. Seiringer, *The BCS critical temperature for potentials with negative scattering length*, Lett. Math. Phys. **84**, no. 2-3, 99–107, 2008.

[54] C. Hainzl and R. Seiringer, *Critical temperature and energy gap for the bcs equation*, Phys. Rev. B **77**, 2008.

[55] C. Hainzl and R. Seiringer, *Spectral properties of the BCS gap equation of superfluidity*, Mathematical results in quantum mechanics, World Sci. Publ., Hackensack, NJ, 117–136, 2008.

[56] C. Hainzl and R. Seiringer, *Low density limit of BCS theory and Bose-Einstein condensation of fermion pairs*, Lett. Math. Phys. **100**, no. 2, 119–138, 2012.

[57] C. Hainzl and R. Seiringer, *The Bardeen-Cooper-Schrieffer functional of superconductivity and its mathematical properties*, J. Math. Phys. **57**, no. 2, 2016.

[58] C. Hainzl and J. Seyrich, *Comparing the full time-dependent Bogoliubov-de-Gennes equations to their linear approximation: a numerical investigation*, Eur. Phys. J. B **89**, no. 5, 2016.

[59] T. S. Han, *Nonnegative entropy measures of multivariate symmetric correlations*, Information and Control **36**(2), 133–156, 1978.

[60] E. Janvresse, *Spectral gap for Kac's model of Boltzmann equation*, Ann. Probab. **29**(1), 288–304, 2001.

[61] M, Kac, *Foundations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, 171–197, University of California Press, 1956.

[62] H. P. McKean Jr., *Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas*, Arch. Rational Mech. Anal. **32**,343–367, 1966.

[63] L. Kelvin, *XXIX. Aepinus atomized*, Philosophical Magazine Series **6** 3.15, 257–283, 1902.

[64] W. Kohn and J. M. Luttinger, *New mechanism for superconductivity*, Phys. Rev. Lett. **15**, 524–526, 1965.

[65] B. J. McLeod and Y. Yang, *The uniqueness and approximation of a positive solution of the Bardeen-Cooper-Schrieffer gap equation*, J. Math. Phys. **41**, no. 9, 6007–6025, 2000.

[66] O. E. Lanford III, *Time evolution of large classical systems*, Dynamical systems, theory and applications, Rencontres, Battelle Res. Inst., Seattle, Wash., 1974.

[67] O. E. Lanford III, *On a derivation of the Boltzmann equation*, International Conference on Dynamical Systems in Mathematical Physics, Rennes, 1975.

[68] A. J. Leggett, *Diatomic molecules and Cooper pairs*, Modern trends in the theory of condensed matter, J. Phys. Colloq. C7-19, 1980.
[69] A. J. Leggett, *Quantum Liquids*, Oxford, 2006.
[70] E. H. Lieb, *Gaussian kernels have only Gaussian maximizers*, Invent. Math. **102**(1), 179–208, 1990.
[71] L. H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. **55**, 961–962, 1949.
[72] D. K. Maslen, *The eigenvalues of Kac's master equation*, Math. Z. **243**(2), 291–331, 2003.
[73] S. Mischler and C. Mouhot, *About Kac's program in kinetic theory*, C. R. Math. Acad. Sci. Paris **349**(23-24), 1245–1250, 2011.
[74] S. Mischler and C. Mouhot, *Kac's program in kinetic theory*, Invent. Math. **193**(1), 1–147, 2013.
[75] E. Nelson,*The free Markoff field*, J. Functional Analysis **12**, 211–227, 1973.
[76] G. Nenciu, *On asymptotic perturbation theory for quantum mechanics: almost invariant subspaces and gauge invariant magnetic perturbation theory.* J. Math. Phys. **43**, no. 3, 1273–1298, 2002.
[77] F. M. Odeh, *An existence theorem for the BCS integral equation*, IBM J. Res. Develop. **8**, 187–188, 1964.
[78] E. Sandier, S. Serfaty, *Vortices in the Magnetic Ginzburg-Landau Model*, Birkhäuser, 2006.
[79] M. Tinkham, *Introduction to superconductivity*, McGraw-Hill Book Co., New York, 1996.
[80] H. Tossounian and R. Vaidyanathan, *Partially thermostated Kac model*, J. Math. Phys. **56**(8), 2015.
[81] A. Vansevenant, *The gap equation in superconductivity theory*, Physica D Nonlinear Phenomena **17**, 1985.
[82] C. Villani, *Cercignani's conjecture is sometimes true and always almost true*, Comm. Math. Phys. **234**(3), 455–490, 2003.
[83] Y. Yang, *On the Bardeen-Cooper-Schrieffer integral equation in the theory of superconductivity*, Letters in Mathematical Physics **22**, no. 1, 27–37, 1991.
[84] Y. Yang, *Mathematical analysis of the multiband BCS gap equations in superconductivity*, Physica D: Nonlinear Phenomena **200**, no. 1-2, 60–74, 2005.
[85] H.-T. Yau, *The logarithmic Sobolev inequality for lattice gases with mixing conditions*, Comm. Math. Phys. **181**, 367–408, 1996.
[86] H.-T. Yau, *The logarithmic Sobolev inequality for generalized simple exclusion processes*, Probab. Theory and Related Fields **109**, 507–538, 1997.