Mathematics of Bose–Einstein condensation

Marcin Napiórkowski

February 2022

Chapter 1

Introduction

The phenomenon of Bose-Einstein condensation was predicted theoretically by Bose and Einstein in 1924. It is a state of matter in which, at low enough temperatures, a macroscopic number of particles occupies the lowest energy state. It is caused by the quantum statistics of bosons only and is therefore purely quantum mechanical. However, for more than ten years, Einstein's prediction has been neglected as purely mathematical conclusion for a fictitious system of non-interacting ideal gas and with little relevance to real physics. The idea remained abstract for a long time, even for the authors themselves. Einstein wrote: *The theory is pretty but is there also some truth to it?*

In 1938, Kapitza [1], and independently Allen and Misener, discovered the phenomenon of superfluidity - a frictionless flow of fluid in liquid ^{4}He. This remarkable phenomenon can be observed by naked eye. However, it cannot be understood in terms of classical physics. In the same year, London [2] had the idea that superfluidity could be an experimental manifestation of Bose–Einstein Condensation.

In 1941, L.D. Landau developed a phenomenological theory of superfluidity in terms of the excitation spectrum of ^{4}He. The first microscopic theory of interacting Bose gases in the context of BEC was formulated in 1947 by Bogoliubov. In spite of the successful development in understanding superfluidity in its early years, it was only after the realization of atomic BEC in 1995 (by Cornell and Wiemann and by Ketterle, all got the Nobel Prize) that the theoretical concepts of Einstein and Bogoliubov have been experimentally confirmed.

The relation between superfluidity and BEC is not obvious until today. In fact, to discuss the connection between these two one first needs to define them. This is related to the concept of the off-diagonal long-range order introduced by Penrose and Onsager.

The experimental realisation of BEC led to a renewed interest in the theoretical and mathematical study of bosonic many-body systems. It would

not be an exaggeration to say that the basis of much of the subsequent work has been laid in the seminal 2002 paper of Lieb and Seiringer [3] in which the authors proved for the first time that the ground state of a dilute, trapped gas exhibits Bose–Einstein condensation.

The goal of this lecture is to discuss some of the topics mentioned above in a mathematically precise way. We will define objects and set up problems in a rigorous way. Some proofs we will be done in details, others will only be sketched. From a mathematical point of view, most of the results we will discuss have been obtained very recently (up to 15 years). During the course we will formulate open problems which could lead to research and/or master/PhD theses.

Chapter 2

Quantum many-body systems

2.1 The Hamiltonian

The central object that we will study throughout the course is the non-relativistic many-body Hamiltonian given by

$$H_N = \sum_{j=1}^{N} \left(-\Delta_{x_j} + V(x_j)\right) + \lambda \sum_{1 \leqslant j < k \leqslant N} w_N(x_j - x_k). \tag{2.1}$$

It acts on the the Hilbert space $\mathfrak{H}^N = \bigotimes_{\mathrm{sym}}^N L^2(\mathbb{R}^3)$ which consists of functions $\psi(x_1, x_2, \ldots, x_N)$ that are square integrable, i.e.

$$\int |\psi(x_1, x_2, \ldots, x_N)|^2 \mathrm{d}x_1 \ldots \mathrm{d}x_N < \infty$$

and, since we consider bosons, are symmetric under the exchange of two particles, i.e.

$$\psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) = \psi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N)$$

for any $i, j = 1, \ldots, N$. Here $x_j \in \mathbb{R}^d$ stands for the coordinate of the j-th particle. We have set the spatial dimension of the problem to be three as we will focus on that case for the majority of the course. But, in principle, physically relevant systems could be also one or two dimensional.

The Hamiltonian (2.1) describes the system of N particles that interact through a two-body potential w_N that in principle can depend on N. It could be positive (repulsive) but it could also have an attractive part. A purely repulsive potential is usually easier to handle. Most of the time we will work with 'nice' (smooth, quickly decaying), symmetric, repulsive potentials which will make our live easier. The factor λ in front of the interaction terms

is assumed to be positive, i.e. $\lambda > 0$. It represents a coupling constant that could also be N dependent. We will soon come back to this.

Finally, the one-body term consists of the kinetic energy operator $-\delta_{x_j}$ of the j-th particle and a external potential $V(x_j)$ which acts on the j-th particle too. On should think of the V as being a trapping potential which keeps the particles located in a certain region of space. One example would be a trapping potential of the form $V(x) = x^2$ which represents a harmonic trap. Another example would be a potential that is of the form

$$V(x) = \begin{cases} 0 & \text{if } x \in \Omega \\ \infty & \text{if } x \notin \Omega \end{cases}$$

In this case the particles would be confined in $\Omega \subset \mathbb{R}^3$.

2.2 Quantities of interest

Given the Hamiltonian H_N of a quantum system, there are many questions one can try to address. The first one might concern its ground state energy

$$E_0(N) = \inf_{\psi \in D(\mathfrak{H}^N), \|\psi\|=1} (\psi, H_N \psi).$$

If $E_0(N)$ is an eigenvalue, the corresponding ground state wave function ψ_0 is determined by Schrödingers equation $H_N \psi_0 = E_0(N)\psi_0$.

One might be interested also in different observables. Recall, that the expectation value of an observable A in the state ψ is given by

$$\langle A \rangle_\psi = (\psi, A\psi).$$

More generally, if the system is at some positive temperature $T > 0$, one would like to compute the free energy of the system, given by

$$F = -T \ln \operatorname{Tr} e^{-H_N/T}.$$

We choose units such that Boltzmanns constant equals 1, and shall often write $T = 1/\beta$. The trace is over the physical Hilbert space, of course, respecting symmetry constraints arising from the indistinguishability of particles. The equilibrium state at temperature T is the Gibbs state

$$\rho_\beta = e^{-\beta(H_N - F)}.$$

It is normalized to have $\operatorname{Tr} \rho_\beta = 1$. The expectation value of an observable A is then given by

$$\langle A \rangle_{\rho_\beta} = \frac{\operatorname{Tr}\left(A e^{-H_N/T}\right)}{\operatorname{Tr} e^{-H_N/T}} = \operatorname{Tr}(A\rho_\beta).$$

For large particle number, it is usually hopeless to try to calculate ρ_β directly, but one can try to investigate properties of the reduced n-particle density matrices, obtained by taking the partial trace of ρ_β over $N - n$ variables. We will introduce some of these objects shortly.

2.3 Grand-canonical ensemble

It is often convenient not to fix the particle number N, but rather work in the grand-canonical ensemble, where one takes a certain average over the number of particles in the system. For simplicity, consider a system of just one species of particles. The N-particle Hilbert space, $\mathfrak{H}^N$, is then the set of square-integrable functions that are totally symmetric under permutations. In the grand-canonical ensemble, one has as Hilbert space the Fock space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathfrak{H}^N.$$

Here $\mathfrak{H}^0 = \mathbb{C}$ by definition and the corresponding vector is called the vacuum vector. As the Hamiltonian on Fock space one simply takes

$$H = \bigoplus_{N=0}^{\infty} H_N.$$

One typically sets $H_0 = 0$ (vacuum has no energy).

For $\mu \in \mathbb{R}$, the grand canonical potential is defined as

$$J = -T \ln \operatorname{Tr}_{\mathcal{F}} e^{-\beta(H-\mu\mathcal{N})}$$

where $\mathcal{N}$ denotes the number operator, i.e.,

$$\mathcal{N} = \bigoplus_{n=0}^{\infty} n.$$

Since H is particle number conserving, we can also write this as

$$J = -T \ln \sum_{N\geq 0} z^N \operatorname{Tr}_{\mathfrak{H}^N} e^{-\beta H_N},$$

where $z = e^{\beta\mu}$ is called the fugacity.

The grand-canonical Gibbs state is

$$\rho_{\beta,\mu} = e^{-\beta(H-\mu\mathcal{N}-J)}.$$

The chemical potential μ is adjusted to achieve a given average particle number $\langle\mathcal{N}\rangle$. The latter equals

$$\langle\mathcal{N}\rangle = \operatorname{Tr} \mathcal{N}\rho_{\beta,\mu} = -\frac{\partial}{\partial\mu} J.$$

2.4 BEC in ideal gas

Let us now consider an non-interacting (ideal) Bose gas, that is confined in a box of side length L, i.e. $\Lambda = [0, L]^3$. The Hamiltonian is given by

$$H_N^{\text{n-in}} = \sum_{i=1}^{N} -\Delta_{x_i}$$

and let us assume periodic boundary conditions. The spectrum of $-\Delta$ equals

$$\left(\frac{2\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2) \tag{2.2}$$

with $(n_x, n_y, n_z) \in \mathbb{Z}^3$. The corresponding eigenstates are plane waves e^{ipx} with $p \in \left(\frac{2\pi}{L}\mathbb{Z}\right)^3$. Let us order (and denote) the eigenvalues in (2.2) by

$$e_0 \leq e_1 \leq e_2 \leq \dots.$$

On Fock space we have

$$H = \sum_{i \geq 0} e_i a_i^* a_i$$

and also

$$\beta(H - \mu\mathcal{N}) = \sum_{i \geq 0} \varepsilon_i a_i^* a_i$$

where $\varepsilon_i = \beta(e_i - \mu)$.

We would like to calculate

$$\ln \operatorname{Tr} e^{-\sum_{i\geq 0} \varepsilon_i a_i^* a_i}.$$

The spectrum of $\sum_{i\geq 0} \varepsilon_i a_i^* a_i$ is of the form $\sum_{i\geq 0} \varepsilon_i n_i$, with $n_i \in \{0, 1, 2, \dots\}$. Summing over all possible occupation numbers is the same as summing over all eigenstates, hence we have

$$\operatorname{Tr} e^{-\sum_{i\geq 0} \varepsilon_i a_i^* a_i} = \prod_i \sum_n e^{-\varepsilon_i n_i} = \prod_i \frac{1}{1 - e^{-\varepsilon_i}}.$$

Here we have to assume that $\varepsilon_i > 0$ for all i for the geometric series to converge. Then

$$\ln \operatorname{Tr} e^{-\sum_{i\geq 0} \varepsilon_i a_i^* a_i} = \sum_i -\ln(1 - e^{-\varepsilon_i}).$$

Thus the grand-canonical potential equals

$$J = T \sum_{p \in \left(\frac{2\pi}{L}\mathbb{Z}\right)^3} \ln(1 - e^{-\beta(p^2 - \mu)}).$$

Note that $\varepsilon_i > 0$ can be achieved by taking $\mu < 0$. This is not really a restriction, however, as any particle number can be achieved even for negative μ. In fact, the average particle number equals

$$\langle \mathcal{N} \rangle = -\frac{\partial}{\partial \mu} J = \sum_{p \in \left(\frac{2\pi}{L}\mathbb{Z}\right)^3} \underbrace{\frac{1}{e^{\beta(p^2-\mu)}-1}}_{\langle a_p^* a_p \rangle}.$$

Here the summands are $\langle a_p^* a_p \rangle$, the average occupation number of momentum p.As μ varies between $(-\infty, 0)$, clearly $\langle \mathcal{N} \rangle$ varies between $(0, \infty)$.

We now perform a thermodynamic limit $L \to \infty$. The sum over p can then be interpreted as a Riemann sum for the corresponding integral. In fact,

$$\frac{1}{L^3} \sum_{p \in \left(\frac{2\pi}{L}\mathbb{Z}\right)^3} \longrightarrow \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} dp$$

as $L \to \infty$. The thermodynamic pressure of the system is thus

$$P = -\lim_{L\to\infty} \frac{J}{L^3} = -\frac{T}{(2\pi)^3} \int_{\mathbb{R}^3} \ln(1 - e^{-\beta(p^2-\mu)}) dp$$

and the average density equals

$$\rho = \frac{\langle \mathcal{N} \rangle}{L^3} = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{1}{e^{\beta(p^2-\mu)}-1} dp. \tag{2.3}$$

Notice that the density stays bounded as $\mu \to 0$! That is

$$\rho_{\mathrm{c}}(\beta) := \lim_{\mu\to 0} \rho = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{1}{e^{\beta p^2}-1} dp < \infty.$$

Hence it appears that the density of the gas can never exceed $\rho_{\mathrm{c}}(\beta)$. What is happening here? Recall that μ has to be chosen as to fix the density and, hence, has to depend on L, in general. If $\rho < \rho_{\mathrm{c}}(\beta)$, then $\mu(L) \to \mu < 0$ in the thermodynamic limit. But when $\rho > \rho_{\mathrm{c}}(\beta)$, then $\mu(L)$ has to tend to zero as $L \to \infty$. In this case, the limits $L \to \infty$ and $\mu \to 0$ must be taken simultaneously and, in particular, do not commute!

In fact, if $\rho > \rho_{\mathrm{c}}(\beta)$, then μ is asymptotically equal to

$$\mu = -(\beta L^3(\rho - \rho_{\mathrm{c}}(\beta)))^{-1} \qquad \text{as} \qquad L \to \infty.$$

For this value of μ we see that

$$\lim_{L\to\infty} \frac{\langle a_0^* a_0 \rangle}{L^3} = \lim_{L\to\infty} \frac{1}{L^3} \frac{1}{e^{-\beta\mu}-1} = \rho - \rho_{\mathrm{c}}(\beta). \tag{2.4}$$

Thus the zero momentum state is occupied by a macroscopic fraction of all the particles. This phenomenon is called Bose–Einstein Condensation (BEC). It occurs when $\rho > \rho_{\mathrm{c}}(\beta)$ or, equivalently, when

$$T < T_{\mathrm{c}}(\rho) = \frac{4\pi}{\zeta(3/2)^{2/3}} \rho^{2/3} \tag{2.5}$$

since $\rho_c(\beta) = \zeta(3/2)(4\pi)^{-3/2}\beta^{-3/2}$. Here ζ denotes the Riemann zeta function

$$\zeta(z) = \sum_{k\geq 1} \frac{1}{k^z}.$$

In other words, BEC occurs below the critical temperature.

Problem 2.4.1. *Derive* (2.5).

We note that only the zero momentum mode is macroscopically occupied, and the other occupations are much smaller. The smallest positive eigenvalue of the Laplacian equals $(2\pi/L)^2$, and

$$\frac{1}{e^{\beta(2\pi/L)^2} - 1} \sim L^2 \ll L^3 \qquad \text{for large} \qquad L.$$

BEC represents a phase transition in the usual sense: the thermodynamic functions exhibit a non-analytic behavior. Consider, for instance, the free energy, which is given in a standard way as the Legendre transform of the pressure. Specifically, the free energy per unit volume equals

$$f(\beta, \rho) = \mu\rho + \frac{T}{(2\pi)^3} \int_{\mathbb{R}^3} \ln(1 - e^{-\beta(p^2-\mu)}) dp,$$

where μ is determined by (2.3) if $\rho < \rho_c(\beta)$, and $\mu = 0$ if $\rho \geq \rho_c(\beta)$. In the latter case, we see that $f(\beta, \rho)$ does not actually depend on ρ, and is constant for $\rho > \rho_c(\beta)$. In particular, f is not analytic. Intuitively, what is happening as one increases the density beyond $\rho_c(\beta)$ is that all additional particles occupy the zero momentum mode and hence do not contribute to the energy or the entropy, hence also not to the free energy.

Finally, let us just mention, that the arguments that we have mentioned above do not work in dimensions $d = 1, 2$. In fact, in that case the ideal Bose gas exhibits BEC only in the ground state ($T = 0$).

2.5 BEC in interacting systems - definition

2.5.1 Reduced densities

The one-particle reduced density associated with the N-body wave function ψ_N is defined by

$$\gamma^{(1)}_{\psi_N} := N \operatorname{Tr}_{2,3,\ldots,N} |\psi_N\rangle\langle\psi_N|, \tag{2.6}$$

where $|\psi_N\rangle\langle\psi_N|$ denotes the orthogonal projection onto ψ_N and $\operatorname{Tr}_{2,3,\ldots,N}$ is the partial trace over the last $(N-1)$ particles. In other words, the one-particle reduced density $\gamma^1_{\psi_N}$ is defined as the non-negative trace-class operator on $L^2(\mathbb{R}^3)$ with integral kernel

$$\gamma^{(1)}_{\psi_N}(x; y) = N \int dx_2 \ldots dx_N \psi_N(x, x_2, \ldots, x_N) \overline{\psi_N(y, x_2, \ldots, x_N)}. \tag{2.7}$$

Notice that we chose the normalization $\operatorname{Tr}\gamma^{(1)}_{\psi_N} = N$.

Analogously, for $k = 2, 3, \ldots, N$, we can define the k-particle reduced density associated with ψ_N by

$$\gamma^{(k)}_{\psi_N} := \binom{N}{k} \operatorname{Tr}_{k+1,\ldots,N} |\psi_N\rangle\langle\psi_N|. \tag{2.8}$$

The integral kernel of the k-particle density matrix is given by

$$\gamma^{(k)}_{\psi_N}(x_1, \ldots, x_k; y_1, \ldots, y_k) = \binom{N}{k} \int dx_{k+1} \ldots dx_N \psi_N(x_1, \ldots, x_k, x_{k+1} \ldots, x_N) \overline{\psi_N(y_1, \ldots, y_k, x_{k+1}, \ldots, x_N)}.$$

The normalization is such that $\operatorname{Tr}\gamma^{(k)}_{\psi_N} = \binom{N}{k}$.

Clearly, for $k < N$, the k-particle reduced density $\gamma^{(k)}_{\psi_N}$ does not contain the full information about the system. Still, $\gamma^{(k)}_{\psi_N}$ is enough to compute the expectation of any k-particle observable. For example, if $k = 1$ and $J^{(1)}$ is a one body operator $J^{(1)} = \sum_{i=1}^N J^{(1)}_i$, then

$$(\psi_N, J^{(1)}\psi_N) = \sum_{i=1}^N (\psi_N, J^{(1)}_i \psi_N) = N(\psi_N, J^{(1)}_1 \psi_N) = \operatorname{Tr} J^{(1)}_1 \gamma^{(1)}_{\psi_N}.$$

The reduced densities can be lifted to the Fock space setting. Note that using (B.3) we see that the integral kernel (2.7) can be written as

$$\gamma^{(1)}_{\psi_N}(x; y) = (\psi_N, a^*_y a_x \psi_N).$$

This leads to the general definition that the one-body reduced density matrix is the one-body operator (acting on the one-body Hilbert space) $\gamma^{(1)}$ defined through

$$(g, \gamma^{(1)} f) = \langle a^*(f) a(g) \rangle, \qquad \forall f, g \in L^2(\mathbb{R}^3).$$

Here $\langle \cdot \rangle$ denotes the expectation value in any state. In particular this definition applies to any state on Fock space, not only thermal equilibrium states. One can also consider states of definite particle number, and hence recover the definition for the canonical ensemble.

2.5.2 Definition of BEC

Let $\gamma^{(1)}$ be the reduced one-body matrix of a bosonic system. It is a positive, trace class operator. Thus, it admits a spectral decomposition with respect to an orthonormal basis $\{u_i\}$ (cf. (A.2)) of the form

$$\gamma^{(1)} = \sum_i \lambda_i |u_i\rangle\langle u_i|,$$

with $\lambda_i > 0$. According to the (generally accepted) definition of Penrose and Onsager, Bose–Einstein Condensation occurs when $\gamma^{(1)}$ has an eigenvalue of order N (or $\langle \mathcal{N} \rangle$). Note that this definition is independent of the fact, whether the system under consideration is interacting or not.

Let us now check, how this definition relates to the concept of BEC derived for the ideal Bose gas. In (2.4) we have shown that above the critical density

$$0 < \rho - \rho_{\mathrm{c}}(\beta) = \lim_{L\to\infty} \frac{\langle a_0^* a_0 \rangle}{L^3} < \lim_{L\to\infty} \frac{\sup_{\|f\|=1} \langle a^*(f) a(f) \rangle_{\beta,\mu}}{L^3}.$$

In particular, $\sup_{\|f\|=1} \langle a^*(f) a(f) \rangle_{\beta,\mu}$ is the largest eigenvalue of $\gamma^{(1)}$ so indeed it has to be of order N (as when taking the thermodynamic limit we assume $\rho = N/L^3$).

2.5.3 Off-diagonal long-range order

Consider the momentum distribution in a given state defined through

$$n_p := \langle a_p^* a_p \rangle.$$

As, by definition, $a_p = \int dx a_x e^{-ipx}/L^{3/2}$, we have

$$\langle a_p^* a_p \rangle = \frac{1}{L^3} \int dx dy \langle a_x^* a_y \rangle e^{-ip(x-y)} = \frac{1}{L^3} \int dx dy \gamma^{(1)}(x,y) e^{-ip(x-y)}.$$

Let us assume that the system under consideration is translation invariant. Then we have $\gamma^{(1)}(x,y) = \gamma^{(1)}(x-y)$ and thus

$$\langle a_p^* a_p \rangle = \frac{1}{L^3} \int dx dy \gamma^{(1)}(x-y) e^{-ip(x-y)}. \tag{2.9}$$

Inverting this relation (inverse Fourier transform), we obtain

$$\gamma^{(1)}(x-y) = \frac{1}{L^3} \sum_{p \in (2\pi/L)\mathbb{Z})^3} \langle n_p \rangle e^{ip(x-y)}.$$

Now, recall our analysis of BEC in the ideal case. It shows that when there is BEC we have (in the thermodynamic limit)

$$\langle n_p \rangle = \delta_0 N_0 + \tilde{n}_p$$

where the singular term arises from the macroscopic occupation of the zero momentum state. Thus, taking the thermodynamics limit in (2.9), we obtain

$$\gamma^{(1)}(x-y) = N_0 + \int dp\, \tilde{n}_p e^{ip(x-y)}.$$

By the expected regularity of $\tilde{n}_p$ and the Riemann-Lebesgue lemma we obtain

$$\gamma^{(1)}(x-y) \to N_0, \qquad \text{as} \qquad x - y \to \infty. \tag{2.10}$$

This behaviour is called off-diagonal (because $x \neq y$) long-range (because $x - y \to \infty$) order of the one-body density.

2.6 The curse of dimensionality

We shall now argue that, due to its high dimensionality, the many-body Schrödinger equation is impossible to solve numerically at a high precision for most physical systems of interest.

To this end let us consider a much simpler, discrete model - a quantum spin system. We place a quantum (say, 1/2) spin on each of the N sites of a square lattice. Since each spin is described the Hilbert space $\mathbb{C}^2$, the Hilbert space of the system is given by the tensor product $\otimes_x \mathbb{C}^2$ where x enumerates all the N sites. Thus, the dimension of the Hilbert space is 2^N. In other words, to encode a general quantum state we need 2^N coordinates. Assume we want consider a square lattice with length 20 and thus with 400 sites altogether. Now, let us assume very optimistically, that one number can be encoded using one bit which is carried on one elementary particle. Thus, we would need 2^{400} elementary particles just to encode on the computer one state of our quantum spin system. How much is 2^{400}? According to recent estimates of astrophysicists, there are 10^{86} elementary particles in the universe. Since $2^{400} > 10^{86}$, in principle we are not able to encode a quantum state on the computer without even doing any operations on it.

That was a crude estimate for a spin system. The models we want to consider are continuous and consist of thousands of particles (like in modern experiments involving Bose–Einstein Condensates). To put the problem on a computer one needs to discretize it and the example discussed above shows that the complexity of this problem will be exponential.

All this forces us to consider effective theories. An effective theory is a proposal to describe a system using less degrees of freedom. Of course, there is price to pay - one looses information about the system. Let us illustrate this idea using an example from classical physics. In classical mechanics, the microscopic theory is given by Newton's equations of motion. Solving these equations would in principle give us information about all positions and momenta of all particles involved. If one wants to understand the behaviour of air in a room, then solving Newton's equations is not a good idea to say the least. But do we really need to know all the positions and velocities of all particles? Maybe it would be enough to know the position and momentum of a *typical* particle - that is, the probability that the particle occupies a given very small region of the phase space (mathematically the volume element $dxdp$) at an instant of time. We would thus be looking for a probability distribution $f(t, x, p)$ which now depends on only three variables. Ludwig Boltzmann derived an equation for f which is now called the Boltzmann equation. A rigorous justification of this equation starting from Newton's microscopic theory is still an active area of research and this field of mathematical physics is called kinetic theory.

In this course we will try to understand to most important properties of quantum many-boson systems by rigorously justifying effective theories. We

will be interested in effective descriptions of the the ground state properties, excitation spectrum, effective dynamics. An effective theory usually is some kind of limiting theory in a certain regime. In the next section we will discuss the scaling regimes we will be studying.

2.7 Scaling regimes

Recall the Hamiltonian in (2.1) given by

$$H_N = \sum_{j=1}^{N} \left(-\Delta_{x_j} + V(x_j)\right) + \lambda \sum_{1 \leqslant j < k \leqslant N} w_N(x_j - x_k),$$

defined on $\mathfrak{H}^N = \bigotimes_{\mathrm{sym}}^N L^2(\mathbb{R}^3)$. We want to find an effective description of the model. Keeping in mind that we are interested in thermodynamic quantities, we have to remember that we we will be taking the $N \to \infty$ limit. We observe that the kinetic energy (or, more generally, the one-body operator) is of order N as it consists of N terms which are independent of N. If we assume that the two-body interaction is N independent, i.e. $w_N \equiv w$, then the interaction term is of order N^2 (for large particle numbers) as there are $N(N-1)/2$ terms in the double sum. In the limit when $N \to \infty$ the interaction term would thus become dominant. But we want to keep the kinetic term as, in some sense, it is more quantum than the pure interaction (recall, BEC has been computed for the ideal gas). Thus, the idea is to balance these two terms (at least, as far as the naive asymptotics in N would suggest). This motivates setting $\lambda = O(N^{-1})$ and this scaling is called *mean-field scaling*. The simplest mean-field Hamiltonian is of the form

$$H_N = \sum_{j=1}^{N} \left(-\Delta_{x_j} + V(x_j)\right) + \frac{1}{N} \sum_{1 \leqslant j < k \leqslant N} w(x_j - x_k).$$

Note, that in this setup there are two length scales. The first one is given by the trapping potential V and basically tells us where the particles can move. It is important for the kinetic term, as it has implications in the spectral gap of the one-body operator.

But for now let us focus on the length scale set by the interaction. When $w_N \equiv w$, then it is independent of N and therefore is of order $O(1)$ (with respect to N). In fact, the length scale of the interaction is characterized by a parameter called the *scattering length*. To define the scattering length, we consider the two-body scattering problem. We assume the two-body interaction is radial, positive and has a finite range. Consider the zero-energy scattering equation

$$-\Delta f + \frac{1}{2} w f = 0 \tag{2.11}$$

with the boundary condition $f(x) \to 1$ as $|x| \to \infty$. Under the assumption of compact support of w one can show that for $|x|$ sufficiently large, the scattering solution satisfies

$$f(x) = 1 - \frac{a_0}{|x|} \tag{2.12}$$

for a constant $a_0 > 0$. This constant is called the scattering length. Equivalently, the scattering length a_0 can also be defined through the integral

$$8\pi a_0 = \int w(x) f(x) dx \tag{2.13}$$

where $f(x)$ is the solution of (2.11). From the point of view of physics, the scattering length a_0 measures the effective range of the interaction potential; two quantum mechanical particles interacting through the potential w, when they are far apart, feel the other particle as a hard sphere with radius a_0 (in particular, the scattering length of a hard sphere potential coincides with the radius of the sphere).

Problem 2.7.1. *Show that* (2.13) *holds true.*

Problem 2.7.2. *Consider the hard sphere potential. Show that the scattering length is in this case equal to the radius of the hard sphere.*

Problem 2.7.3. *Let f be the solution of the scattering equation for a positive interaction potential. Show that $0 \leq f \leq 1$. Deduce that*

$$a_0 \leq \frac{\int w}{8\pi}.$$

Having defined the scattering length, let us come back to the mean-field model. Assuming the trapping potential restricts the particles to a volume of order $O(1)$, we see that the density is of order N and the average distance between the particles is $N^{-1/3}$. Thus it is much smaller that the effective range of the interaction. In other words, each particle feels and interacts with many other particles, but the strength of interactions is weak (of order $O(1)$). This situation corresponds to a high density regime where the particles meet very often but interact only a little bit each time.

To model a different situation, which will call the dilute regime (a typical situation in experiments with ultracold gases), we can scale the interaction and make it N dependent. Assume $w_N(x) = N^2 w(Nx)$. By rescaling $x = N\tilde{x}$ we see from (2.11) that

$$\left(-\frac{1}{N^2} \Delta_{\tilde{x}} + \frac{1}{2} w(N\tilde{x}) \right) f(N\tilde{x}) = 0$$

Thus $f(N\cdot)$ solves the scattering equation

$$\left(-\Delta + \frac{1}{2} N^2 w(N\cdot) \right) f(N\cdot) = 0.$$

It follows that as $|x| \to \infty$ we have

$$f(Nx) = 1 - \frac{a_0}{N|x|} = 1 - \frac{a_0/N}{|x|}$$

and thus $a = a_0/N$ is the scattering length of the rescaled potential $w_N(x) = N^2 w(Nx)$. This scaling corresponds to the dilute limit because the effective range of the interaction is $O(N^{-1}) \ll N^{-1/3}$ is much smaller than the average inter-particle distance. Thus, this scaling models a very short range interaction that is very strong. Indeed, for a nice function w, the scaled w_N converges in a vague sense to a Dirac delta. Each particle encounters rare but strong collisions. This scaling is called the *Gross–Pitaevskii limit*.

Finally, one can interpolate between these two extreme cases. This is done by introducing $\beta \in [0, 1]$ (do not confuse with inverse temperature). For a general β we define

$$w_N(x) = N^{3\beta} w(N^\beta x).$$

Now, heuristically, the effective range of the interaction is of order $N^{-\beta}$ (heuristically, because the scattering length does not scale nicely for intermediate β's). A special case corresponds to $\beta = 1/3$ (or, more generally $\beta = 1/d$ in d dimensions) because for $\beta < 1/3$ the average distance between the particles is much bigger than the effective range of interaction (mean-field limit), while for $\beta > 1/3$ the average distance between the particles is much smaller than the effective range of interaction (GP/dilute limit). These two regimes are very different but they lead to somewhat similar effects. This similarity is both very useful and confusing. Generally speaking, the larger β is, the more difficult the problem is.

Finally, let us mention a slightly different scaling. While the setup described above fits well when one wants to describe a trapped gas o bosons, a different approach is usually considered when one wants to describe the thermodynamic in its classical formulation, that is when the system is restricted to a box of length L and in the end one that the limit $N \to \infty, L \to \infty$ with fixed $\rho = N/L^3$. In this case, the dilute limit is characterized by the relation

$$a\rho^{1/3} \ll 1, \tag{2.14}$$

which means that the effective range of the interaction a is much smaller than the average distance between the particles $\rho^{-1/3}$.

Chapter 3

Ground state properties

We shall now discuss the properties of the ground state energy (i.e. $T = 0$) of a system of interacting bosons, i.e.

$$E_0(N) = \inf_{\psi \in L^2_{\mathrm{sym}}(\mathbb{R}^{3N}), \|\psi\|=1} (\psi, H_N \psi)$$

where

$$H_N = \sum_{j=1}^{N} \underbrace{\big(-\Delta_{x_j} + V(x_j)\big)}_{=:h_{x_j}} + \frac{1}{N-1} \sum_{1 \leqslant j < k \leqslant N} w_N(x_j - x_k)$$

with $w_N(x) = N^{3\beta} w(Nx)$.

3.1 Mean-field scaling

3.1.1 Ground state energy

Let us first consider the mean-field scaling, i.e. when $\beta = 0$. The mean-field Hamiltonian is of the form

$$H_N = \sum_{j=1}^{N} \big(-\Delta_{x_j} + V(x_j)\big) + \frac{1}{N-1} \sum_{1 \leqslant j < k \leqslant N} w(x_j - x_k). \tag{3.1}$$

Recall, that the mean-field scaling correspond to a situation when the interaction between the particles is frequent, but weak. We know that the ground state of a non-interacting system (i.e. when $w \equiv 0$) is given by a product state $\tilde{u}^{\otimes N}(x_1, \ldots, x_N) := \tilde{u}(x_1) \cdots \tilde{u}(x_N)$ where $\tilde{u}$ is the (normalized) ground state of the one-body operator. In the mean-field scaling the interactions are supposed to be weak, so a simple minded argument suggests that maybe the ground-state will not be changed dramatically and will be close, in some sense, also to a product state.

Let us make this ansatz, i.e. $\psi(x_1,\dots,x_N) = u^{\otimes N}(x_1,\dots,x_N)$ for some normalized function $u \in L^2(\mathbb{R}^3)$. Computing the energy leads to the following formula

$$(u^{\otimes N}, H_N u^{\otimes N}) = N\mathcal{E}(u)$$

where

$$\mathcal{E}(u) = (u, hu) + \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} w(x-y)|u(x)|^2|u(y)|^2 dxdy. \tag{3.2}$$

The functional (3.2) is called the *Hartree functional*.

Problem 3.1.1. *Derive the Hartree functional.*

Thus we obtain the following upper bound on the ground-state energy

$$E_0(N) = \inf_{\psi\in L^2_{\mathrm{sym}}(\mathbb{R}^{3N}),\|\psi\|=1} (\psi, H_N\psi) \le N \inf_{u\in L^2(\mathbb{R}^3),\|u\|=1} \mathcal{E}(u). \tag{3.3}$$

Problem 3.1.2. *Let us assume that the Hartree functional has a minimizer and let us denote it by u_0. It satisfies the so-called Hartree equation*

$$hu_0 + \left(|u_0|^2 * w\right) u_0 = \varepsilon_0 u_0 \tag{3.4}$$

where $(f*g)(x) = \int_{\mathbb{R}^3} f(x-y)g(y)dy$ denotes the convolution and ε_0 is a Lagrange multiplier due to the normalization constraint.

Note that in general there is no reason to expect that minimizers will be unique. Uniqueness can be broken because of the one-body Hamiltonian h or due to the interaction if it has an attractive part.

Let us denote by e_{H} the infimum of the Hartree functional, i.e.

$$e_{\mathrm{H}} := \mathcal{E}(u_0).$$

From (3.3) we obtain

$$\frac{E_0(N)}{N} \le e_{\mathrm{H}}.$$

A natural question whether a matching lower bound holds. This is indeed true.

Theorem 3.1.3 (Convergence of ground-state energy)**.** *Under appropriate assumptions on h and w one has*

$$\lim_{N\to\infty} \frac{E_0(N)}{N} = e_{\mathrm{H}}.$$

We will not state the most general assumptions on h and w that were used to prove the most general version of Theorem 3.1.3 given by Lewin–Nam–Serfaty–Solovej in 2014. Let us just mention that the one-body operator h can be much more general than in (3.1).

We will give a proof in a simpler case. We follow the proof of Lewin (2015). Let us start with the simplest case when the following two conditions are satisfied:

- h is symmetric and positive preserving: $(u, hu) \geq (|u|, h|u|)$;
- w is positive-definite, i.e. $w \geq 0$ and $\hat{w} \geq 0$.

These two properties can be used through the following two lemmas.

Lemma 3.1.4 (Hoffmann–Ostenhof inequality)**.** *Let h be symmetric and positive preserving. Then, for any bosonic N-body wave-function ψ_N we have*

$$(\psi_N, \sum_{i=1}^{N} h_{x_i} \psi_N) \geq N(\sqrt{\rho_{\psi_N}}, h\sqrt{\rho_{\psi_N}})$$

where ρ_{ψ_N} is the one-body density defined by

$$\rho_{\psi_N}(x) = \int |\psi_N(x, x_2, \ldots, x_N)|^2 dx_2 \ldots dx_N.$$

Proof. We know from the definition of the reduced one-body matrix that

$$(\psi_N, \sum_{i=1}^{N} h_{x_i} \psi_N) = \mathrm{Tr}(h\gamma^{(1)}_{\psi_N}) = \sum_k n_k (u_k, hu_k) \geq \sum_k n_k (|u_k|, h|u_k|)$$

where we used the decomposition $\gamma^{(1)}_{\psi_N} = \sum_k n_k |u_k\rangle\langle u_k|$ and the assumption on h. For real functions u_1, u_2 we have

$$(u_1, hu_1) + (u_2, hu_2) = (u_1 + iu_2, h(u_1 + iu_2)) \geq \left(\sqrt{u_1^2 + u_2^2}, h\sqrt{u_1^2 + u_2^2}\right).$$

Using this inductively we obtain

$$(\psi_N, \sum_{i=1}^{N} h_{x_i} \psi_N) \geq \sum_k n_k (|u_k|, h|u_k|) \geq \left(\sqrt{\sum_k n_k |u_k|^2}, h\sqrt{\sum_k n_k |u_k|^2}\right).$$

Using the fact that $\sum_k n_k |u_k|^2 = \gamma^{(1)}_{\psi_N}(x, x) = N\rho_{\psi_N}(x)$ we get the desired result. □

Lemma 3.1.5 (Estimating the two-body interaction by a one-body term)**.** *If $0 \leq \hat{w} \in L^1(\mathbb{R}^3)$, the for all $\eta \in L^1(\mathbb{R}^3)$ we have*

$$\sum_{1 \leq i < j \leq N} w(x_i - x_j) \geq \sum_{i=1}^{N} \eta * w(x_i) - \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} w(x - y)\eta(x)\eta(y) dxdy - \frac{N}{2} w(0).$$

Proof. We use

$$\iint_{\mathbb{R}^3 \times \mathbb{R}^3} w(x - y) f(x) f(y) dxdy = (2\pi)^{3/2} \int_{\mathbb{R}^3} \hat{w}(k) |\hat{f}(k)|^2 dk \geq 0$$

with $f = \sum_{i=1}^{N} \delta_{x_i} - \eta$ and expanding all terms. □

We are ready to give the proof of Theorem 3.1.3.

Proof. Take $\eta = N\rho_{\psi_N}$. Then the two lemmas above imply

$$(\psi_N, H_N\psi_N) \geq N\mathcal{E}(\sqrt{\rho_{\psi_N}}) - \frac{Nw(0)}{2(N-1)} \geq Ne_{\mathrm{H}} - \frac{Nw(0)}{2(N-1)}.$$

Minimizing over ψ_N and recalling the upper bound gives the theorem. □

3.1.2 Ground state

Let us now give a sketch of the proof that the ground state in the mean-field scaling exhibits Bose–Einstein Condensation. In fact, the proof below will also give the statement about the ground state-energy as a by-product.

Before we start let us introduce a rescaled definition of the reduced density matrices. Let us assume for the rest of this subsection that the reduced density matrices are normalized differently than in (2.5.1), i.e. we shall assume for now that

$$\gamma^{(k)}_{\psi_N}(x_1,\dots,x_k;y_1,\dots,y_k) = \int dx_{k+1}\dots dx_N \psi_N(x_1,\dots,x_k,x_{k+1}\dots,x_N)\overline{\psi_N(y_1,\dots,y_k,x_{k+1},\dots,x_N)}.$$

The normalization is then such that $\operatorname{Tr}\gamma^{(k)}_{\psi_N} = 1$.

Theorem 3.1.6 (Convergence of ground states)**.** *Assume that h and w satisfy the inequality*

$$-C_1(T_1+T_2) - C \leq w(x_1-x_2) \leq C_2(T_1+T_2) + C$$

for some constants $0 \leq C_1, C_2 < 1$. Assume $V(x) \to \infty$ as $|x| \to \infty$ and that w is symmetric and smooth enough. Let ψ_N be the ground state for H_N and $\gamma^{(n)}_{\psi_N}$ its n-th reduced density matrix with the normalization that

Then there exists a subsequence and a probability measure μ on the set of minimizers of the Hartree functional $\mathcal{M}$ such that

$$\gamma^{(k)}_{\psi_{N_j}} \to \int_{\mathcal{M}} |u^{\otimes k}\rangle\langle u^{\otimes k}| d\mu(u) \tag{3.5}$$

strongly in the trace-class as $N_j \to \infty$, for all $k \geq 1$.

The probability measure μ describes the fragmented Bose–Einstein condensation. For instance, if there are only two Hartree minimizers, then μ will give their relative occupations. The simplest case is when $\mathcal{M} = \{u_0\}$, where there will always be complete Bose–Einstein condensation on u_0.

The main tool for proving the theorems is the quantum de Finetti theorem. The latter is an abstract result which says that, at the level of reduced density matrices, only product states remain in the limit $N \to \infty$, for any sequence of bosonic states.

Theorem 3.1.7 (quantum de Finetti)**.** *Let $\mathcal{H}$ be a separable Hilbert space and let $\Gamma^{(n)}$ be a sequence of positive, trace-class operators satisfying*

$$\Gamma^{(0)} = 1, \mathrm{Tr}_{\mathfrak{H}^n_{sym}}[\Gamma^{(n)}] = 1, \mathrm{Tr}_{n+1}[\Gamma^{(n+1)}] = \Gamma^{(n)}$$

where Tr_{n+1} is the partial trace with respect to the last variable in $\mathfrak{H}^{n+1}$. Then there exists a unique Borel probability measure μ on the sphere $S\mathfrak{H} = \{u \in \mathfrak{H} : \|u\| = 1\}$ of $\mathfrak{H}$, invariant with respect to multiplication by a phase, such that

$$\Gamma^{(n)} = \int_{S\mathfrak{H}} |u^{\otimes n}\rangle\langle u^{\otimes n}| d\mu(u)$$

We will not give a proof of this statement now.

Proof of Theorem 3.1.6. We start by writing

$$\begin{aligned}\frac{E_0(N)}{N} &= \frac{1}{N}(\psi_N, H_N\psi_N)_{\mathfrak{H}^N_{\mathrm{sym}}} = \mathrm{Tr}_{\mathfrak{H}}[T\gamma^{(1)}_{\psi_N}] + \frac{1}{2}\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[w\gamma^{(2)}_{\psi_N}] \\ &= \frac{1}{2}\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w)\gamma^{(2)}_{\psi_N}]\end{aligned}$$

Notice that after this rewriting the the N-dependence now appears only in $\gamma^{(2)}_{\psi_N}$. To describe the limit of this object we will use the quantum de Finetti theorem.

To this end notice that, by definition, the sequences $\gamma^{(n)}_{\psi_N}$ are bounded uniformly (by 1) in trace class. This allows us to use the Banach–Alaoglu theorem, to find a converging subsequence. Indeed, the Banach–Alaoglu theorem states that, if X is a separable Banach space, then to any sequence (x^*_n) in X^* (the dual of X) which is bounded, i.e. $\|x^*_n\| \leq M$, there exists a weak-$*$ convergent subsequence $(x^*_{n_k})$. Weak-$*$ convergent means that there exists a $x^* \in X^*$ such that $x^*_{n_k}(x) \to x^*(x)$ as $k \to \infty$ for all $x \in X$. Moreover, $\|x^*\| \leq M$.

In our case X^* is the space of trace class operators. X is then the space of compact operators. Thus for any compact operator K_n on $\mathfrak{H}^n$ we have

$$\mathrm{Tr}_{\mathfrak{H}^n}[\gamma^{(n)}_{\psi_N} K_n] \to \mathrm{Tr}_{\mathfrak{H}^n}[\gamma^{(n)} K_n]$$

along a subsequence. We will show that this convergence is actually strong. To this end it is enough (this is easy to prove in Hilbert space setting, for trace class it is not so obvious, but still true) to show that $\mathrm{Tr}[\gamma^{(n)}] = 1$, i.e. no mass is lost in the limit.

Using the assumptions of the theorem, we compute

$$\begin{aligned}e_{\mathrm{H}} &\geq \frac{1}{2}\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w)\gamma^{(2)}_{\psi_N}] \\ &\geq (1 - C_1)\frac{1}{2}\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2)\gamma^{(2)}_{\psi_N}] - C\,\mathrm{Tr}[\gamma^{(2)}_{\psi_N}] \\ &= (1 - C_1)\,\mathrm{Tr}_{\mathfrak{H}}[T\gamma^{(1)}_{\psi_N}] - C.\end{aligned}$$

It follows that $\mathrm{Tr}_{\mathfrak{H}}[T\gamma^{(1)}_{\psi_N}]$ is uniformly bounded and thus so is $\mathrm{Tr}_{\mathfrak{H}}[(T + C_0)\gamma^{(1)}_{\psi_N}]$ which by the cyclicity if the trace yields (possibly, up to a further subsequence) that

$$(T + C_0)^{1/2}\gamma^{(1)}_{\psi_N}(T + C_0)^{1/2} \rightharpoonup^* (T + C_0)^{1/2}\gamma^{(1)}(T + C_0)^{1/2}$$

for some C_0. Consequently,

$$\begin{aligned} 1 = \mathrm{Tr}_{\mathfrak{H}}[\gamma^{(1)}_{\psi_N}] &= \mathrm{Tr}_{\mathfrak{H}}[(T + C_0)^{-1}(T + C_0)^{1/2}\gamma^{(1)}_{\psi_N}(T + C_0)^{1/2}] \\ &\to \mathrm{Tr}_{\mathfrak{H}}[(T + C_0)^{-1}(T + C_0)^{1/2}\gamma^{(1)}(T + C_0)^{1/2}] = \mathrm{Tr}_{\mathfrak{H}}[\gamma^{(1)}] \end{aligned}$$

since $(T + C_0)^{-1}$ is by the assumptions on $V(x)$ a compact operator. The fact that $\mathrm{Tr}_{\mathfrak{H}^n}[\gamma^{(n)}] = 1$ follows from a similar argument using the fact that

$$\mathrm{Tr}_{\mathfrak{H}}[T\gamma^{(1)}_{\psi_N}] = \frac{1}{n}\,\mathrm{Tr}_{\mathfrak{H}^n}\left[\sum_{j=1}^{n} T_j\gamma^{(n)}_{\psi_N}\right]$$

is also uniformly bounded in N and that $\sum_{j=1}^{n} T_j$ also has compact resolvent which allows for similar argument.

Thus we have proven for any $n \in \mathbb{N}$ that

$$\gamma^{(n)}_{\psi_N} \to \gamma^{(n)}$$

strongly in trace class. Testing this convergence against a bounded operator $B_{n+1} = B_n \otimes \mathbb{1}$ we obtain

$$\mathrm{Tr}_{\mathfrak{H}^{n+1}}[\gamma^{(n+1)}_{\psi_N}(B_n \otimes \mathbb{1})] \to \mathrm{Tr}_{\mathfrak{H}^{n+1}}[\gamma^{(n+1)}(B_n \otimes \mathbb{1})] = \mathrm{Tr}_{\mathfrak{H}^n}[(\mathrm{Tr}_{n+1}\,\gamma^{(n+1)})B_n]$$

but also

$$\mathrm{Tr}_{\mathfrak{H}^{n+1}}[\gamma^{(n+1)}_{\psi_N}(B_n\otimes\mathbb{1})] = \mathrm{Tr}_{\mathfrak{H}^n}[(\mathrm{Tr}_{n+1}\,\gamma^{(n+1)}_{\psi_N})B_n] = \mathrm{Tr}_{\mathfrak{H}^n}[\gamma^{(n)}_{\psi_N}B_n] \to \mathrm{Tr}_{\mathfrak{H}^n}[\gamma^{(n)}B_n].$$

This implies

$$\mathrm{Tr}_{n+1}\,\gamma^{(n+1)} = \gamma^{(n)}$$

and thus we can use Theorem 3.1.7. Before that, using the fact that (by assumptions) $T_1 + T_2 + w$ is bounded below by, say, C_T, we write

$$\begin{aligned} \liminf_{N\to\infty} \frac{1}{2}\,\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w)\gamma^{(2)}_{\psi_N}] &= \frac{1}{2}\,\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w - 2C_T)\gamma^{(2)}_{\psi_N}] + C_T \\ &\geq \frac{1}{2}\,\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w - 2C_T)\gamma^{(2)}] + C_T \\ &= \frac{1}{2}\,\mathrm{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}}[(T_1 + T_2 + w)\gamma^{(2)}] \end{aligned}$$

where we used Fatou's lemma for positive operators. Using the quantum de Finetti theorem we obtain

$$e_{\mathrm{H}} \geq \liminf_{N\to\infty} \frac{E_0(N)}{N} \geq \int_{u\in S\mathfrak{H}} \frac{1}{2} \operatorname{Tr}_{\mathfrak{H}^2_{\mathrm{sym}}} [(T_1 + T_2 + w)|u^{\otimes 2}\rangle\langle u^{\otimes 2}|] d\mu(u)$$
$$\geq \int_{u\in S\mathfrak{H}} \mathcal{E}_{\mathrm{H}}(u) d\mu(u) = e_{\mathrm{H}}.$$

The second part of the theorem follows from the fact that all inequalities become equalities. □

Thus we have the following Corollary:

Corollary 3.1.8 (Bose–Einstein condensation in the mean-field limit)**.** *Assume there exists a unique minimizer u_0 for the Hartree functional. Then for a sequence of ψ_N of* (3.1) *satisfying*

$$(\psi_N, H_N\psi_N) = Ne_{\mathrm{H}} + o(N)$$

we have

$$\gamma^{(1)}_{\psi_N} \to |u_0\rangle\langle u_0|.$$

Thus the ground state exhibits Bose–Einstein condensation.

Proof. It follows from taking the partial trace in (3.5). □

3.2 Gross–Pitaevskii scaling

One would like to know, whether statements similar to Theorems 3.1.3 and 3.1.6 hold true also in the case when the interaction becomes short range, i.e. when

$$H_N = \sum_{j=1}^{N} \left(-\Delta_{x_j} + V(x_j)\right) + \frac{1}{N-1} \sum_{1\leqslant j<k\leqslant N} w_N(x_j - x_k). \tag{3.6}$$

with $w_N(x) = N^{3\beta} w(Nx)$ and $\beta \in (0, 1]$. For strictly positive β, the function w_N provides an approximation to the identity (up to normalization), i.e.

$$w_N(x) = N^{3\beta} w(Nx) \to_{N\to\infty} \left(\int w\right) \delta_0$$

in the sense that

$$|u_0|^2 * w_N(x) \to \left(\int w\right) |u_0|^2(x).$$

Consequently, in the scaling regime when $\beta > 0$ the Hartree equation (3.4) has to be modified and the new effective equation reads

$$hu_0 + \left(\int w \right) |u_0|^2 u_0 = \varepsilon_0 u_0. \tag{3.7}$$

This equation is sometimes called the nonlinear Schrödinger (NLS) equation. It turns out that the NLS equation provides a good effective description only when $\beta < 1$. When $\beta = 1$, then a subtle phenomenon takes place which results in the appearance of the scattering length in the effective equation, the so-called Gross–Pitaevskii equation

$$hu_0 + 8\pi a_0 |u_0|^2 u_0 = \varepsilon_0 u_0.$$

The corresponding functional is the Gross–Pitaevski functional

$$\mathcal{E}^{\mathrm{GP}}(u) = (u, hu) + 4\pi a_0 \int_{\mathbb{R}^3} |u(x)|^4 dx. \tag{3.8}$$

If one assumes the two-body interaction is positive, then Lieb and Seiringer proved that statements analogous to Theorems 3.1.3 and 3.1.6 hold true when $\beta = 1$ with the effective theory now being given by the Gross–Pitaevskii functional. The presence of the scattering length is, from a physics perspective, crucial as it is a parameter which experimentalists can manipulate.

One can interpret these results as a universality result. The effective theory does not depend on the details of the underlying microscopic system, but rather on an quasi-macroscopic quantity that is accessible in the lab.

Finally, let us mention that the Gross–Pitaevski functional cannot be derived using a product wave-function as an ansatz. This would to the NLS functional. It follows that the trial wave-function must include correlations between the particles. Those correlations appear on very short length scales and are therefore difficult to capture. We will not discuss the related details in this course and refer to the original papers.

Chapter 4

Bogoliubov theory

4.1 Landau criterion for superfluidity

It is observed experimentally that a moving superfluid doesn't dissipate its kinetic energy. Landaus theory of superfluids is based on the Galilean transformation of energy and momentum.

Let E and P be the energy and momentum of the fluid in a reference frame K. If we try to express the energy and momentum of the same fluid but in a moving frame K', which has a relative velocity V with respect to a reference frame K, we have the following relations:

$$\begin{aligned} P' &= P - MV, \\ E' = P'^2/(2M) &= \frac{1}{2M}|P - MV|^2 = E - PV + \frac{1}{2}MV^2 \end{aligned} \tag{4.1}$$

where M is the total mass of the fluid.

We first consider a fluid at zero temperature, in which all particles are in the ground state and flowing along a capillary at constant velocity v. If the fluid is viscous, the motion will produce dissipation of energy via friction with the capillary wall and decrease of the kinetic energy. We assume that such dissipative processes take place through the creation of elementary excitation. Let us first describe this process in the reference frame K which, rather confusingly, moves with the same velocity v of the fluid. In this reference frame, the fluid is at rest and its energy is the ground state energy that we denote by E_0 . If a single elementary excitation with a momentum p appears in the fluid, the total energy of the fluid in the reference frame K is $E_0 + \epsilon(p)$, where $\epsilon(p)$ is the energy of the excitation with momentum p. Let us move to the moving frame K' in which the fluid moves with a velocity v but the capillary is at rest. In this moving frame K' which moves with the velocity v with respect to the fluid, the energy and momentum of

the fluid are given by setting $V = -v$ in (4.1). We obtain

$$P' = p + Mv,$$
$$E' = E_0 + \epsilon(p) + pv + \frac{1}{2}Mv^2.$$

The above results indicate that the changes in energy and momentum caused by the appearance of one elementary excitation are $\epsilon(p) + pv$ and p, respectively.

Spontaneous creation of elementary excitations, i.e. energy dissipation, can occur if and only if such a process is energetically favorable. This requires

$$\epsilon(p) + pv < 0.$$

This is satisfied when $|v| > \frac{\epsilon(p)}{|p|}$ and $pv < 0$, i.e. when the elementary excitation has the momentum p opposite to the fluid velocity v and the fluid velocity $|v|$ exceeds the critical value

$$v_{\rm cr} = \min_p \frac{\epsilon(p)}{|p|}. \tag{4.2}$$

Thus, superfluidity will occur only if the critical velocity is strictly positive. In particular, the ideal Bose gas has $\epsilon(p) = p^2$ and thus $v_{\rm cr} = 0$ so it is not superfluid. In particular, the particle-particle interaction is a crucial requirement for the appearance of superfluidity in a bosonic system.

4.2 Bogoliubov approximation

Our starting point is the Bose gas enclosed in a cubic box of length L with periodic boundary conditions. The second quantized Hamiltonian is given by (we do not introduce any scaling of the interaction at this point)

$$H = \sum_p p^2 a_p^* a_p + \frac{1}{2L^3} \sum_{p,k,q} \hat{w}(k) a_{p+k}^* a_{q-k}^* a_p a_q.$$

We are interested in the excitation spectrum of the system, that is the structure of the energies above the ground state energy.

If there were no interaction then we know that the ground state would be given by all particles in the zero momentum mode. The interaction w has the property that it converts a pair of particles with momenta p and q into a pair with momenta $p + k$ and $q - k$. Starting with all N particles having momentum zero, we would first get $(N - 2)$ with momentum zero, together with one pair having momenta k and k. When the potential is applied again we could get two possibilities: one would be two pairs k, k and q, q; the other would be a genuine triplet k, q, r, such that $k + q + r = 0$. But the probability of the former relative to the latter would be $(N - 2)(N - 3)/4$

because there are $(N-1)$ particles with zero momentum and only 2 with nonzero momentum. Applying w over and over again we will ultimately get a finite fraction of triplets, quartets, etc. as well as pairs, but hopefully if the interaction is weak enough we need consider explicitly only pairs in the ground state wave function. This suggests that only terms with pairs should be relevant in the "effective" Hamiltonian.

These were the arguments that led N. N. Bogoliubov to drop all terms in the Hamiltonian involving more than two creation/annihilation operators of a non-zero mode. Doing this we obtain

$$\begin{aligned} H &\approx \frac{\hat{w}(0)}{2L^3} a_0^* a_0^* a_0 a_0 + \sum_{k\neq 0} \left(k^2 + \frac{a_0^* a_0}{L^3} (\hat{w}(k) + \hat{w}(0) \right) a_k^* a_k \\ &+ \sum_{k\neq 0} \frac{\hat{w}(k)}{2L^3} (a_0^* a_0^* a_k a_{-k} + a_k^* a_{-k}^* a_0 a_0) \\ &= \frac{\hat{w}(0)\rho}{2}(N-1) + H_{\mathrm{Bog}} + R \end{aligned}$$

where

$$\rho = N/L^3;$$

$$H_{\mathrm{Bog}} := \sum_{k\neq 0} (k^2 + \rho\hat{w}(k)) a_k^* a_k + \frac{1}{2} \sum_{k\neq 0} \rho\hat{w}(k)(a_k^* a_{-k}^* + a_k a_{-k});$$

$$R := \frac{-\hat{w}(0)}{2L^3}(N-N_0)(N-N_0-1) + \sum_{k\neq 0} \frac{\hat{w}(k)}{2L^3} \left((a_0^* a_0^* - N) a_k a_{-k}) + a_k^* a_{-k}^* (a_0 a_0 - N) \right).$$

We used

$$a_0^* a_0^* a_0 a_0 = N_0(N_0-1) = N(N-1) - 2N_0(N-N_0) - (N-N_0)(N-N_0-1).$$

We argue that R is small, because

$$a_0^* a_0^* \approx a_0 a_0 \approx N_0 \approx N.$$

The last step is called the c-number substitution. Thus the effective Hamiltonian is given by a quadratic operator (quadratic in creation and annihilation operators) that does not conserve the number of particles. It can however be diagonalized. To this end we use a Bogoliubov transformation By introducing

$$b_p = c_p a_p + s_p a_{-p}^*, b_p^* = c_p a_p^* + s_p a_{-p}$$

with appropriate c_p, s_p (see exercises) such that $c_p^2 - s_p^2 = 1$. Then

$$H_{\mathrm{Bog}} = E_{\mathrm{Bog}} + \sum_{p\neq 0} \epsilon(p) b_p^* b_p \tag{4.3}$$

with

$$\epsilon(p) = |p| \sqrt{p^2 + 2\rho\hat{w}(p)}.$$

Problem 4.2.1. *Derive* (4.3)

Note that it follows that

$$\frac{\epsilon(p)}{|p|} = \sqrt{p^2 + 2\rho\hat{w}(p)}$$

and thus the critical velocity $v_{\rm cr} > 0$. Thus Bogoliubov's computations shows that the interacting Bose gas satisfies Landau's criterion for superfluidity.

4.3 Validity of the Bogoliubov approximation

4.3.1 Exciting the Hartree state

We have shown already, that a mean-field Bose gas exhibits Bose–Einstein condensation at zero temperature with the condensate wave-function given by the minimizer of the Hartree functional. Thus the Hartree state plays the role of the macroscopically occupied state that in Bogoliubov's computation was represented by the zero momentum mode. The excited states should therefore correspond to the situation when a particle is outside the Hartree state. To implement this idea rigorously, we will introduce the so-called excited Fock space.

Consider now any (real-valued) orthonormal basis $u_0, u_1, \ldots$ of $L^2(\Omega)$ containing the Hartree minimizer u_0. Then, it is known that $\{u_{i_1} \otimes_{\rm s} u_{i_2} \otimes_{\rm s} \ldots u_{i_N}\}$ is an orthogonal basis of the symmetric space $\mathfrak{H}^N$, where

$$u_1 \otimes_{\rm s} u_2 \ldots \otimes_{\rm s} u_N = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} u_{\sigma(1)}(x_1) \cdots u_{\sigma(N)}(x_N).$$

Let us define

$$\mathfrak{H}_0^N := \mathrm{Span}(\underbrace{u_0 \otimes \ldots \times u_0}_{N \text{ times}})$$

and for $k > 0$

$$\mathfrak{H}_k^N := \mathrm{Span}(\underbrace{u_0 \otimes \ldots \times u_0}_{N-k \text{ times}}) \otimes_{\rm s} \bigotimes_{\rm sym}^{k} \mathfrak{H}_+ = u_0^{\otimes(N-k)} \otimes_{\rm s} \mathfrak{H}_+^k$$

where $\mathfrak{H}_+ = \{u_0\}^\perp = \mathrm{Span}\{u_1, u_2, \ldots\} \in L^2(\Omega)$. It follows that the many-body Hilbert space $\mathfrak{H}^N$ can be written as a direct sum of the form

$$\mathfrak{H}^N = \mathfrak{H}_0^N \oplus \mathfrak{H}_1^N \oplus \cdots \oplus \mathfrak{H}_N^N.$$

Thus, any N-body wave function $\Psi \in \mathfrak{H}^N$ can be written as

$$\Psi = \psi_0 u_0^{\otimes N} + u_0^{\otimes(N-1)} \otimes_{\rm s} \psi_1 + u_0^{\otimes(N-2)} \otimes_{\rm s} \psi_2 + \cdots + \psi_N$$

with $\psi_k \in \mathfrak{H}_+^k$. It is easy to check that

$$\langle u_0^{\otimes(N-k)} \otimes_{\mathrm{s}} \psi_k, u_0^{\otimes(N-l)} \otimes_{\mathrm{s}} \psi_l \rangle_{\mathfrak{H}^N} = \delta_{kl} \langle \psi_k, \psi_l \rangle_{\mathfrak{H}^l}$$

from which it follows that

$$\|\Psi\|^2 = |\psi_0|^2 + \sum_{k=1}^{N} \|\psi_k\|^2_{\mathfrak{H}_+^k}.$$

Therefore we see that the linear map

$$U_N : \mathfrak{H}^N \to \mathcal{F}_+^{\leq N} := \bigoplus_{n=0}^{N} \mathfrak{H}_+^n$$

given by

$$U_N(\Psi) = \psi_0 \oplus \psi_1 \oplus \psi_2 \oplus \ldots \oplus \psi_N$$

is a unitary operator from $\mathfrak{H}^N$ onto the truncated, excited Fock space $\mathcal{F}_+^{\leq N}$. The latter can always be seen as being embedded in the the full excited Fock space given by

$$\mathcal{F}_+ := \bigoplus_{n=0}^{\infty} \mathfrak{H}_+^n.$$

The full Fock space of excited particles appears as the limit of the truncated excited Fock space when $N \to \infty$.

The operator U_N has some important properties. In particular we have that U_N can be written as

$$U_N(\Psi) = \bigoplus_{j=0}^{N} Q^{\otimes j} \left(\frac{a_0^{N-j}}{\sqrt{(N-j)!}} \Psi \right)$$

for all $\Psi \in \mathfrak{H}^N$. Here $Q = 1 - |u_0\rangle\langle u_0|$ is the projection onto the excited space. Similarly

$$U_N^* \left(\bigoplus_{j=0}^{N} \psi_j \right) = \sum_{j=0}^{N} \frac{(a_0^*)^{N-j}}{\sqrt{(N-j)!}} \phi_j$$

for all $\phi_j \in \mathfrak{H}_+^j$. These operators satisfy the following indentities on $\mathcal{F}_+^{\leq N}$:

$$\begin{aligned} U_N a_0^* a_0 U_N^* &= N - \mathcal{N}_+, \\ U_N a^*(f) a_0 U_N^* &= a^*(f) \sqrt{N - \mathcal{N}_+}, \\ U_N a_0^* a(f) U_N^* &= \sqrt{N - \mathcal{N}_+} a(f), \\ U_N a^*(f) a(g) U_N^* &= a^*(f) a(g). \end{aligned} \tag{4.4}$$

Here $\mathcal{N}_+ = \sum_{m \geq 1} a_m^* a_m$ is the operator counting the number of excited particles.

4.3.2 The Bogoliubov Hamiltonian

We define the Bogoliubov Hamiltonian in the following way:

$$\mathbb{H} = \sum_{m,n\geq 1} \langle u_m, (h+K_1)u_n\rangle a_m^* a_n dxdy + \frac{1}{2}\langle u_m \otimes u_n, K_2\rangle a_m^* a_n^* + \frac{1}{2}\langle K_2, u_m \otimes u_n\rangle a_m a_n \tag{4.5}$$

where $K_1 : \mathcal{H}_+ \to \mathcal{H}_+$ and $K_2 : \overline{\mathcal{H}_+} \to \mathcal{H}_+$ are operators defined by

$$\begin{aligned} \langle u, K_1 v\rangle &= \iint_{\Omega\otimes\Omega} \overline{u(x)}v(y)u_0(x)\overline{u_0(y)}w(x-y)dxdy, \\ \langle u, K_2\overline{v}\rangle &= \iint_{\Omega\otimes\Omega} \overline{u(x)v(y)}u_0(x)u_0(y)w(x-y)dxdy \end{aligned} \tag{4.6}$$

for all $u, v \in \mathfrak{H}_+$. Finally, h is the one-body operator given by

$$h = -\Delta + V + |u_0|^2 * w - e_{\mathrm{H}} - \frac{1}{2}\iint |u_0(x)|^2|u_0(y)|^2 w(x-y)$$

which comes from the Hartree equation.

While we shall treat K_1 as a one-body operator, we should really think of K_2 as its integral kernel $K_2(x,y) = (Q\otimes Q)(u_0 \otimes u_0 w(.-.))(x,y)$ which is the two-body function obtained by projecting the symmetric function $u_0(x)u_0(y)w(x-y)$ onto $\mathfrak{H}_+^2$.

4.3.3 Rigorous statement about the Bogoliubov approximation

We shall now make several assumptions that will be used to prove the validity of the Bogoliubov approximation.

A1 - interaction. Assumptions on the interaction potential are the same as in Theorem 3.1.6.

A2 - Hartree theory. We assume that the Hartree minimizer is unique and nondegenerate.

A3 - Bose–Einstein Condensation. We assume that there is BEC on u_0 in the sense that for any constant $R > 0$, there exists a function $\epsilon_R : \mathcal{N} \to [0,\infty)$ with $\lim_{N\to\infty} \epsilon_R = 0$ such that, for any wave function $\Psi_N \in \mathfrak{H}^N$ satisfying $\langle \Psi_N, H_N \Psi_N\rangle \leq E(N) + R$ we have

$$\frac{\langle u_0, \gamma^{(1)}_{\Psi_N} u_0\rangle}{N} \geq 1 - \epsilon_R.$$

Note that during the course we have already showed that these assumptions are satisfied in many situations.

Using assumptions (A1) and (A2) one can prove that

$$c\mathrm{d}\Gamma(h+1) - C \leq \mathbb{H} \leq \mathrm{d}\Gamma(h+C) + C \tag{4.7}$$

for some constants $c, C > 0$. We will now state the main result about Bogoliubov approximation, which heuristically can be written as

$$U_N(H_N - Ne_{\mathrm{H}})U_N^* \to \mathbb{H}.$$

Here it is.

Theorem 4.3.1 (Validity of Bogoliubov approximation)**.** *Under the assumptions (A1)-(A3) the following holds true:*
1 - weak convergence to $\mathbb{H}$***.*** *For any* Φ *and* Φ' *in the (quadratic form) domain of* $\mathbb{H}$ *we have*

$$\lim_{N\to\infty} \langle \Phi', U_N(H_N - Ne_{\mathrm{H}})U_N^*\Phi\rangle = \langle \Phi', \mathbb{H}\Phi\rangle_{\mathcal{F}_+}.$$

2 - convergence of eigenvalues. *Let* $\lambda_i(A)$ *denote the* i*-th eigenvalue of the operator* A*. Then*

$$\lim_{N\to\infty} (\lambda_L(H_N) - Ne_{\mathrm{H}}) = \lambda_L(\mathbb{H}).$$

3 - convergence of eigenvectors. *Let* Ψ_N^L *and* Φ^L *denote the eigenvectors corresponding to the* L*-th eigenvalue of* H_N *and* $\mathbb{H}$ *respectively. Then*

$$\lim_{N\to\infty} U_N\Psi_N^L = \Phi^L$$

and the convergence is strong in $\mathcal{F}_+$*.*

Let us now sketch the main ideas behind the proof of this theorem. Roughly, it contains two main steps. First, one has to compute $U_N(H_N - Ne_{\mathrm{H}})U_N^*$ and rewrite it as

$$U_N(H_N - Ne_{\mathrm{H}})U_N^* = \mathbb{H} + R \tag{4.8}$$

where R will be treated as a rest term.

In the second step one has to show that the rest term is in some sense small. This will then lead to the desired conclusion by the min-max principle, which roughly states states that if self-adjoint operators A and B satisfy $A \leq B$, then the n-th eigenvalue of A is smaller than the n-th eigenvalue of B.

Let us now discuss how to get (4.8). First we write

$$H_N = \sum_{m,n\geq 0} T_{mn}a_m^*a_n + \frac{1}{2(N-1)} \sum_{m,n,p,q\geq 0} W_{mnpq}a_m^*a_n^*a_pa_q$$

where $T_{mn} = \langle u_m, (-\Delta + V)u_n\rangle$ and $W_{mnpq} = \langle u_m \otimes u_n, w(x-y)u_p \otimes u_q\rangle$.

A tedious but straightforward computation using the relations (4.4) shows that

$$U_N(H_N - Ne_{\mathrm{H}})U_N^* = \sum_{j=0}^{4} A_j$$

where

$$
\begin{aligned}
A_0 &= \frac{1}{2} W_{0000} \frac{\mathcal{N}_+(\mathcal{N}_+ + 1)}{N-1}, \\
A_1 &= \sum_{m \geq 1} \left(T_{0m} + W_{000m} \frac{N - \mathcal{N}_+ - 1}{N-1} \right) \sqrt{N - \mathcal{N}_+} a_m + \text{h.c.}, \\
A_2 &= \sum_{m,n \geq 1} T_{mn} a_m^* a_n - (T_{00} + W_{0000}) \mathcal{N}_+ \\
&\quad + \sum_{m,n \geq 1} \langle u_m, (|u_0|^2 * w + K_1) u_n \rangle a_m^* a_n \frac{N - \mathcal{N}_+}{N-1} \\
&\quad + \left(\frac{1}{2} \sum_{m,n \geq 1} \langle u_m \otimes u_n, K_2 \rangle a_n^* a_m^* \frac{\sqrt{(N - \mathcal{N}_+)(N - \mathcal{N}_+ - 1)}}{N-1} + \text{ h.c.} \right), \\
A_3 &= \frac{1}{N-1} \sum_{m,n,p \geq 1} W_{mnp0} a_m^* a_n^* a_p \sqrt{N - \mathcal{N}_+} + \text{ h.c.}, \\
A_4 &= \frac{1}{2(N-1)} \sum_{m,n,p,q \geq 1} W_{mnpq} a_m^* a_n^* a_p a_q.
\end{aligned}
$$

Now we have to identify the Bogoiubov Hamiltonian. This we do by looking at A_2. We see that

$$
\begin{aligned}
A_2 - \mathbb{H} &= \mathrm{d}\Gamma(Q(|u_0|^2 * w)Q + K_1) \frac{1 - \mathcal{N}_+}{N-1} + \\
&\quad + \Re \sum_{m,n \geq 1} \langle u_m \otimes u_n, K_2 \rangle \left(a_n^* a_m^* \left(\frac{\sqrt{(N - \mathcal{N}_+)(N - \mathcal{N}_+ - 1)}}{N-1} - 1 \right)) \right).
\end{aligned}
$$

We will show that when evaluated on a state $\Phi \in \mathcal{F}_+^{\leq M}$, then $A_2 - \mathbb{H}$ can be made small. To this end we will use to following lemma whose proof follows basically from (4.7).

Lemma 4.3.2. *Under the assumptions on the interaction we have the following estimates*

$$
\begin{aligned}
d\Gamma(QTQ) &\leq \frac{1}{1 - \alpha_1} \mathbb{H} + C\mathcal{N}_+ + C, \\
\mathrm{d}\Gamma(Q(|u_0|^2 * w)Q &\leq \frac{\alpha_2}{1 - \alpha_1} \mathbb{H} + C\mathcal{N}_+ + C
\end{aligned}
\tag{4.9}
$$

where $1 > \alpha_1 > 0$ *and* $\alpha_2 > 0$.

Let $\Phi \in \mathcal{F}_+^{\leq M}$. Then

$$
\begin{aligned}
\langle A_2 \rangle_\Phi - \langle \mathbb{H} \rangle_\Phi &= -\left\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w)Q + K_1) \frac{\mathcal{N}_+ - 1}{N-1} \right\rangle_\Phi \\
&\quad + \Re \sum_{m,n \geq 1} \langle u_m \otimes u_n, K_2 \rangle \langle a_n^* a_m^* X \rangle_\Phi
\end{aligned}
\tag{4.10}
$$

where $X = \frac{\sqrt{(N-\mathcal{N}_+)(N-\mathcal{N}_+-1)}}{N-1} - 1$.

Let us rewrite $w = w_+ - w_-$ where $w_\pm \geq 0$. By linearity, using the triangle inequality, we have

$$\begin{aligned}\Big| - \Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w)Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big| \leq \Big|\Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w_+)Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big| \\ + \Big|\Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w_-)Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big|.\end{aligned}$$

Since the convolution of two non-negative functions is a non-negative function and as the multiplication operator by a non-negative function is non-negative operator, we have that

$$d\Gamma(Q(|u_0|^2 * w_\pm)Q) := A_\pm \geq 0$$

is a positive operator that commutes with $\mathcal{N}_+$. It follows that $A_\pm = (A_\pm^* A_\pm)^{1/2}$ and we can write

$$A_\pm \frac{\mathcal{N}_+ - 1}{N-1} = \sqrt{A_\pm^*}\frac{\mathcal{N}_+ - 1}{N-1}\sqrt{A_\pm}.$$

Using the fact that $A_\pm$ (and therefore $\sqrt{A_\pm}$) does not change the number of particles of a given state we have

$$\Big\langle \sqrt{A_\pm^*}\frac{\mathcal{N}_+ - 1}{N-1}\sqrt{A_\pm}\Big\rangle_\Phi \leq \frac{M}{N-1}\langle A_\pm\rangle_\Phi$$

for $\Phi \in \mathcal{F}_+^{\leq M}$. Thus

$$\begin{aligned}\Big| - \Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w)Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big| &\leq \frac{M}{N-1}\Big|\Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * (w_+ + w_-))Q)\Big\rangle_\Phi\Big| \\ &\leq \frac{M}{N-1}\langle \frac{\alpha_2}{1-\alpha_1}\mathbb{H} + C\mathcal{N}_+ + C\rangle_\Phi\end{aligned}$$

where in the last step we used (4.9).

Furthermore, since for any ψ with $\|\psi\| = 1$ we have by the Cauchy-Schwarz inequality

$$|\langle \psi, \pm K_1\psi\rangle| \leq \|K_1\psi\| \leq \|K_1\|$$

it follows that

$$\|K_1\|\mathbb{1} + K_1 \geq 0.$$

Thus, similarly as above we get

$$\begin{aligned}\Big| - \Big\langle \mathrm{d}\Gamma(QK_1Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big| &= \Big|\Big\langle \mathrm{d}\Gamma(Q(\underbrace{\|K_1\|\mathbb{1} + K_1}_{\geq 0} - \|K_1\|\mathbb{1})Q)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi\Big| \\ &\leq \frac{M}{N-1}\Big\langle \mathrm{d}\Gamma(Q(\|K_1\|\mathbb{1} + K_1)Q)\Big\rangle_\Phi\Big| + \frac{M}{N-1}\Big\langle \mathrm{d}\Gamma(Q\|K_1\|\mathbb{1}Q)\Big\rangle_\Phi\Big| \\ &\leq \frac{3M\|K_1\|}{N-1}\langle \mathcal{N}_+\rangle_\Phi.\end{aligned}$$

It follows that for the first term on the RHS of (4.3.3) we get the bound

$$\Big| - \Big\langle \mathrm{d}\Gamma(Q(|u_0|^2 * w)Q + K_1)\frac{\mathcal{N}_+ - 1}{N-1}\Big\rangle_\Phi \Big| \leq \frac{CM}{N-1}\langle \mathbb{H} + \mathcal{N}_+ 1\rangle_\Phi$$

where the constant $C > 0$ is independent of N.

It remains to look at the second term in (4.3.3). Recall $X = \frac{\sqrt{(N-\mathcal{N}_+)(N-\mathcal{N}_+-1)}}{N-1} - 1$. By the Cauchy-Schwarz inequality we have

$$\Big| \sum_{m,n\geq 1} \langle u_m \otimes u_n, K_2\rangle \langle a_n^* a_m^* X\rangle_\Phi \Big| \leq \Big(\sum_{m,n\geq 1} |\langle u_m \otimes u_n, K_2\rangle|^2\Big)^{1/2} \Big(\sum_{m,n\geq 1} |\langle a_n^* a_m^* X\rangle_\Phi|^2\Big)^{1/2}$$
$$\leq \Big(\iint |K_2(x,y)|^2 dxdy\Big)^{1/2} \Big(\sum_{m,n\geq 1} \langle a_n^* a_m^* a_m a_n\rangle_\Phi\Big)^{1/2} \langle X^2\rangle_\Phi^{1/2}.$$

Here we used that K_2 is a Hilbert-Schmidt operator. Now, one can check that

$$X^2 \leq C\frac{(\mathcal{N}_+ + 1)^2}{(N-1)^2}.$$

Furthermore, since $a_n^* a_m^* a_m a_n = a_n^* a_n a_m^* a_m - \delta_{nm} a_n^* a_m$, we get

$$\sum_{m,n\geq 1} \langle a_n^* a_m^* a_m a_n\rangle_\Phi \leq \langle \mathcal{N}_+^2\rangle_\Phi.$$

Altogether this gives

$$\Big| \sum_{m,n\geq 1} \langle u_m \otimes u_n, K_2\rangle \langle a_n^* a_m^* X\rangle_\Phi \Big| \leq C\|K_2\|_2 \langle \mathcal{N}_+^2\rangle_\Phi^{1/2} \left(\frac{(\mathcal{N}_+ + 1)^2}{(N-1)^2}\right)^{1/2} \leq \frac{CM}{N-1}\langle \mathcal{N}_+ + 1\rangle_\Phi.$$

Thus we see that

$$|\langle A_2\rangle_\Phi - \langle \mathbb{H}\rangle_\Phi| \leq\leq \frac{CM}{N-1}\langle \mathbb{H} + \mathcal{N}_+ 1\rangle_\Phi$$

and thus is relatively small as compared with $\mathcal{H}$ if $M \ll N$. Similar estimates can be obtained for other terms that are left from $U_N(H_N - Ne_{\mathrm{H}})U_N^* = \sum_{j=0}^4 A_j$ after removing A_2. Then one can indeed show that the contributions coming from sectors of the Fock space with large particle numbers are indeed small. The details require some more subtle arguments and can be found in the 2015 CPAM paper of Lewin, Nam, Serfaty and Solovej (see also the paper of Grech and Seiringer CMP 2013).

Appendix A

Hilbert spaces and operators

A.1 Operators

Recall that a Hilbert space $\mathcal{H}$ is a vector space endowed with a sesquilinear map $(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ (i.e., a map which is conjugate linear in the first variable and linear in the second) such that $\|\phi\| = \sqrt{(\phi,\phi)}$ defines a norm on $\mathcal{H}$ which makes $\mathcal{H}$ into a complete metric space.

One basic property of Hilbert spaces that we will use is the fact that for any closed subspace $V \subset \mathcal{H}$ there corresponds the orthogonal complement $V^\perp$ such that $V \oplus V^\perp = \mathcal{H}$.

Another one goes under the name of the Riesz representation theorem: to any continuous linear functional $f : \mathcal{H} \to \mathbb{C}$ there is a unique $\psi \in \mathcal{H}$ such that $f(\phi) = (\psi, \phi)$ for all $\phi \in \mathcal{H}$.

We shall always assume that our Hilbert spaces are separable and therefore that they have countable orthonormal bases.

Definition A.1.1. *By an operator (or more precisely densely defined operator) A on a Hilbert space $\mathcal{H}$ we mean a linear map $A : D(A) \to \mathcal{H}$ defined on a dense subspace $D(A) \subset \mathcal{H}$. Dense refers to the fact that the norm closure $\overline{D(A)} = \mathcal{H}$.*

Definition A.1.2. *If A and B are two operators such that $D(A) \subseteq D(B)$ and $A\psi = B\psi$ for all $\psi \in D(A)$ then we write $A \subset B$ and say that B is an extension of A.*

Note that the domain is part of the definition of the operator. In defining operators one often starts with a domain which turns out to be too small and which one then later extends.

Definition A.1.3. *We say that A is a symmetric operator if*

$$(\phi, A\psi) = (A\phi, \psi) \tag{A.1}$$

for all $\psi, \phi \in D(A)$.

It is a fact that (A.1) holds if and only if $(\psi, A\psi) \in \mathbb{R}$ for all $\psi \in D(A)$. This property is of great importance in quantum mechanics as expectations values of observables need to be real.

Definition A.1.4. *An operator A is said to be bounded on the Hilbert space $\mathcal{H}$ if $D(A) = \mathcal{H}$ and A is continuous, which by linearity is equivalent to*

$$\|A\| = \sup_{\phi, \|\phi\|=1} \|A\phi\| < \infty.$$

The number $\|A\|$ is called the norm of the operator A. An operator is said to be unbounded if it is not bounded.

Definition A.1.5. *If A is an operator we define the adjoint A^* of A to be the linear map $A^* : D(A^*) \to \mathcal{H}$ defined on the space*

$$D(A^*) = \{\phi \in \mathcal{H} | \sup_{\psi \in D(A), \|\psi\|=1} |(\phi, A\psi)| < \infty\}$$

and with $A^*\phi$ defined such that

$$(A^*\phi, \psi) = (\phi, A\psi)$$

for all $\psi \in D(A)$. The existence of $A^*\phi$ for $\phi \in D(A^*)$ is ensured by the Riesz representation theorem. If $D(A^*)$ is dense in $\mathcal{H}$ then A^* is an operator on $\mathcal{H}$.

Definition A.1.6. *A bounded operator $K : \mathcal{H} \to \mathcal{H}$ is called compact if and only if for every bounded sequence $\{\phi_n\} \subset \mathcal{H}$, $\{K\phi_n\}$ has a subsequence convergent in $\mathcal{H}$.*

An equivalent definition says, that compact operators are the closure of the set of finite rank operators.

Theorem A.1.1. *Let K be a compact operator on $\mathcal{H}$. Then there exist orthonormal bases $\{u_n\}$ and $\{v_n\}$ of $\mathcal{H}$ and a sequence of real numbers $\{\lambda_n\}$ with $\lambda_n \to_{n\to\infty} 0$ such that*

$$K\phi = \sum_n \lambda_n (u_n, \phi) v_n$$

for all $\phi \in \mathcal{H}$.

Definition A.1.7. *An operator A defined on a subspace $D(A)$ of $\mathcal{H}$ is said to be positive (or positive definite) if $(\psi, A\psi) > 0$ for all non-zero $\psi \in D(A)$. It is said to be positive semi-definite if $(\psi, A\psi) \geq 0$ for all $\psi \in D(A)$. In particular, such operators are symmetric.*

Definition A.1.8. *If A and B are two operators with $D(A) = D(B)$ then we say that A is (strictly) less than B and write $A < B$ if the operator $B - A$ (which is defined on $D(B - A) = D(A) = D(B)$ is a positive definite operator. We write $A \leq B$ if $B - A$ is positive semi-definite.*

We conclude that if K is a positive, compact operator then

$$K = \sum_i \lambda_i |u_i\rangle\langle u_i| \tag{A.2}$$

where now λ_i's can be chosen positive.

Among compact operators two special classes are of particular importance: trace class operators and Hilbert-Schmidt operators.

Definition A.1.9. *Let $\mathcal{H}$ be a separable Hilbert space and $\{\phi_n\}$ its orthonormal basis. Then for any positive operator A we define its trace to be*

$$\operatorname{Tr} A = \sum_n (\phi_n, A\phi_n).$$

It is easy to see that if $0 \le A \le B$ then $\operatorname{Tr} A \le \operatorname{Tr} B$. Also, $\operatorname{Tr}(UAU^{-1}) = \operatorname{Tr} A$ for any unitary operator U. Finally, if A is trace class and B is bounded, then $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$.

Definition A.1.10. *Any (bounded) operator A on $\mathcal{H}$ is called trace class if and only if $\operatorname{Tr}|A| < \infty$. Here $|A| = \sqrt{A^*A}$.*

A trace class operator is compact. It follows from the representation in Theorem A.1.1 that a compact operator is trace class if and only if $\sum_i |\lambda_i| < \infty$. In particular, a positive, trace class operator has a representation of the form (A.2).

Theorem A.1.2. *Let $\|\cdot\|_1$ be defined for trace class operators by $\|A\|_1 = \operatorname{Tr}|A|$. Then the space of trace class operators equipped with the norm $\|\cdot\|_1$ is a Banach space. Furthermore $\|A\| \le \|A\|_1$.*

Let us now mention the other important class of compact operators.

Definition A.1.11. *An operator A on $\mathcal{H}$ is called Hilbert–Schmidt if and only if $\operatorname{Tr} A^*A < \infty$.*

As in the case of trace class operators, a Hilbert–Schmidt operator is compact. As before this leads the condition that a compact operator is Hilbert–Schmidt if and only if $\sum_i \lambda_i^2 < \infty$.

Among Hilbert–Schmidt operators one can introduce the norm $\|A\|_2 = \sqrt{\operatorname{Tr}(A^*A)} = \sqrt{(A,A)_2}$. With this scalar product the space of Hilbert–Schmidt becomes a Hilbert space. We also have $\|A\| \le \|A\|_2 \le \|A\|_1$. Another important fact is the following

Theorem A.1.3. *Let (M,μ) be a measure space and $\mathcal{H} = L^2(M, d\mu)$. Then A is Hilbert–Schmidt if and only if there is a function $K \in L^2(M \times M, d\mu \otimes d\mu)$ with*

$$(Af)(x) = \int K(x,y) f(y) d\mu(y).$$

Moreover,

$$\|A\|_2^2 = \int |K(x,y)|^2 d\mu(x) d\mu(y).$$

A.2 Tensor product of Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. Consider $\mathcal{Z} = \mathcal{H} \times \mathcal{K}$. We form a vector space which has $\mathcal{Z}$ as a basis. To this end we consider the vector space of functions from $\mathcal{Z} \to \mathbb{C}$ and identify (v, w) with the function that takes the value 1 on (v, w) and 0 otherwise. Let $\mathcal{Z}_0$ be the subspace spanned by elements of the form

$$\begin{aligned}
&(u, v_1 + v_2) - (u, v_1) - (u, v_2),\\
&(u_1 + u_2, v) - (u_1, v) - (u_2, v),\\
&(\lambda u, v) - \lambda(u, v),\\
&(u, \lambda v) - \lambda(u, v)
\end{aligned}$$

where $\lambda \in \mathbb{C}$.

Definition A.2.1. *The algebraic tensor product is defined by*

$$\mathcal{H} \otimes^{al} \mathcal{K} = \mathcal{Z}/\mathcal{Z}_0$$

and the image of (v, w) in this quotient is denoted $v \otimes w$ and is called the tensor multiplication.

We introduce the inner product that satisfies

$$(u_1 \otimes v_1, u_2 \otimes v_2)_{\mathcal{H}\otimes^{\mathrm{al}}\mathcal{K}} = (u_1, u_2)_{\mathcal{H}}(v_1, v_2)_{\mathcal{K}}.$$

We set

$$\mathcal{H} \otimes \mathcal{K} := (\mathcal{H} \otimes^{\mathrm{al}} \mathcal{K})^{\mathrm{cpl}}.$$

Thus, $\mathrm{Span}\{u \otimes v | u \in \mathcal{H}, v \in \mathcal{K}\}$ is dense in $\mathcal{H} \otimes \mathcal{K}$. We call the vectors of the form $u \otimes v$ pure tensor products.

If we have an operator A on the Hilbert space $\mathcal{H}$ and an operator B on the Hilbert space $\mathcal{K}$, then we may form the tensor product operator $A \otimes B$ on $\mathcal{H} \otimes \mathcal{K}$ with domain

$$D(A \otimes B) = \mathrm{Span}\{\phi \otimes \psi | \phi \in D(A), \psi \in D(B)\}$$

and acting on pure tensor products as

$$A \otimes B(\phi \otimes \psi) = (A\phi) \otimes (B\psi).$$

The tensor product may in a natural way be extended to more than two Hilbert spaces. In particular, we may for $N = 1, 2, \ldots$ consider the N-fold tensor product $\bigotimes^N \mathcal{H}$ of a Hilbert space $\mathcal{H}$ with itself. On this space we have a natural action of the symmetric group S_N. I.e., if $\sigma \in S_N$, then we have a unitary map $U_\sigma : \bigotimes^N \mathcal{H} \to \bigotimes^N \mathcal{H}$ defined uniquely by the following action on the pure tensor products

$$U_\sigma u_1 \otimes \ldots \otimes u_N = u_{\sigma^{-1}(1)} \otimes \ldots \otimes u_{\sigma^{-1}(N)}.$$

We shall denote by $\mathrm{Ex} : \bigotimes^N \mathcal{H} \to \bigotimes^N \mathcal{H}$ the unitary corresponding to a simple interchange of the two tensor factors.

Definition A.2.2. *Let P_+ be the orthogonal projection given by*

$$P_+ = (N!)^{-1} \sum_{\sigma \in S_N} U_\sigma.$$

We define the symmetric tensor of Hilbert spaces as

$$\bigotimes_{sym}^{N} \mathcal{H} = P_+ (\bigotimes^{N} \mathcal{H}).$$

Recall that in (2.1) we defined the Hamiltonian that in a more general notation can be written as

$$H_N = \sum_{j=1}^{N} h_j + \sum_{1 \leqslant j < k \leqslant N} w_{jk}.$$

The one-body operators h_j should be identified with

$$h_j = \mathbb{1} \otimes \ldots \otimes \mathbb{1} \otimes \underbrace{h_j}_{j\text{th slot}} \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1}.$$

The operator

$$H_N^{\text{n-in}} = h_1 + \ldots + h_N$$

may then be defined on the domain

$$D(H_N^{\text{n-in}}) = \text{Span}\{\phi_1 \otimes \ldots \otimes \phi_N | \phi_1 \in D(h_1), \ldots, \phi_N \in D(h_N)\}$$

which is dense in $\bigotimes^N \mathcal{H}$ if $D(h_i)$ is dense in $\mathcal{H}$ for each i.

Similarly, the two-body operator w_{kj} is acting on the Hilbert space $\mathcal{H}_j \otimes \mathcal{H}_k$ and in the context of the many-body Hamiltonian w_{jk} has to be understood as being tensored with identities on the remaining variables.

Appendix B

Fock spaces and second quantization

B.1 Fock space

Recall that we introduced the Fock space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathfrak{H}^N.$$

For $\Psi, \Phi \in \mathcal{F}$ we define the inner product to be

$$(\Psi, \Phi)_{\mathcal{F}} = \sum_{n \geq 0} (\psi^{(n)}, \phi^{(n)})_{\mathfrak{H}^n}.$$

For $\Psi \in \mathcal{F}$ we denote by $\|\Psi\|$ the corresponding norm given by

$$\|\Psi\|_{\mathcal{F}}^2 = \sum_{n \geq 0} \|\psi^{(n)}\|_2^2.$$

The Fock space allows to describe states of the system where the number of particles is not fixed. A normalized vector $\Psi = \{\psi^{(n)}\}_{n \geq 0} \in \mathcal{F}$ describes a state that with probability $\|\psi^{(n)}\|_2^2$ has n particles. In particular, states with exactly N particles are embedded in the Fock space; they are described by vectors of the form $\{0, \ldots, 0, \psi_N, 0, \ldots\} \in \mathcal{F}$ having only one non-zero component. These vectors are eigenvectors of the particles number operator $\mathcal{N}$ defined by

$$(\mathcal{N}\Psi)^{(n)} = n\psi^{(n)}$$

for any $\Psi = \{\psi^{(n)}\}_{n \geq 0}$ such that

$$\sum_{n \geq 0} n^2 \|\psi^{(n)}\|_2^2 < \infty.$$

The vacuum vector $\Omega = \{1, 0, 0, \ldots\}$ plays a special role; it is an eigenvector of $\mathcal{N}$ with eigenvalue zero and describes a state with no particles.

B.2 Creation and annihilation operators

For any one-particle wave-function $f \in L^2(\mathbb{R}^3)$ we define the creation operator $a^*(f)$ and the annihilation operator $a(f)$ by setting

$$(a^*(f)\Psi)^{(n)}(x_1,\dots,x_n) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} f(x_j)\psi^{(n-1)}(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n),$$

$$(a(f)\Psi)^{(n)}(x_1,\dots,x_n) = \sqrt{n+1}\int \overline{f(x_{n+1})}\psi^{(n+1)}(x_1,\dots,x_n,x_{n+1})dx_{n+1}$$

In the case of the creation operator we have to assume $n \geq 1$. To get a physically consistent interpretation, in the definition we need to assume that the vacuum ($n = 0$) component of the Fock space vector is zero after the action of the creation operator. Also,

$$a(f)\Omega = 0.$$

Creation and annihilation operators satisfy canonical commutation relations (CCR)

$$[a(f), a(g)] = [a^*(f), a^*(g)] = 0, \quad [a(f), a^*(g)] = \langle f, g\rangle, \quad \forall f, g \in L^2(\mathbb{R}^3). \tag{B.1}$$

Problem B.2.1. *Check* (B.1).

It is also convenient to introduce operator-valued distributions a_x^* and a_x so that

$$a^*(f) = \int_{\mathbb{R}^3} f(x)a_x^* dx, \quad a(f) = \int_{\mathbb{R}^3} \overline{f(x)}a_x dx, \quad \forall f \in L^2(\mathbb{R}^3). \tag{B.2}$$

Expressed through these operator-valued distributions, the CCR take the form

$$[a_x^*, a_y^*] = [a_x, a_y] = 0, \quad [a_x, a_y^*] = \delta(x-y), \quad \forall x, y \in \mathbb{R}^3.$$

The interpretation is straightforward: $a^*(f)$ creates a new particle with wave function f, while $a(f)$ annihilates such a particle. Creation and annihilation operators are closed densily defined operators on $\mathcal{F}$; moreover, $a^*(f)$ is the formal adjoint of $a(f)$ (as the notation suggests). Notice also that $a^*(f)$ is linear while $a(f)$ is antilinear.

Problem B.2.2. *Check that $a^*(f)$ and $a(f)$ are formal adjoints, i.e.*

$$(\Psi, a^*(f)\Phi)_{\mathcal{F}} = (a(f)\Psi, \Phi)_{\mathcal{F}}$$

for $\Phi, \Psi \in \mathcal{F}$ and $f \in L^2(\mathbb{R}^3)$.

Despite the bosonic creation and annihilation operators being unbounded, usually domain questions are unproblematic since one can take the domain of a sufficiently large power of the number operator to easily make sense of most expressions - see (B.5). One has to be more careful in this respect when working with the operator-valued distributions as formally a_x is obtained form $a(f)$ by taking f to be the Dirac δ-function which is not an element of the one-body Hilbert space.

Notice that if $\Psi = \{\psi^{(n)}\}_{n\geq 0}$, then

$$(a_x\Psi)^{(n)}(x_1,\ldots,x_n) = \sqrt{n+1}\psi^{(n+1)}(x,x_1,\ldots,x_n). \tag{B.3}$$

Hence

$$\begin{aligned}(\Psi,\mathcal{N}\Phi) &= \sum_{n\geq 0} n\int dx_1\ldots dx_n\overline{\psi}^{(n)}(x_1,\ldots,x_n)\phi^{(n)}(x_1,\ldots,x_n)\\ &= \sum_{n\geq 1}\int dx dx_1\ldots dx_{n-1}\overline{(a_x\Psi)}^{(n-1)}(x_1,\ldots,x_{n-1})(a_x\Phi)^{(n-1)}(x_1,\ldots,x_{n-1})\\ &= \int dx(a_x\Psi,a_x\Phi).\end{aligned}$$

It follows that as a quadratic form we can write

$$\mathcal{N} = \int dx\, a_x^* a_x. \tag{B.4}$$

This expression for $\mathcal{N}$ suggests that, although creation and annihilation operators are unbounded operators, they can be bounded with respect to the square root of the number of particles operator.

Indeed, in norm we have the bounds

$$\begin{aligned}\|a(f)\Psi\|_{\mathcal{F}} &\leq \|f\|\|\mathcal{N}^{1/2}\Psi\|_{\mathcal{F}},\\ \|a^*(f)\Psi\|_{\mathcal{F}} &\leq \|f\|\|(\mathcal{N}+1)^{1/2}\Psi\|_{\mathcal{F}}\end{aligned} \tag{B.5}$$

for any $f\in L^2(\mathbb{R}^3)$. To prove the first bound, we observe that by the Cauchy–Schwarz inequality

$$\|a(f)\Psi\|_{\mathcal{F}} \leq \int dx|f(x)|\|a_x\Psi\|_{\mathcal{F}} \leq \|f\|\left(\int dx\|a_x\Psi\|_{\mathcal{F}}^2\right)^{1/2} = \|f\|\|\mathcal{N}^{1/2}\Psi\|.$$

The second estimate follows from the first one and the canonical commutation relations.

B.3 Second quntization of operators

Similarly to the expression (B.4), we can express the second quantization of any one-particle operator in terms of the operator-valued distributions

a_x^* and a_x. Let $J^{(1)}$ be an operator on the one-particle space $L^2(\mathbb{R}^3)$. The second quantization of $J^{(1)}$ is the operator $\mathrm{d}\Gamma(J^{(1)})$ on $\mathcal{F}$ defined by the requirement that

$$(\mathrm{d}\Gamma(J^{(1)})\Psi)^{(n)} = \sum_{i=1}^{n} J_i^{(1)}\psi^{(n)}$$

where $J_i^{(1)}$ denotes the operator acting on $L^2(\mathbb{R}^{3n})$ as $J^{(1)}$ on the i-th particle and as the identity on the other $(n-1)$ particles. If the one-particle operator $J^{(1)}$ has the integral kernel $J^{(1)}(x;y)$, we can write

$$\begin{aligned}
(\Phi,\mathrm{d}\Gamma(J^{(1))}\Psi) &= \sum_{n\geq 1}\sum_{j=1}^{n}(\phi^{(n)}, J_j^{(1)}\psi^{(n)}) \\
&= \sum_{n\geq 1} n\int dxdydx_2\ldots dx_n\overline{\phi^{(n)}(x,x_2,\ldots,x_n)}J^{(1)}(x;y)\psi^{(n)}(y,x_2,\ldots,x_n) \\
&= \sum_{n\geq 1}\int dxdydx_2\ldots dx_n J^{(1)}(x;y)\overline{(a_x\Phi)^{(n-1)}(x_2,\ldots,x_n)}(a_y\Psi)^{(n-1)}(x_2,\ldots,x_n) \\
&= \int dxdyJ^{(1)}(x;y)(a_x\Phi, a_y\Psi).
\end{aligned}$$

Thus, as a quadratic form, we have

$$\mathrm{d}\Gamma(J^{(1)}) = \int dxdyJ^{(1)}(x;y)a_x^*a_y.$$

Problem B.3.1. *Let $J^{(2)}$ be an operator on the two-particle Hilbert space $L^2_{sym}(\mathbb{R}^6)$. We define the second quantization $\mathrm{d}\Gamma(J^{(2)})$ of $J^{(2)}$ by*

$$(\mathrm{d}\Gamma(J^{(2)})\Psi)^{(n)} = \sum_{\{i_1,i_2\}} J^{(2)}_{\{i_1,i_2\}}\psi^{(n)}$$

where the sum runs over all sets $\{i_1,i_2\}$ of two different indices in $\{1,\ldots,n\}$ and where $J^{(2)}_{\{i_1,i_2\}}$ denotes the operator on $L^2_{sym}(\mathbb{R}^{3n})$ acting as $J^{(2)}$ on the variables i_1,i_2 and as identity on the other $(n-2)$ variables. Show that if $J^{(2)}$ has the integral kernel $J^{(2)}(x_1,x_2;y_1,y_2)$ then

$$\mathrm{d}\Gamma(J^{(2)}) = \int dx_1dx_2dy_1dy_2J^{(2)}(x_1,x_2;y_1,y_2)a_{x_1}^*a_{x_2}^*a_{y_1}a_{y_2}.$$

Problem B.3.2. *Consider the Hamiltonian*

$$H_N = \sum_{i=1}^{N} -\Delta_i + \sum_{i<j} w(x_i - x_j).$$

Show that its second quantization equals

$$H = \int dx\ \nabla_x a_x^*\nabla_x a_x + \frac{1}{2}\int w(x-y)a_x^*a_y^*a_xa_y\ dxdy.$$

Problem B.3.3. *Consider the many-body Hamiltonian of the form*

$$H_N = \sum_{i=1}^{N} T_i + \sum_{i<j} W_{ij}.$$

Let $\{u_n\}$ be an orthonormal basis of the one-body Hilbert space. Show that the second quantization of H_N equals

$$H = \sum_{m,n} (u_m, Tu_n) a_m^* a_n + \frac{1}{2} \sum_{k,l,m,n} (u_k \otimes u_l, W u_m \otimes u_n) a_k^* a_l^* a_m a_n$$

where $a_m^* = a^*(u_m)$.

Problem B.3.4. *Consider the operator on the Fock space of the form*

$$A = \sum_i e_i a_i^* a_i.$$

Show that the spectrum of A is of the form

$$\sum_I e_i n_i, \qquad n_i \in \{0, 1, 2, \ldots\}.$$

Bibliography

[1] Kapitza, P. (1938). "Viscosity of Liquid Helium Below the λ-Point". Nature 141 (3558)

[2] F. London (1938). "The λ-Phenomenon of Liquid Helium and the Bose-Einstein Degeneracy". Nature. 141 (3571)

[3] E. H. Lieb and R. Seiringer, Proof of Bose-Einstein Condensation for Dilute Trapped Gases, Phys. Rev. Lett., 88 (2002), p. 170409.