International School for Advanced Studies

Mathematical Physics Sector

October 2007

Ph.D. thesis

Bose-Einstein condensation: analysis of problems and rigorous results

by

Alessandro Michelangeli

supervisor

prof. Gianfausto Dell'Antonio

external referee

prof. Jakob Yngvason

Contents

Introduction and outline

This work is an up-to-date and partially improved analysis of the basics of the mathematical description of Bose-Einstein condensation in terms of first principles of Quantum Mechanics.

The main aims are:

① to provide a compact, yet coherent overview

- ✓ of the basic mathematical tools used to formalise Bose-Einstein condensation,
- ✓ of the mathematical techniques for studying several features of this physical phenomenon,
- ✓ and of how such means emerge as the natural ones in connection with their physical interpretation;

② to discuss and to place in the above perspective some new contributions and improvements

- ✓ on equivalent characterizations of Bose-Einstein condensation,
- ✓ on the strength of the convergence for some currently available asymptotic results,
- ✓ and on the effects of interparticle correlations on the energy and the dynamics of the many-body condensate.

Framework

As a general feature of the Physics of large (many-body) systems, mathematically precise statements about the implications of the equations of Quantum Mechanics *have to be* the necessary counterpart of formal derivations, perturbative approximations and numerical treatments, which come all three into the game in the absence of exact solutions to the equations, due to the overwhelming complexity of the problem. Indeed the validity of a formal as well as a perturbative approach often remains to be studied and rigorously proved and the feasibility of numerical solutions is heavily limited by the computational power of even modern computers.

In this perspective, we will be concerned with the rigorous description of a large number of bosons of certain chemical species, mutually interacting via a purely repulsive pair interaction of short range, confined in some container of macroscopic size, at a suitably high dilution and low temperature. These are the systems that can be actually considered in the experiments nowadays, after a long-lasting theoretical investigation,

and which exhibit bizarre properties that go under the name of BOSE EINSTEIN CONDENSATION (B.E.C.): not only they remain in a gaseous phase at all temperatures down to absolute zero, but also they appear *as if* almost all particles behave as one, condensing in a condensed cloud (possibly surrounded by a thermal cloud) where, quantum mechanically, particles are in the same one-body quantum state.

This is a purely quantum phenomenon that does not find any comprehension within a classical picture and this makes B.E.C. one of the most striking evidences of Quantum Mechanics. In fact, B.E.C. is a consequence of quantum statistics only, namely of undistinguishability of identical bosons. In the simple case of an ideal quantum gas, all thermodynamic properties of the system can be computed explicitly and the condensation mechanism can be clearly depicted and interpreted. In nature, however, particles exhibit forces on each other and these interactions among the particles complicate the theory immensely. In the presence of interaction, B.E.C. is far from being completely understood, although one has a clear theoretical evidence of how it originates. Hence, the comprehension in terms of first principles of all B.E.C. features showing up in the most recent experiments is a major challenge from a Mathematical Physics viewpoint.

A new and reach mainstream in the mathematical analysis of B.E.C. and related topics is boomed in the last decade, after the first experimental observations in 1995. It is remarkable, instead, that the theoretical discover of this phenomenon dates back to the 1920s, even before the full establishment of Quantum Mechanics, while the mathematical structure, before the current new flurry of interest, had not being improving significantly since the semi-rigorous treatment of the 1950s and 1960s.

In this framework, a number of rigorous (and beautiful) results is already established in the literature. They cover an amount of connected fields, such as

- ✓ the stationary description of the ground state and the ground state energy of a dilute Bose gas, both with short-range and with Coulomb interactions ('charged' gas),
- ✓ the Bose-Einstein quantum phase transition in optical lattice models,
- ✓ the emergence of vortices in rapidly rotating condensates and their characteristic pattern
- ✓ one- and two-dimensional limits of the three-dimensional theory,
- ✓ the dynamics of a Bose gas freely expanding from an originally confined and condensed phase.

Altogether, they constitute a still now incomplete description – a general proof of Bose-Einstein condensation for interacting gases still eludes us – and they all are asymptotic results in some limit of infinite number of particles, for which a full control of the errors is still lacking.

In the present work, our interest will be limited to the basics of such a rigorous description:

- ① mathematical and physical well-posedness of the definition of B.E.C. and some equivalent versions,
- ② nature of the customarily adopted scaling limits,

③ control of the convergence in the corresponding asymptotic results,

④ emergence and persistence in time of correlations in the many-body condensate.

The physical picture will be that of Bose gases at zero temperature, i.e., in their ground state. In real experiments the temperature is not exactly zero, of course, but low enough to allow for a ground state description.

These topics, stemming from Condensed Matter Physics, Statistical Physics, and Kinetic Theory, from the mathematical point of view are functional analytic problems involving both standard mathematical techniques of advanced Functional Analysis, Calculus of Variations, and Operator Theory, and *ad hoc* techniques specifically developed in this context (generalized Poincaré inequalities, hard-to-soft potential transformation by partially sacrificing the kinetic energy, localization of the energy, a priori estimates on higher power of the energy with cut-off techniques, diagrammatic control of Duhamel-like expansions, just to mention some of them).

Our investigation and contributions place themselves after the analysis and the results of the following contributors: Lieb, Seiringer, Solovej, and Yngvason, for the time-independent picture, and Elgart, Erdős, Schlein, and Yau, and Adami, Mauser, Golse, and Teta, for the time-dependent one. These, in turn, rely on a vast landscape of related researches which will be thoroughly recalled.

Synopsis

The material is organized as follows.

Chapter 1 provides a short, self-consistent description of the essential physical features of B.E.C. as they have been consolidated through a 80-years-long investigation. The purpose here is to focus on the theoretical and experimental basis of the concept of B.E.C. as a macroscopic occupation of a single-particle quantum state.

Chapter 2 is, in turn, a mathematical detour on the properties of the crucial and ubiquitous tool throughout this work: the reduced density matrix. Physically, it encodes the description of a subsystem of the original system one starts with. The standard notions of kernels, trace, partial trace are here reviewed, together with some other useful results for the following.

Chapter 3 presents the rigorous formalisation of B.E.C., from the mathematics to describe large many-body boson systems, to the emergence of the reduced density matrices as the natural tool, and the conceptual and technical need of a limit of infinite number of particles. The outcoming notion of B.E.C. will be, from this moment on, an intrinsically asymptotic notion, accounting for the convergence of the one-body reduced density matrix to the projection onto the one-body condensate wave function. Some distinct natural topologies are here shown to be equivalent for controlling such a convergence. To see how fertile this framework is, we report the main achievements it has recently led to: the rigorous derivation of the stationary Gross-Pitaevskiĭ theory for a dilute Bose gas and the proof of existence of asymptotic B.E.C. for a model of bosons paired by a short-range and hard-core interaction.

Chapter 4 puts the emphasis on the large-size limits used in studying many-body systems. It is already well known that these are scaling limits, namely where the interaction is scaled with the number of particles according to some prescriptions, yet the subject deserves further consideration. By scaling, one handles a more and more populated system which shares some physical features with the original one. The collection of these features identifies the scaling of interest. In particular, a scaling limit is not simply a thermodynamic limit. Any asymptotic result on B.E.C., as well as its definition itself, depends actually on the scaling adopted to derive it, namely under which physical regime it is derived, and should be accompanied by a control of the error terms in the asymptotics, which is currently a major open issue. In particular, the role of the Gross-Pitaevskiĭ scaling is reviewed, as *the* limit of ultra-high dilution which still enables one to study the dynamics of the gas.

Chapter 5 is centred on the issue of the persistence in time of the condensed phase, after the gas is released from the trap that was confining it. The underlying mathematical problem is the derivation of the time-dependent non-linear equation for the one-body condensate wave function, starting from the linear many-body dynamics. A formal derivation, rather commonly accepted as satisfactory, together with the fit of the experimental data, show that this equation is the cubic non-linear Schrödinger equation known as the Gross-Pitaevskiĭ equation. The path towards a rigorous derivation, instead, snakes around a much longer sequence of intermediate achievements, within a framework of classical and quantum kinetic equations which is even more general than our B.E.C.-related problem. Such a derivation method involves the study of an infinite hierarchy of evolutionary equations for reduced density matrices, which is eventually solved by all tensor powers of the projections onto the solution of the Gross-Pitaevskiĭ equation. Physically, this corresponds to describe the time evolution of all finite portions of the system of interest, while letting it enlarge to infinity in some scaling limit. Both the mathematical scheme of this method and the previous achievements are reviewed, until the most recent and somehow conclusive results currently known. These are obtained within the same conceptual scheme, but through different techniques: the core contribution of this chapter is to provide some strengthened unified version for the convergence of reduced density matrices, lifting it to a trace norm convergence at any fixed time.

Chapter 6 touches the problem of determining interparticle correlations establishing in the true many-body state as a consequence of the interaction. If this state undergoes B.E.C., correlations shows up in a peculiar form and one can monitor their dramatic influence on the energy and the dynamics of the condensate. The net effect is the presence of a typical short scale correlation structure built up (and preserved in time) by the two-body process. This is actually a crucial issue in understanding the properties of a condensate. In an asymptotic analysis where, in the limit $N \to \infty$, marginals factorise and correlations disappear, a strategy is needed, both at the many-body and at the one-body level, to mime the true time evolution of the system through some suitable trial state.

Chapter 7 deals with a number of similar characterizations of B.E.C. that one usually ends up with because they are easier to handle, for technical reasons. The substantial equivalence of all them is proved and discussed. First, it is pointed out that

the asymptotic convergence of the one-body reduced density matrix to a rank-one projection is the same as the corresponding factorisation for any fixed number of variables. After that, a couple of typical asymptotic factorisations of the many-body state are proved to be equivalent to the standard definition of B.E.C. in terms of reduced density matrices. This provides some homogeneity in the hypotheses of several theorems of time-stability of condensation. All these mathematical characterizations share the same physical interpretation of a condensate where essentially all particles behave as one.

Chapter 8 collects some concluding remarks and outlines some open mathematical problems related. Some of them are are just improvements with respect to the currently known results, where certain non-trivial simplifications are believed to be relaxable. Some others are indeed major open problems, even on basics of rigorous B.E.C. Among them, a control of the errors in the asymptotic rigorous results, a quantum statistical treatment which includes the temperature, a deeper insight in the interparticle correlations, a full treatment of true delta interactions at least in the one-dimensional case.

Chapter 1

Physical preliminaries

1.1 Bose-Einstein condensation theoretically

Bose-Einstein condensation is a familiar and rather well understood quantum phenomenon that physicists have been dealing with for the last 80 years. One can refer to a number of comprehensive references on the whole subject, which discuss the theoretical physical background that will be understood throughout in this work. Our choice is unavoidably towards the most recent and complete of them:

- ✓ Castin [26] (2000),
- ✓ Leggett [71] (2001),
- ✓ Pethick and Smith [96] (2002),
- ✓ Pitaevskiĭ and Stringari [97] (2003),
- ✓ Yukalov [111] (2004).

Not surprisingly, these are not Mathematical Physics references. Rigorous analysis on B.E.C. has received a new flurry of interest in the last decade, parallel to the experimental realizations that finally have been possible and to the comprehension of the most appropriate mathematical tools to investigate the subject. One could say that while from the theoretical Condensed Matter Physics perspective B.E.C. realizations have represented the apex of huge experimental efforts and the confirm of a long theoretical research, culminating with the 2001 Nobel prize in Physics to Cornell, Ketterle, and Wieman [31, 63], from the Mathematical Physics viewpoint they have stressed the need of a fully rigorous comprehension of the phenomenology in terms of first principles.

It is difficult to forsee whether this recent and rich mathematical mainstream will advance towards the understanding of the newest related phenomena (rotating condensates, quantized vortices, B.E.C./B.C.S. transitions, optical lattices, coherence phenomena, interference and Josephson effects, just to mention some of them). For instance, temperature itself has still to enter in many of these rigorous treatments, which are set at $T = 0$.

Instead, what can be emphasized now, and presumably for the future, is the prolific interaction between this so-called "Mathematics of the Bose gas" and the Physics behind. The former is developing new advanced tools combining their own abstract interest with their physical applications. The latter provides true experimental data, as

Figure 1.1. Cover of *Science* magazine of December 22, 1995, declaring the Bose condensate as the "molecule of the year".

well as prescriptions and ansatz for the expected mathematical results, and benefits from a comprehension and a predictive description on a rigorous basis.

The standard few-words description of B.E.C. is

> the macroscopic occupation of the same single-particle state in a many-body system of bosons.

This is illustrated by the cover of *Science* magazine of December 22, 1995, in which the Bose condensate is declared "molecule of the year" and pictured as a platoon of soldiers marching in lockstep: each particle in the condensate shares a quantum mechanical wave function and so they all move as one; particles outside the condensate move faster and in all directions (see Fig. 1.1). Actually this picture is misleading, rigorously speaking: the condensate is described by an essentially factorised many-body wave function $\Psi_N \sim \varphi^{\otimes N}$ which is interpreted as the occupation of the common one-body state φ. This, in turn, extends to the whole region: one cannot distinguish among distinct φ's marching close together.

The ubiquitous and celebrated Fig. 1.2 gives the typical qualitative explanation of this phenomenon. It provides at least a semi-quantitative prediction: occurrence of B.E.C., as a consequence of undistinguishability, namely, of quantum statistics, at a temperature and a dilution such that the thermal (de Broglie) wavelength becomes comparable to the mean interparticle distance. This is in contrast with other phase transitions (like melting or crystallization), which depend on the interparticle interactions.

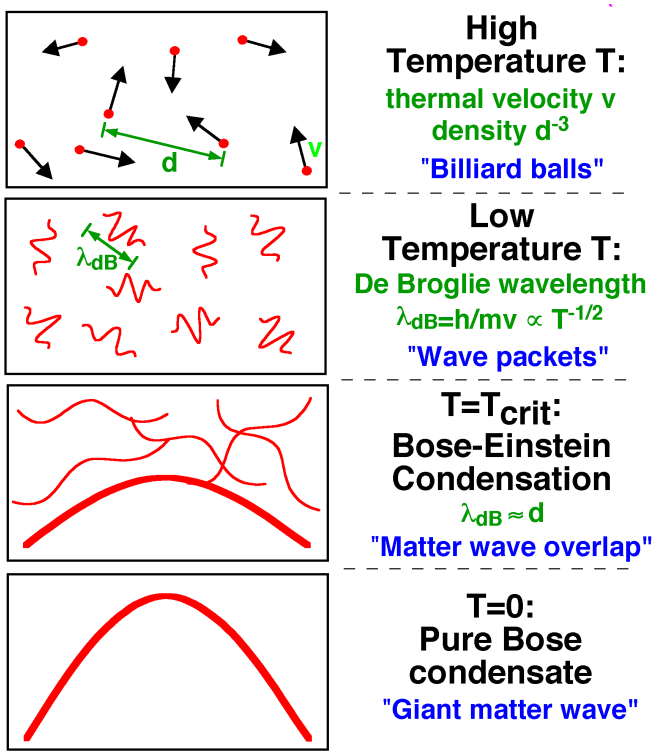


Figure 1.2. Semi-qualitative explanation for Bose-Einstein condensation

Ideal gas picture

A first analytical explanation of such a mechanism is customarily done for an ideal Bose gas, that is, N undistinguishable non-interacting non-relativistic quantum particles, which for concreteness are assumed to be in a box Ω with volume $|\Omega|$. To treat the most natural scenario, *hereafter in this section, a three-dimensional setting will be assumed.*

The emergence of this phenomenon, rigorously speaking, requires a thermodynamic limit. In the grand-canonical description where the system is put in thermal equilibrium at temperature T with a reservoir with chemical potential μ, as a consequence of the statistics of bosons one finds that the total occupation number is

$$\overline{N} = \sum_{i\geqslant 0} \frac{1}{e^{\beta(\varepsilon_i-\mu)}-1} \equiv \overline{N}_0 + \overline{N}_{\text{therm}} \tag{1.1}$$

$$\overline{N}_0 := \frac{1}{e^{\beta(\varepsilon_0-\mu)}-1}\,, \tag{1.2}$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature in terms of the Boltzmann's constant k_B and ε_i is the energy of the i-th excited state of the single-particle Hamiltonian

$$H^{(1)} = \frac{\mathbf{p}^2}{2m}\,, \tag{1.3}$$

while the Hamiltonian of the many-body system is

$$H = \sum_k H_k^{(1)} = \sum_k \frac{\mathbf{p}_k^2}{2m}\,. \tag{1.4}$$

Here each $\mathbf{p}_k$ is the quantum operator $\mathbf{p}_k = -i\,\hbar\nabla_{\mathbf{r}_k}$ acting on the one-body Hilbert space $L^2(\Omega)$ with appropriate boundary conditions. As usual, $\hbar$ and m are the Plank's constant and the mass of each particle, respectively. For instance, the well known eigenvalues of $H^{(1)}$ for a three-dimensional cubic box of side length L with periodic boundary conditions are

$$\varepsilon_{\mathbf{n}} = \frac{(2\pi\hbar)^2}{2mL^2}\mathbf{n}^2\,, \qquad \mathbf{n} \in \mathbb{N}^3\,. \tag{1.5}$$

Thus, one distinguishes among the occupancy $\overline{N}_0$ of the ground state and the remaining thermal component $\overline{N}_{\text{therm}}$, and in the thermodynamic limit $\varepsilon_0 \to 0$.

For *fixed* $\mu < 0$, in the thermodynamic limit,

$$\frac{\overline{N}}{|\Omega|} \xrightarrow{|\Omega|\to\infty} \Big(\frac{2\pi}{\hbar}\Big)^3 \int_{\mathbb{R}^3} \frac{1}{e^{\beta(\frac{\mathbf{p}^2}{2m}-\mu)}-1}\,\mathrm{d}\mathbf{p} \;=:\; \rho \tag{1.6}$$

which is monotonously increasing with μ and bounded, as $\mu \to 0$, by

$$\rho_{\text{crit}} \;:=\; g_{3/2}(1)\Big(\frac{m}{2\pi\hbar^2\beta}\Big)^{3/2}. \tag{1.7}$$

Here

$$g_{3/2}(z) := \frac{2}{\sqrt{\pi}}\int_0^{+\infty}\frac{\sqrt{x}}{z^{-1}\,e^x-1}\,\mathrm{d}x = \sum_{\ell=1}^{\infty}\frac{z^\ell}{\ell^{3/2}}\,, \tag{1.8}$$

whence $g_{3/2}(1) \simeq 2.612$. Thus, *condensation arises* as the phenomenon where particles exceeding the above critical number all go into the lowest energy state. To fix the density at some value larger that ρ_{crit} one has to simultaneously let $|\Omega| \to \infty$ and $\mu \to 0$. Limit (1.6) can be controlled separating the contribution from the lowest energy level and approximating the contribution from the remaining terms by an integral, and the net result is

$$\rho = \rho_{\text{cond}} + \rho_{\text{crit}} \qquad\qquad (\rho > \rho_{\text{crit}}) \tag{1.9}$$

where

$$\rho_{\text{cond}} \;:=\; \lim_{|\Omega|\to\infty}\frac{\overline{N}_0}{|\Omega|} \tag{1.10}$$

is the DENSITY OF THE CONDENSATE.

Transition between the non condensed regime and B.E.C. corresponds to the transition between a zero value and a positive value for ρ_{cond}, hence, it is determined by the condition $\rho_{\text{crit}} = \rho$ in (1.7), which leads to the CRITICAL TEMPERATURE

$$T_c = \frac{2\pi\hbar^2}{mk_B}\Big(\frac{\rho}{g_{3/2}(1)}\Big)^{2/3}. \tag{1.11}$$

This corresponds to

$$\lim_{L\to\infty}\frac{1}{L^3}\frac{1}{e^{\beta(\varepsilon_0(L)-\mu(L))}-1} = 0 \tag{1.12}$$

where L is the length of the side of the box and the dependence $\mu = \mu(L)$ is determined by (1.1), writing $\overline{N} = \rho\,L^3$ with fixed ρ. The critical value T_c is the *smallest* T_c for which (1.12) is true. Condensation occurs for $T \leqslant T_c$. One usually restates this criterion in the much more familiar form

$$\rho\,\lambda_{\text{dB}}^3 = g_{3/2}(1) \simeq 2.612 \tag{1.13}$$

where $\lambda_{\rm dB}$ is the thermal (or de Broglie) wavelength

$$\lambda_{\rm dB} := \sqrt{\frac{2\pi\hbar^2}{mk_BT}}\,. \tag{1.14}$$

Condensation occurs when

$$\rho\,\lambda_{\rm dB}^3 \gtrsim 2.612\,, \tag{1.15}$$

that is, below the critical value of the temperature such that the thermal wavelength is of the same order of the mean interparticle distance $\rho^{-1/3}$. In the same picture one proves that the ground state occupancy in the condensed regime is

$$\frac{\overline{N}_0}{\overline{N}} = 1 - \Big(\frac{T}{T_c}\Big)^{3/2} \qquad (T \leqslant T_c)\,. \tag{1.16}$$

This way, one recovers the interpretation of Figure 1.2.

Among other peculiar features, the condensate phase turns out to have infinite compressibility and zero entropy and one has a first order gas/condensate phase transition.

In particular, at zero temperature there is 100% B.E.C.: *all* atoms are in the ground state φ of the single-particle Hamiltonian $H^{(1)}$. Therefore, φ is called the condensate (or macroscopic) wave function.

Interacting gas picture

The explanation above is limited to ideal Bose gases. In the presence of interaction a similar exact treatment is no longer possible. Since the many-body Hamiltonian is not just the sum of single-particle Hamiltonians, the ground state many-body wave function is not a product of N single-particle wave functions.

The first systematic and semi-rigorous treatment of interacting Bose fluids was due to Bogolubov in 1947, based on a new perturbation technique (see the historical perspective in the next section, as well as that in Sec. 4.2 concerning the intense mainstream after Bogolubov in the determination of the ground state energy of a cold dilute Bose gas). Bogolubov's theory relies on some unproved assumptions, yet provides a successful description of the phenomenology. Today it has become a standard textbook subject and the above references treat it extensively. For a comprehensive review in the current case of *cold and dilute Bose gases* we also refer to the 2001 Physics Report exhaustive presentation by Zagrebnov and Bru [113].

What one finds is that the ideal gas picture retains validity even when we include weak interactions. Here below we summarize the main points.

① The dilute nature of the gas allows one to describe the effects of the interaction in a rather fundamental way, via a single physical parameter, the s-wave scattering length a of the two-body scattering process. In fact, an important simplification for the alkalis and for H, is that the values of thermal energy k_BT characterizing B.E.C. conditions are small compared to the characteristic energies at which the two-body scattering process takes place. As a consequence, in the description of the scattering one legitimately neglects all partial waves, but the s-wave. The value of a could be in principle determined if one knew the interatomic potential with high precision. One can extract it from the experimental measurements; recently it has become possible in some cases to tune the value of a by applying an external magnetic potential, so to probe how theoretical results vary with a.

② Under ultra-high dilution, that is, when the mean interparticle distance is far larger than the scattering length,

$$\rho\, a^3 \ll 1\,, \tag{1.17}$$

one can neglects all but the two-body interactions among particles. The Hamiltonian is then

$$H \;=\; \sum_{k=1}^{N} \Big(\frac{\mathbf{p}_k^2}{2m} + U(\mathbf{r}_k) \Big) \;+ \sum_{1 \leqslant k < h \leqslant N} V(\mathbf{r}_k - \mathbf{r}_h)\,, \tag{1.18}$$

where the potential V has scattering length a and U is the trap. The ground state energy of H, provided that the gas is sufficiently diluted ($\rho\, a^3 \ll 1$) and populated (large N), turns out to be[1]

$$E^{\text{g.s.}} \sim \; \frac{4\pi\hbar^2 \rho\, a}{m}\, N\,. \tag{1.19}$$

③ A HEALING (or INDETERMINACY) LENGTH is defined as

$$\ell := \frac{1}{\sqrt{8\pi\rho\, a}}\,. \tag{1.20}$$

It is the length over which the perturbing effect of a confining wall is "healed" (whence its name): in the presence of repulsive interactions, it is energetically favourable to make the density of the gas nearly constant in the bulk of the box and to fall off to zero exponentially with characteristic length $\sim \ell$, as we approach a wall. In the low density regime it is impossible to localize the particles relative to each other (even though ρ is small), since

$$a \ll \rho^{-1/3} \ll \ell_c \tag{1.21}$$

(indeed $\frac{\rho^{-1/3}}{\ell_c} \sim \frac{\rho^{-1/3}}{(\rho a)^{-1/2}} = (\rho a^3)^{1/6} \ll 1$). Bosons in their ground state are therefore smeared out over distances large compared to the mean particle distance and their individuality is completely lost. They cannot be localized with respect to each other without changing the kinetic energy enormously.

④ As suggested by such a "delocalization picture", at zero temperature the many-body ground-state wave function turns out to be still, to a very good approximation, a product of N copies of the same single-particle wave function φ^{GP}.

$$\begin{aligned} \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) &= W(\mathbf{r}_1, \dots, \mathbf{r}_N)\, \varphi^{\text{GP}}(\mathbf{r}_1) \cdots \varphi^{\text{GP}}(\mathbf{r}_N) \\ &\sim \varphi^{\text{GP}}(\mathbf{r}_1) \cdots \varphi^{\text{GP}}(\mathbf{r}_N) \end{aligned} \tag{1.22}$$

This is the phenomenon of condensation. Such a φ^{GP} turns out to solve a nonlinear Schrödinger equation, the so-called Gross-Pitaevskiĭ equation

$$-\frac{\hbar^2}{2m}\Delta\varphi(\mathbf{r}) + U(\mathbf{r})\varphi(\mathbf{r}) + \frac{4\pi\hbar^2 a}{m}\, |\varphi(\mathbf{r})|^2 \varphi(\mathbf{r}) = \mu\, \varphi(\mathbf{r})\,, \tag{1.23}$$

[1] It is worth noticing that, within Bogolubov's theory, estimate (1.19) follows *assuming B.E.C.* Only 40 years after, the same result has been proved rigorously (see Sec. 4.2) for any dilute Bose gas with no additional assumption of condensation.

with $\int |\varphi(\mathbf{r})|^2 \, \mathrm{d}r = 1$, and the many-body ground state energy turns out to be

$$E^{\text{g.s.}} \sim N \frac{4\pi\hbar^2 \rho\, a}{m} \sim N \mathcal{E}^{\text{GP}}\big[\varphi^{\text{GP}}\big] \tag{1.24}$$

where

$$\mathcal{E}^{\text{GP}}[\varphi] := \int \Big(\frac{\hbar^2}{2m} |\nabla\varphi|^2 + U|\varphi|^2 + \frac{2\pi\hbar^2 a}{m} |\varphi|^4 \Big) \, \mathrm{d}\mathbf{r} \tag{1.25}$$

and φ^{GP} actually minimizes such an energy functional among all φ's such that $\int |\varphi(\mathbf{r})|^2 \, \mathrm{d}r = 1$. Here μ is fixed by normalisation and has the meaning of a chemical potential. The nonlinear term in (1.23) accounts for an on-site (i.e., local) self interaction, as if each particle would be subject to an additional potential given by the particle density itself, and describe the effect of two-body collisions. The issue of the *rigorous* derivation of (1.23), beyond the typical *formal* treatment that we will present Sec. 5.2, has been central in Mathematical Physics until the very recent rigorous results reported in Sec. 3.5 and 5.3, and still many major related problems remain open.

⑤ In the event of condensation, (1.22) can still be interpreted quantitatively as the presence of a "fraction $\overline{N}_0/N$ of particles" *in the condensate*, that is, each in the one-particle state φ^{GP}. The remaining "fraction of particles" $1 - \overline{N}_0/N$ is called the QUANTUM DEPLETION. This fraction is positive even at zero temperature, due to the presence of repulsive interaction, and one cannot have 100% B.E.C. like for the ideal case. At $T > 0$ depletion is enhanced by thermal effects and one speaks in addition of thermal depletion. In Bogolubov's theory

$$\text{quantum depletion} = \frac{8}{3\sqrt{\pi}} \sqrt{\rho a^3} \, , \tag{1.26}$$

typically 1% or less for the alkali condensates. This means that even for the interacting gases, we can, with 99% of accuracy, regard all the atoms to have the same single-particle wave function.[2]

⑥ This "imperfect Bose gas" turns out not to be infinitely compressible as the ideal counterpart. The transition gas/condensate can be regarded to be of the second order, provided that one neglects effects of higher order in a/λ_{dB} and $\rho a \lambda_{\text{dB}}^2$.

Hence, summarizing, occurrence of B.E.C. at zero temperature for interacting Bose gases is governed by the condition

$$\begin{cases} \text{negligible contribution of the } l \neq 0 \text{ waves to the two-body cross section} \\ \text{ultra-high dilution } \ \rho\, a^3 \ll 1 \end{cases} \tag{1.27}$$

which replaces criterion (1.15) for ideal gases.

The density distribution of a condensate can be directly observed in a nondestructive way. Such observations can be regarded as a direct visualization of the magnitude of the macroscopic wave function. In B.E.C., one simultaneously realizes millions of identical copies of the same one-body wave function and the entire many-body wave function can

[2] This is in contrast to liquid helium in which the quantum depletion is about 90%.

thus be measured while affecting only a small fraction of the condensed atoms. The resulting dramatic visualizations of wave functions are an appealing aspect of experimental studies of B.E.C.

In turn, experimental data can be fit with the numerical results, with an excellent agreement. Numerical analysis is essentially the only way to explicitly obtain thermodynamic quantities of the many-body picture (when they are computationally accessible) and the one-body condensate wave function. A number of techniques and tools has been developed across a huge literature: we refer to [8, 86, 106].

1.2 B.E.C. historically

The basic idea of B.E.C. dates back to 1925 when Einstein, on the basis of a 1924 paper by the Indian physicist Bose [18], devoted to the statistical description of the quanta of light, predicted the occurrence of a phase transition in a gas of non-interacting atoms [38, 39]. This phase transition is the consequence of quantum statistical effects. Bose's paper dealt with the statistics of photons, for which the total number is not a fixed quantity. He sent his paper to Einstein asking for comments.[3] Recognising its importance, Einstein translated the paper and submitted it for publication. Subsequently, he extended Bose's treatment to massive particles, whose total number is fixed.

For a long time these predictions had no practical impact and were viewed as a mathematical curiosity with little experimental interest, however. In 1938 London [81], immediately after the discovery of superfluidity in liquid helium, had the intuition that superfluidity could be a manifestation of B.E.C. (an issue which was highly controversial between him and Landau's "quantum hydrodynamics").

The first self-consistent theory of superfluids was developed by Landau [67] in 1941 in terms of the spectrum of elementary excitations of the fluid. In 1947 Bogolubov [15] developed the first microscopic theory of *interacting* Bose gases, based on the concept of B.E.C.: that theory, while intuitively appealing and undoubtedly correct in many aspects, has major gaps and some flaws. In the 1950s, after Landau and Lifshitz [68], Penrose [94], and Penrose and Onsanger [95] introduced the concept of the "off diagonal long range order" and discussed its relationship with B.E.C., intense theoretical work was developed,

[3] While presenting a lecture at the University of Dhaka on the photoelectric effect and the ultraviolet catastrophe, Bose intended to show his students that the current theory was inadequate, because it predicted results not in accordance with experimental results. During this lecture, Bose committed an error in applying the theory, which unexpectedly gave a prediction that agreed with the experiments. The error was a simple mistake – similar to arguing that flipping two fair coins will produce two heads one-third of the time – that would appear obviously wrong to anyone with a basic understanding of statistics. However, the results it predicted agreed with experiment and Bose realized it might not be a mistake at all. He for the first time took the position that the Maxwell-Boltzmann distribution would not be true for microscopic particles where fluctuations due to Heisenberg's uncertainty principle will be significant. Thus he stressed in the probability of finding particles in the phase space each having volumes h^3 and discarding the distinct position and momentum of the particles. Physics journals refused to publish Bose's paper. It was their contention that he had presented to them a simple mistake and Bose's findings were ignored. Discouraged, he wrote to Albert Einstein, who immediately agreed with him. His theory finally achieved respect when Einstein sent his own paper in support of Bose's to Zeitschrift für Physik, asking that they be published together (see Ref. [30, 62]). The reason Bose's "mistake" produced accurate results was that since photons are indistinguishable from each other, one cannot treat any two photons having equal energy as being two distinct identifiable photons. By analogy, if in an alternate universe coins were to behave like photons and other bosons, the probability of producing two heads would indeed be one-third (tail-head = head-tail).

aimed to better understand the relationship between B.E.C. and superfluidity. While theoretical intuition benefited greatly from this activity, the mathematical structure did not significantly improve.

In the same years, experimental studies on superfluid helium had become more and more refined, checking Landau's predictions for the excitation spectrum and providing the first measurements of the condensate fraction through the determination of the *momentum* distribution. An important development in the field took place with the prediction of quantized vortices by Onsager [93] (1949) and Feynman [48] (1955), and with their experimental discovery by Hall and Vinen [55] (1956).

Superfluid liquid ^{4}He is the prototype Bose-Einstein condensate and it has played a unique role in the development of physical concepts. However, interaction between helium atoms is strong and this reduces the number of atoms in the zero-momentum state even at absolute zero. Consequently, it is difficult to measure directly the occupancy (which is experimentally known today to be less than 1/10). The fact that interactions in liquid helium reduce it dramatically led to the search for *weakly interacting Bose gases* with a higher condensate fraction. The difficulty in most substances is that, at low temperatures, they do not remain gaseous but form solids or, in the case of the helium isotopes, liquids, and the effects of interaction thus become large. In other examples atoms first combine to form molecules, which subsequently solidify.

As long ago as in 1959 Hecht [56] argued that spin-polarized hydrogen would be a good candidate for a weakly interacting Bose gas. The attractive interaction between two hydrogen atoms with their electronic spins aligned was then estimated to be so weak that there would be no bound state. Thus, a gas of hydrogen atoms in a magnetic field would be stable against formation of molecules and, moreover, would not form a liquid, but remain a gas to arbitrarily low temperatures. Hecht's paper was ahead of its time and received little attention. The experimental studies on the dilute atomic gases were developed much later, starting from the 1970s, profiting from the new techniques developed in atomic physics based on magnetic and optical trapping, and advanced cooling mechanisms. In a series of experiments hydrogen atoms were first cooled in a dilution refrigerator, then trapped by a magnetic field and further cooled by evaporation, coming very close to B.E.C., although this approach was still limited by recombination of individual atoms to form molecules.

In the 1980s, laser-based techniques, such as laser cooling and magneto-optical trapping, were developed to cool and trap neutral atoms. Laser cooling have been the major experimental breakthrough, in contrast to so many other proposals which in practice work less well than that predicted theoretically. *Alkali atoms* are well suited to laser-based methods because their optical transition can be excited by available lasers and because they have a favourable internal energy-level structure for cooling to very low temperatures. Once they are trapped, their temperature can be lowered further by evaporative cooling.

By combining the different cooling techniques, the experimental teams of Cornell and Wieman at Boulder and of Ketterle at MIT eventually succeeded in 1995 in reaching the temperature and the densities required to observe B.E.C. in vapours of ^{87}Rb [6] and ^{23}Na [34], respectively. In the same year, first signatures of the occurrence of condensation in vapours of ^{7}Li were also reported [19]. B.E.C. was later achieved in other atomic species, including spin-polarized hydrogen, metastable ^{4}He [49], and ^{41}K [89].

In the end, the successful approach was to use laser cooling only as pre-cooling for

magnetic trapping and evaporative cooling. Laser cooling opened a new route to ultra-low temperature physics. Also, the number of atomic species which can be studied at ultra-low temperatures was greatly extended from helium and hydrogen to all of the alkali atoms, metastable rare gases, several earth-alkali atoms, and others (the list of laser cooled atomic species is still growing). In this sense, the 2001 Nobel prize in Physics to Cornell, Ketterle, and Wieman "for the achievement of Bose-Einstein condensation in dilute gases of alkali atoms, and for early fundamental studies of the properties of the condensates" is along the same mainstream of the 1997 Nobel prize in Physics to Chu, Cohen-Tannoudji and Phillips "for development of methods to cool and trap atoms with laser light".

Let us conclude this historical outlook by mentioning that in the course of time the concept of B.E.C. have found applications in many system other than liquid He and boson alkali atoms. For example, many aspects of the behaviour of superconductors may be understood qualitatively on the basis of the idea that pairs of electrons form a Bose-Einstein condensate (although their properties are quantitatively very different from those of weakly interacting gas of pairs). B.E.C. of pairs of fermions is also observed experimentally in atomic nuclei, where the effect of the neutron-neutron, proton-proton, and neutron-proton pairing may be seen in the excitation spectrum as well as in reduced moments of inertia. Theoretically, B.E.C. of nucleon pairs is expected to play an important role in the interiors of neutron stars and observations of glitches in the spin-down rate of pulsars have been interpreted in terms of neutron superfluidity. The possibility of mesons, either pions or kaons, forming a Bose-Einstein condensate in the cores of neutron stars has been widely discussed, since it would have far-reaching consequences for theories of supernovae and the evolution of neutron stars. In the field of nuclear and particle Physics the ideas of Bose-Einstein condensation also find application in the understanding of the vacuum as a condensate of quark-antiquark ($u\bar{u}$, $d\bar{d}$, and $s\bar{s}$) pairs, the so-called chiral condensate. This condensate gives rise to particle masses in much the same way as the condensate of electron pairs in a superconductor gives rise to the gap in the electronic excitation spectrum.

1.3 B.E.C. experimentally

The current explosion of activity con ultra-cold atoms and Bose-Einstein condensation is well documented by the huge number of groups working on it, partially listed at

`http://www.uibk.ac.at/exphys/ultra-cold/atomtraps.html` ,

and by the vast bibliography of several thousands papers related to B.E.C. written in the lat decade, that one can find collected at

`http://bec01.phy.georgiasouthern.edu/bec.html/bibliography.html` .

So far B.E.C. has been realized in ^{7}Li, ^{23}Na, ^{41}K, ^{87}Rb, as well as in spin-polarized H and metastable ^{4}He. The range of values for the relevant physical parameter in the condensed regime are collected in Table 1.1.

In particular, in these systems B.E.C. shows up not only in momentum space but also in *coordinate space*, making the direct experimental investigation of the condensation feasible and providing new opportunities for interesting studies, like the temperature dependence of the condensate, energy and density distributions, interference phenomena, frequencies of collective excitations, and so on.

Table 1.1. Typical values for the relevant physical parameters in the regime of condensation (expressed in the convenient units for each quantity)

	typical B.E.C. values
Temperature	50 nK $\div$ 50 μK
Number of atoms	$10^2 \div 10^{11}$
Trap size	10 $\div$ 50 μm
Mean density	$10^{11} \div 10^{15}$ cm^{-3}
Scattering length	300 $\div$ 300 a.u.
Dilution (ρa^3)	$10^{-4} \div 10^{-6}$

Trapping, cooling, detecting condensation

Making and probing Bose-Einstein condensates turns out to be a subtle art [64]. At the ultra-low temperatures and densities of the experiments, the equilibrium configuration would be the solid phase. Thus, in order to observe B.E.C., one has to preserve the system in a metastable gas phase for a sufficiently long time.

In order to create a Bose condensate in a dilute gas, atoms must be cooled and compressed in a trap until the thermal de Broglie wavelength is comparable to the spacing between atoms. In addition, atoms must be thermally isolated from all material walls. This is done by trapping atoms with magnetic fields or with laser light inside ultra-high vacuum chambers. Such traps can store atoms for seconds or even minutes, which is enough time to cool them. Efforts have to face the challenge of simultaneously achieving effective laser cooling and trapping, which work best at low atomic densities, and efficient evaporative cooling, which requires high densities.

Pre-cooling is a prerequisite for trapping because conservative atom traps can only confine neutral atoms with a maximum energy of one kelvin at best (and in many cases the trap depth is just a few millikelvins). The pre-cooling is done by laser cooling and the final cooling by evaporation. These cooling techniques together reduce the temperature of the atoms by a factor of a billion (from $10^2 \div 10^3$ K in the oven, to $1 \div 10$ μK after laser cooling, to $\sim 10^2$ nK after evaporative cooling). In almost all B.E.C. experiments, pre-cooling is done in a magneto-optical trap loaded from a slowed beam.

To keep the atoms tightly compressed during cooling and hold the condensate for study, the gas is confined by conservative atom traps that have a sufficiently small heating rate. The requirements for the trap during cooling are more stringent than they are for holding condensates. First, the time for cooling (typically 30 sec for evaporative cooling) is usually much longer than the time for performing experiments on B.E.C., requiring low heating and trap loss rates. Furthermore, for cooling, the trap needs to be sufficiently deep to hold the initial (pre-cooled) cloud, and must accommodate a cooling scheme able to reach B.E.C. temperatures. In particular, suitable magnetic traps (e.g., quadrupole-type traps, Ioffe-Pritchard traps, etc.) accommodate the pre-cooled atoms and compress them in order to achieve high collision rates and efficient evaporative cooling.

After that, evaporative cooling (like hot water that cools down in a bathtub or in a cup of coffee) removes continuously the high-energy tail of the thermal distribution from the trap. The evaporated atoms carry away more than the average energy, which means that the temperature of the remaining atoms decreases. The high energy tail must be

constantly repopulated by collisions, thus maintaining thermal equilibrium and sustaining the cooling process. Evaporation of atoms is controlled by radio frequency radiation ("rf induced evaporation") that flips the atomic spin. As a result, the attractive trapping force turns into a repulsive force and expels the atoms from the trap. Evaporative cooling proved to work much better for alkali atoms than for hydrogen, for which the technique was originally developed. In fact, it turns out that for alkali atoms, in contrast to hydrogen, the so-called good collisions (elastic collisions necessary for the evaporation process) highly dominate over the so-called bad collisions (inelastic two- and three-body collisions).

The resulting condensates can then be manipulated to study their properties in different shapes and symmetries and to explore their dynamic behaviour. Direct-contact manipulation is excluded, due to the extremely low temperature of the condensate and the high sticking probability of atoms on cold surfaces. Instead, special tools have been developed to shape, slice, kick, shake, and stir condensates. Typically, by magnetic fields one adiabatically expands the condensate and changes its aspect ratio. Focused off-resonant laser beams are used for "microsurgery" on condensates. and to create deformed (non-parabolic) trapping potentials.

We remark that everything we know about gaseous Bose condensates has been obtained by optical diagnostics. "Contact probes" cannot be used because the samples are much smaller ($\sim 10^7$ atoms) than even a 10 μm sized probe ($\sim 10^{13}$ atoms), which would cause the atoms to equilibrate with the probe rather than the opposite. With the "cooling power" of rf evaporation one can cool at most 10^8 atoms per minute and it would take several months just to cool the sensor tip. Fortunately, optical diagnostics are very versatile and the ease with which light scattering methods are implemented for dilute atomic samples is a major advantage over condensed matter systems.

The two most important techniques for observing Bose-Einstein condensates are in-situ and time-of-flight imaging. In both cases, one obtains an image which reflects the density distribution of the atoms either in a trapped state or in ballistic expansion. Absorption imaging is done by illuminating the atoms with a laser beam and imaging the shadow cast by the atoms onto a CCD camera. To image a transparent object, information encoded in the phase shift of the light must be converted into intensity information which can be detected by a photosensor. All dispersive methods rely on the ability to distinguish scattered and unscattered components of the probe light and manipulate them independently.

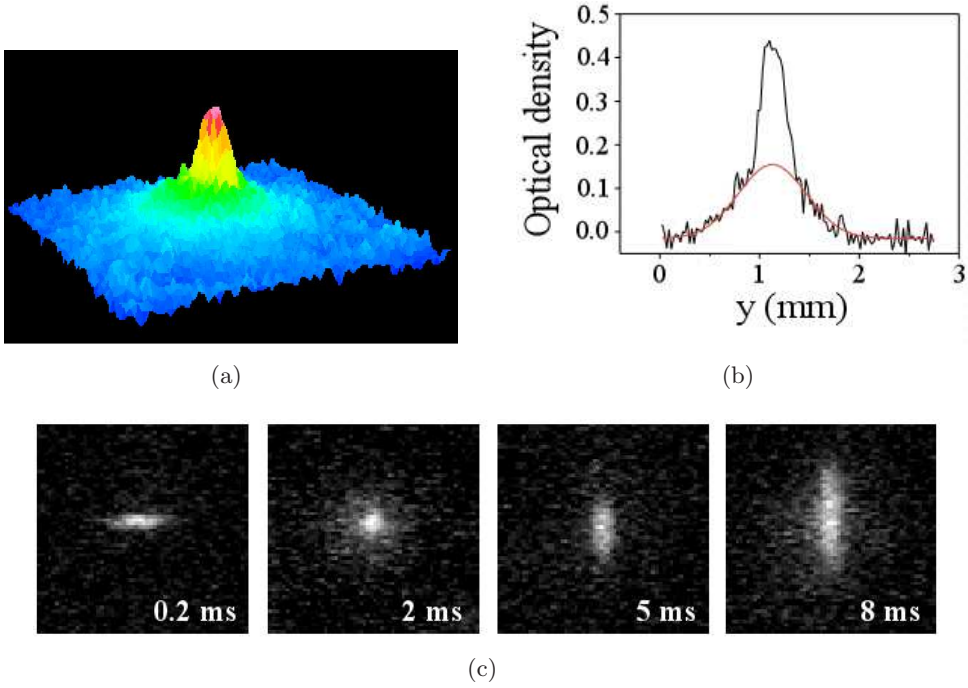
A discernible image with 30×30 pixels and 100 detected photons per pixel involves 10^5 photons. Even if each photon would "knock" an atom out of the condensate, this would be "non-perturbative" if the sample had many millions of atoms. This explains why the large Bose condensates can be observed non-destructively with absorption spectroscopy.

The current generation of low-noise CCD cameras meets all the needs of current B.E.C. experiments. With about 5 electrons rms noise per pixel, the camera is already shot-noise limited for more than about 50 detected photons per pixel. The low noise of the camera usually requires a slow readout over several seconds. Rapid sequencing of non-destructive images is achieved by storing the whole sequence on the chip and then reading it out slowly. For this purpose, the sensor is divided into an area for exposure and a much larger area for storage.

Next figures give a visible flavour of the very subtle experimental expedients to make condensation possible. Fig. 1.3 refers to an observation of B.E.C. with metastable helium

at Laboratoire Kastler Brossel, Paris.[4] Fig. 1.4 describe a typical magnetic trap for confining atoms in the cooling process. Fig. 1.5, 1.6, 1.7, and 1.8 survey the main steps and features of an experiment of condensation with ^{87}Rb realized by the experimental group of York University, Toronto.[5]

[4] http://www.lkb.ens.fr/recherche/atfroids/anglais/activite_an.html
[5] http://www.yorku.ca/wlaser/projects/projects_BEC.htm



(a) (b)

(c)

Figure 1.3. Observation of a condensate of metastable helium atoms. (a) 3D image. (b) 1D profile of the same cloud after 4ms of expansion time. One can clearly see the difference between the condensed and the thermal fraction which is fitted by a bosonic distribution in the 1D profile. (c) Images of a pure condensate containing 400 000 atoms for different expansion time after switching off the trap. Initially the elliptical shape of the condensate corresponds to the geometry of the trap. As the expansion time increases, the ellipticity of the condensate undergoes an inversion due to mean-field interactions between atoms. This inversion is the strongest evidence for the existence of a Bose-Einstein condensate.

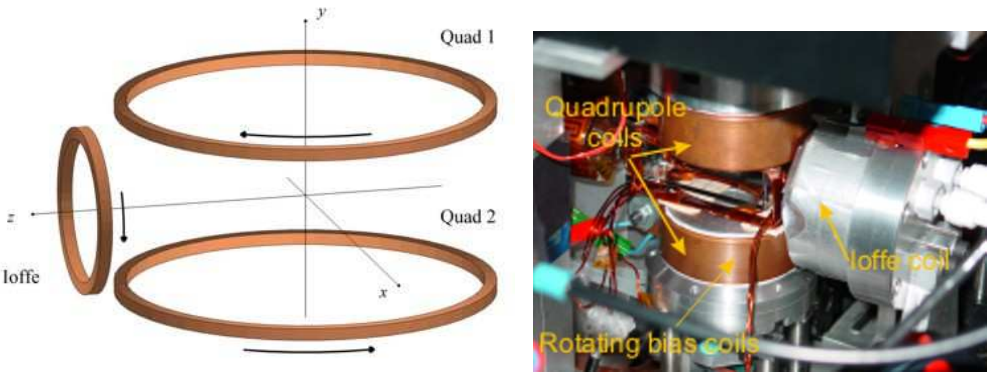


Figure 1.4. QUadrupole-Ioffe Configuration (QUIC). It consists of two quadrupole coils and Ioffe coil, with current directions shown by the black arrows in the left side. The magnetic fields of the Ioffe coil and the quadrupole pair are in the opposite direction on the z axis, so that they almost cancel. If the currents in all the coils are equal then the design of the coil geometry means a harmonic single well exists. If the Ioffe coil current is less than that of the quadrupole pair a double-well potential is created. For scale, the quadrupoles are 80 mm in diameter. The coils are driven at up to 27 Amps and are water cooled. The vacuum system, which is at a pressure of less than 10^{-11} torr, extends from the left between the two quadrupole coils.

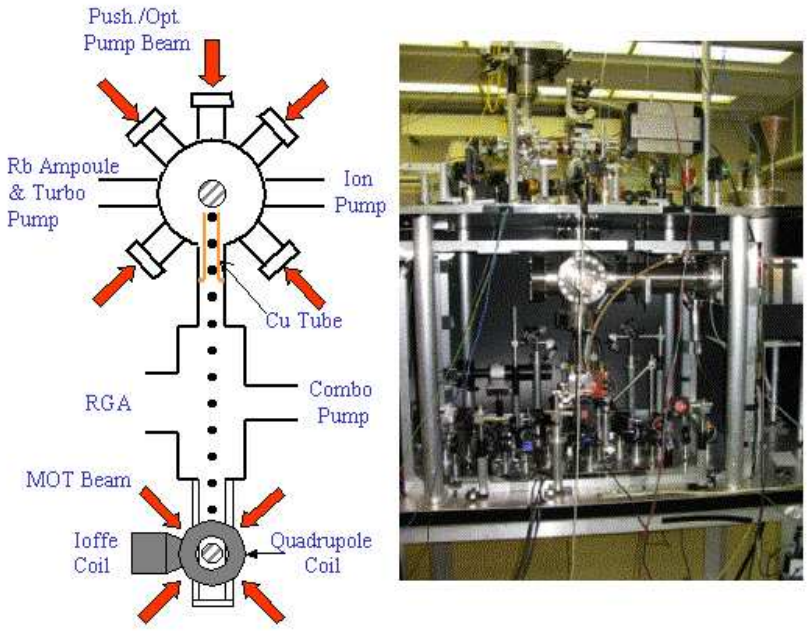


Figure 1.5. Apparatus: upper and lower vacuum chambers. Atoms are first cooled in the upper vapour cell magneto-optical trap. A laser then pushes the atoms into the lower chamber which is at a pressure of about $2 \cdot 10^{-11}$ torr. The atoms are then trapped in a second magneto-optical trap and subsequently loaded into a so called QUIC trap (QUadrupole-Ioffe Configuration). Evaporative cooling is finally used to achieve Bose Einstein Condensation.

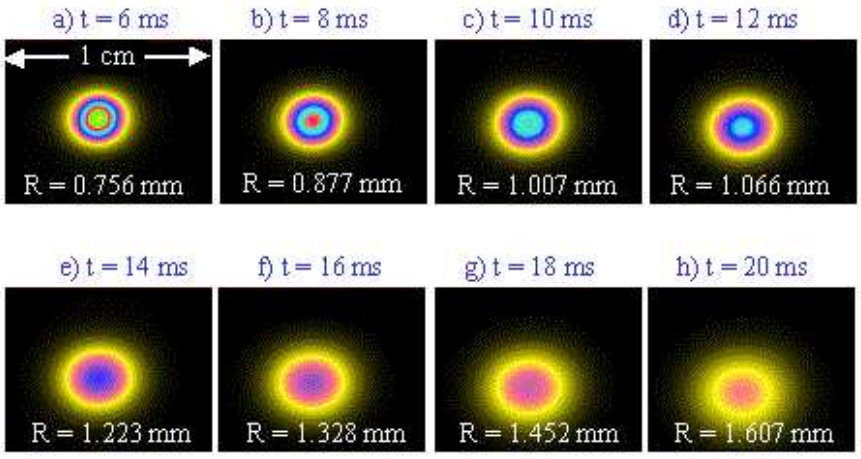


Figure 1.6. Temperature determination of ultra-cold atoms. The atom temperature is determined by switching off the magnetic fields and observing the expansion of the atom cloud. A hotter sample of atoms expands faster than a colder sample. The data below show actual pictures (not modelled data) of such an expanding cloud: from the plot of cloud radius vs time, temperature is found to be 52 microkelvins.

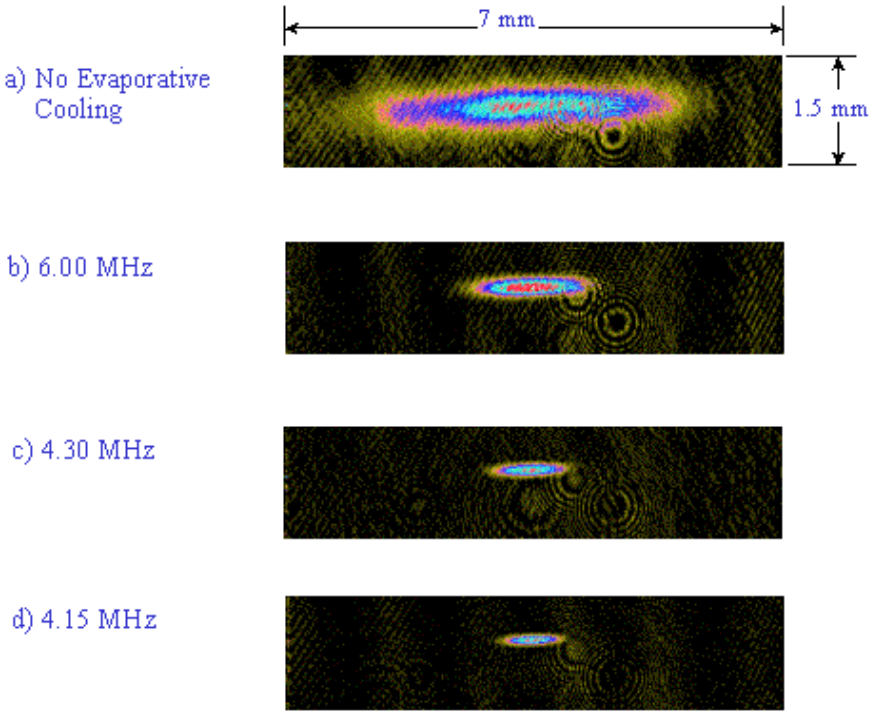


Figure 1.7. Evaporative cooling. The final stage of atom cooling is to switch off all laser beams and apply a radio frequency signal using a small 1 loop antenna. The purpose of this frequency is to flip the spins of the hot atoms which are then expelled from the trap resulting in a cooler collection of trapped atoms. The radio frequency is swept from 20 MHz to the lower frequencies shown in the figure. The cloud size clearly gets smaller as the lower frequency decreases causing the temperature to decrease accordingly.

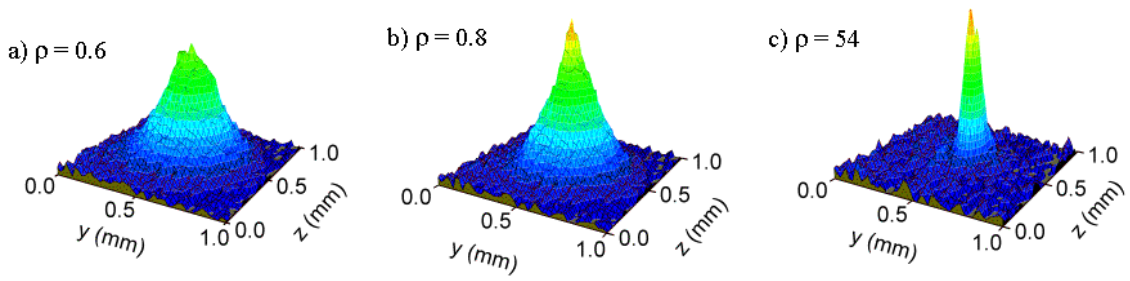


Figure 1.8. Transition to B.E.C. (a) Thermal cloud with $N = 1.9 \cdot 10^6$ atoms at temperature $T = 450$ nK. (b) Mixed thermal atom cloud and B.E.C. where $N = 1.8 \cdot 10^6$ and $T = 400$ nK. (c) Pure condensate where $N = 4.2 \cdot 10^5$ atoms and $T < 60$ nK.

Chapter 2

Mathematical preliminaries

This chapter collects some general theoretic and preparatory material, which is suitably re-organized with respect to standard references as [98, 99]. The scheme is that of Ref. [83]. *Density matrices* will be introduced and discussed as special trace class, Hilbert-Schmidt and compact operators on a Hilbert space, and the trace will be discussed in abstract, as well as concretely in terms of *kernels* (Sec. 2.1, 2.2). Then the partial trace operation will be introduced and characterized, leading to the key notion of *reduced density matrix* (Sec. 2.3). Such material is classic and statements will be quoted without proof. Also we inserted the not-so-known Brislawn's rigorous control of the well-posedness of the trace of a kernel operator (a density matrix, for our purposes) in terms of the integral of the diagonal of its kernel. The chapter is concluded by Sec. 2.4 with a discussion on the distance of states and of reduced density matrices.

2.1 Density matrices: definition and general properties

The aim of this section is to introduce the main notion of the present chapter, i.e., a *density matrix*. Notoriously, it is the mathematical tool for describing states in quantum systems. Density matrices are special trace class operators. Let us recall first some useful basic facts about the trace class.

Let $\mathcal{H}$ be a separable Hilbert space with scalar product $(\,\cdot\,,\,\cdot\,)$ and norm $\|\cdot\|$. It will be always assumed that the Hilbert structure is on $\mathbb{C}$. Separability will be needed throughout to guarantee the existence of a countable complete orthonormal system $\{\varphi_j\}_j$ of $\mathcal{H}$, usually referred to simply as an orthonormal basis. The finite or countable set the index j runs in will be possibly omitted, so that we will write $\{\varphi_j\}_{j=0}^{\infty}$, $\{\varphi_j\}_{j=1}^{\infty}$, $\{\varphi_j\}_{j=0}^{M}$, or simply $\{\varphi_j\}_j$, depending on the context. In the absence of other specifications, we will be working with an infinite-dimensional $\mathcal{H}$. Denote by $\mathcal{L}(\mathcal{H})$ and $\mathrm{Com}(\mathcal{H})$ the spaces of bounded and of compact operators on $\mathcal{H}$ respectively.

Definition–Theorem 2.1.1 (the trace class). Let $T \in \mathcal{L}(\mathcal{H})$. Then the sequence $\{(\varphi_j, |T|\varphi_j)\}_j$ is summable or not independently of the orthonormal basis $\{\varphi_j\}_j$ of $\mathcal{H}$. Denote its sum by

$$\mathrm{Tr}\,|T| := \sum_{i=1}^{\infty}(\varphi_j, |T|\varphi_j)\,. \tag{2.1}$$

The TRACE CLASS $\mathcal{L}^1(\mathcal{H})$ is the space

$$\mathcal{L}^1(\mathcal{H}) := \{T \in \mathcal{L}(\mathcal{H}) \,:\, \mathrm{Tr}\,|T| < +\infty\} \tag{2.2}$$

Properties:

① $\mathcal{L}^1(\mathcal{H})$ is a $*$-ideal of $\mathcal{L}(\mathcal{H})$. It is closed under the usual Banach topology of $\mathcal{L}(\mathcal{H})$, that is, the operator norm topology, iff $\dim \mathcal{H} < \infty$.

② When endowed with the norm

$$\|T\|_{\mathcal{L}^1} := \operatorname{Tr} |T| \,, \tag{2.3}$$

then $\mathcal{L}^1(\mathcal{H})$ is a Banach space (in fact, a Banach $*$-algebra) and

$$\|T\| \leqslant \|T\|_{\mathcal{L}^1} \,. \tag{2.4}$$

Also, the finite rank operators on $\mathcal{H}$ are $\|\cdot\|_{\mathcal{L}^1}$-dense in $\mathcal{L}^1(\mathcal{H})$.

③ $\mathcal{L}^1(\mathcal{H}) \subset \mathrm{Com}(\mathcal{H})$ densely in the operator norm topology.

If $T \in \mathcal{L}^1(\mathcal{H})$ then $\sum_{i=1}^{\infty}(\varphi_j, T\varphi_j)$ converges absolutely and the limit is independent of the orthonormal basis $\{\varphi_j\}_j$. Thus, on the trace class one defines the operation of TRACE as

$$\begin{aligned} \mathcal{L}^1(\mathcal{H}) &\longrightarrow \mathbb{C} \\ T &\longmapsto \operatorname{Tr}[T] := \sum_{i=1}^{\infty} (\varphi_j, T\varphi_j) \,. \end{aligned} \tag{2.5}$$

④ Trace is linear; furthermore

$$\begin{aligned} \operatorname{Tr}[T^*] &= \operatorname{Tr}[T] \\ \operatorname{Tr}[TS] &= \operatorname{Tr}[ST] \\ |\operatorname{Tr}[T]| &\leqslant \|T\|_{\mathcal{L}^1} \\ |\operatorname{Tr}[TS]| &\leqslant \|S\| \, \|T\|_{\mathcal{L}^1} \end{aligned} \tag{2.6}$$

$\forall\, T \in \mathcal{L}^1(\mathcal{H})$ and $S \in \mathcal{L}(\mathcal{H})$.

⑤ The map $T \mapsto \operatorname{Tr}[T\,\cdot\,]$ is an isometric isomorphism of $\mathcal{L}^1(\mathcal{H})$ onto $\mathrm{Com}(\mathcal{H})^*$, whence $\mathcal{L}^1(\mathcal{H}) = \mathrm{Com}(\mathcal{H})^*$.

⑥ The map $S \mapsto \operatorname{Tr}[S\,\cdot\,]$ is an isometric isomorphism of $\mathcal{L}(\mathcal{H})$ onto $\mathcal{L}^1(\mathcal{H})^*$, whence $\mathcal{L}(\mathcal{H}) = \mathcal{L}^1(\mathcal{H})^*$.

The following familiar notation will turn out to be useful in the sequel.

Definition–Theorem 2.1.2 (Dirac notation). $\forall\, \varphi, \psi \in \mathcal{H}$ one defines the operator

$$\begin{aligned} |\varphi\rangle\langle\psi| : \ \mathcal{H} &\longrightarrow \mathcal{H} \\ \xi &\longmapsto |\varphi\rangle\langle\psi|\xi\rangle := (\psi, \xi)\, \varphi \,. \end{aligned} \tag{2.7}$$

Its range has dimension 0 or 1 and its norm is

$$\| \, |\varphi\rangle\langle\psi| \, \| = \|\varphi\| \, \|\psi\| \,. \tag{2.8}$$

Further,

$$\begin{aligned} |\varphi\rangle\langle\psi|^* &= |\psi\rangle\langle\varphi| \\ |\lambda\varphi\rangle\langle\psi| &= \lambda\,|\varphi\rangle\langle\psi| \\ |\varphi\rangle\langle\lambda\psi| &= \bar\lambda\,|\varphi\rangle\langle\psi| \\ |\varphi+\varphi'\rangle\langle\psi| &= |\varphi\rangle\langle\psi| + |\varphi'\rangle\langle\psi| \\ |\varphi\rangle\langle\psi+\psi'| &= |\varphi\rangle\langle\psi| + |\varphi\rangle\langle\psi'| \\ |\varphi'\rangle\langle\psi'|\,|\varphi\rangle\langle\psi| &= (\psi',\varphi)\,|\varphi'\rangle\langle\psi| \\ S\,|\varphi\rangle\langle\psi| &= |S\varphi\rangle\langle\psi| \\ |\varphi\rangle\langle\psi|\,S &= |\varphi\rangle\langle S^*\psi| \end{aligned} \tag{2.9}$$

$\forall\,\varphi,\varphi',\psi,\psi'\in\mathcal{H}$, $\forall\,\lambda\in\mathbb{C}$, $\forall\,S\in\mathcal{L}(\mathcal{H})$.

Definition–Theorem 2.1.3 (bounded operators with canonical representation $\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j|$). Let $\{\varphi_j\}_j$ and $\{\psi_j\}_j$ be two orthonormal (not necessarily complete) families of vectors in $\mathcal{H}$ and let $\{\lambda_j\}_j$ be a family of complex numbers, indexed by the same set $\{j\}$. The sequence

$$\{\,\lambda_j|\varphi_j\rangle\langle\psi_j|\xi\rangle\,\}_j = \{\,\lambda_j(\psi_j,\xi)\varphi_j\,\}_j$$

is summable *for every* $\xi\in\mathcal{H}$ iff $\{\lambda_j\}_j$ is bounded. Whenever this happens, the sequence of operators $\{\lambda_j|\varphi_j\rangle\langle\psi_j|\}_j$ sums in $\mathcal{L}(\mathcal{H})$ to an operator denoted by $\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j|$, with norm

$$\Big\|\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j|\Big\| = \sup_j|\lambda_j|\,. \tag{2.10}$$

In the usual notation

$$\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j| = \sum_j \lambda_j(\psi_j,\,\cdot\,)\,\varphi_j\,. \tag{2.11}$$

Properties:

① One has

$$\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j| = \mathbb{O} \quad\Leftrightarrow\quad \lambda_j = 0 \quad\forall\,j \tag{2.12}$$

whence

$$\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j| = \sum_j \mu_j|\varphi_j\rangle\langle\psi_j| \quad\Leftrightarrow\quad \lambda_j=\mu_j \quad\forall\,j\,. \tag{2.13}$$

② If $\lambda_j\neq 0\ \forall\,j$, then

$$\begin{aligned} \sum_j \lambda_j|\varphi_j\rangle\langle\psi_j| \quad\text{is injective} \quad&\Leftrightarrow\quad \{\psi_j\}_j \quad\text{is complete} \\ \sum_j \lambda_j|\varphi_j\rangle\langle\psi_j| \quad\text{has dense range} \quad&\Leftrightarrow\quad \{\varphi_j\}_j \quad\text{is complete}\,. \end{aligned} \tag{2.14}$$

③ Let $S=\sum_j \lambda_j|\varphi_j\rangle\langle\psi_j|$ as above; then

$$\begin{aligned} S^* &= \sum_j \bar\lambda_j|\psi_j\rangle\langle\varphi_j| \\ S^*S &= \sum_j |\lambda_j|^2|\psi_j\rangle\langle\psi_j| \\ |S| &= \sum_j |\lambda_j|\,|\psi_j\rangle\langle\psi_j| \end{aligned} \tag{2.15}$$

④ $S \in \mathcal{L}(\mathcal{H})$ is an orthogonal projection iff $S = \sum_j |\varphi_j\rangle\langle\varphi_j|$, its range being the subspace spanned by $\{\varphi_j\}_j$.

⑤ $S \in \mathcal{L}(\mathcal{H})$ is unitary iff $S = \sum_j |\varphi_j\rangle\langle\psi_j|$ where both $\{\varphi_j\}_j$ and $\{\psi_j\}_j$ are orthonormal basis.

⑥ $S \in \mathcal{L}(\mathcal{H})$ is a partial isometry iff $S = \sum_j |\varphi_j\rangle\langle\psi_j|$, where $\{\varphi_j\}_j$ and $\{\psi_j\}_j$ are not necessarily complete, and in this case the projections onto the initial and final space are $S^*S = \sum_j |\psi_j\rangle\langle\psi_j|$ and $SS^* = \sum_j |\varphi_j\rangle\langle\varphi_j|$, respectively.

⑦ A given $S = \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j|$ is invertible iff both $\{\varphi_j\}_j$ and $\{\psi_j\}_j$ are orthonormal basis and $0 < c_1 < |\lambda_j| < c_2 < +\infty$ for some positive constant c_1, c_2 and $\forall\, j$; in this case $S^{-1} = \sum_j \frac{1}{\lambda_j} |\psi_j\rangle\langle\varphi_j|$.

Theorem 2.1.4 (**compacts**). Let $C \in \mathcal{L}(\mathcal{H})$.

① $C \in \mathrm{Com}(\mathcal{H})$ iff it admits the following so called CANONICAL FORM:

$$\begin{aligned} C = & \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j| \\ & \{\varphi_j\}_j \text{ and } \{\psi_j\}_j \text{ orthonormal systems} \\ & \lambda_j > 0 \quad \forall\, j\,. \end{aligned} \tag{2.16}$$

Series in (2.16) converges in operator norm. Both the orthonormal systems are not necessarily complete in $\mathcal{H}$, so that the j's run on a finite or countably infinite set of integers; in this last case, necessarily $\lambda_j \to 0$. The λ_j's are called the SINGULAR VALUES of C. By (2.15), $\{\lambda_j\}_j$ is precisely the set of nonzero eigenvalues of $|C|$.

② $C \in \{\text{finite rank operators}\}$ iff for some positive integer m

$$\begin{aligned} C = & \sum_{j=1}^{m} \lambda_j |\varphi_j\rangle\langle\psi_j| \\ & \{\varphi_j\}_{j=1}^m \text{ and } \{\psi_j\}_{j=1}^m \text{ orthonormal systems} \\ & \lambda_j > 0 \quad \forall\, j = 1, \ldots, m\,. \end{aligned} \tag{2.17}$$

③ Let, in addition, $C = C^* \in \mathrm{Com}(\mathcal{H})$. Then there exists an orthonormal basis $\{\varphi_j\}_{j=1}^{\infty}$ of $\mathcal{H}$ of eigenvectors for C (i.e., $C\varphi_j = \lambda_j \varphi_j$). Its nonzero eigenvalues have finite multiplicity and $\lambda_j \to 0$ as $j \to \infty$. In Dirac notation

$$C = \sum_{j=1}^{\infty} \lambda_j |\varphi_j\rangle\langle\varphi_j|\,. \tag{2.18}$$

(This is the Hilbert-Schmidt theorem.)

Theorem 2.1.5 (**trace class and compacts**). ① $\forall\, \varphi, \psi \in \mathcal{H}$, $|\varphi\rangle\langle\psi| \in \mathcal{L}^1(\mathcal{H})$ and

$$\begin{aligned} \mathrm{Tr}[\,|\varphi\rangle\langle\psi|\,] &= (\psi, \varphi) \\ \|\,|\varphi\rangle\langle\psi|\,\|_{\mathcal{L}^1} &= \|\varphi\|\,\|\psi\|\,. \end{aligned} \tag{2.19}$$

② Let $T \in \mathrm{Com}(\mathcal{H})$ and let $T = \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j|$ be its canonical form. Then $T \in \mathcal{L}^1(\mathcal{H})$ iff $\sum_j \lambda_j < +\infty$. If so,

$$\begin{aligned} \mathrm{Tr}[T] &= \sum_j \lambda_j (\psi_j, \varphi_j) \\ \|T\|_{\mathcal{L}^1} = \mathrm{Tr}|T| &= \sum_j \lambda_j \,. \end{aligned} \tag{2.20}$$

With the background presented so far one introduces the main mathematical object of this work.

Definition 2.1.6. A DENSITY MATRIX is a positive trace class operator on $\mathcal{H}$ with trace one.

As a consequence of theorems 2.1.3, 2.1.4, and 2.1.5, one has the following.

Definition–Theorem 2.1.7 (pure and mixed states). ① Any density matrix γ on $\mathcal{H}$ can be written in diagonal form as

$$\gamma = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j| \,, \qquad \lambda_j > 0 \,, \qquad \sum_j \lambda_j = 1 \,, \tag{2.21}$$

where $\{\psi_j\}_j$ is an orthonormal system of $\mathcal{H}$ and $\{\lambda_j\}_j$ is the corresponding sequence of eigenvalues.

② When in (2.21) at least two nonzero λ_j's enter, then γ is said to be a MIXED STATE. If so,

$$\begin{aligned} \gamma^2 &= \sum_j \lambda_j^2 \, |\psi_j\rangle\langle\psi_j| \\ \mathrm{Tr}[\gamma^2] &= \sum_j \lambda_j^2 < 1 \,. \end{aligned} \tag{2.22}$$

③ Conversely, if γ has only one nonzero eigenvalue, which is necessarily 1, then it takes the form of the rank-one orthogonal projection $\gamma = |\psi\rangle\langle\psi| = \gamma^* = \gamma^2$ and it is said a PURE STATE. For pure states $\mathrm{Tr}[\gamma^2] = 1$. Thus, there is a one-to-one correspondence between pure states and rays of $\mathcal{H}$, where by RAY $\{\psi\}$ one defines the subspace generated by $\psi \in \mathcal{H}$.

2.2 Trace of kernels

In the standard physical applications, the Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^n)$ (or $L^2(\Omega)$ for some measure space Ω), and the nature of density matrices as operators with an integral kernel arises naturally. The perspective is then to regard density matrices as special Hilbert-Schmidt operators. In this section Hilbert-Schmidt operators are reviewed as kernel operators and the notion of trace is given in terms of kernels.

Definition–Theorem 2.2.1 (Hilbert-Schmidt operators). The HILBERT-SCHMIDT class is the space

$$\mathcal{L}^2(\mathcal{H}) := \{S \in \mathcal{L}(\mathcal{H}) \,:\, \mathrm{Tr}[S^*S] < +\infty\} \tag{2.23}$$

that is, S is Hilbert-Schmidt iff S^*S is trace class. Properties:

① $\mathcal{L}^2(\mathcal{H})$ is a $*$-ideal of $\mathcal{L}(\mathcal{H})$. It is closed under the usual Banach topology of $\mathcal{L}(\mathcal{H})$ iff $\dim \mathcal{H} < \infty$.

② $T \in \mathcal{L}^1(\mathcal{H})$ iff $T = S_1 S_2$ with $S_1, S_2 \in \mathcal{L}^2(\mathcal{H})$. Thus, the product of two Hilbert-Schmidt is traceable. When endowed with the scalar product

$$(S_1, S_2)_{\mathcal{L}^2} := \mathrm{Tr}[S_1^* S_2]\,, \tag{2.24}$$

then $\mathcal{L}^2(\mathcal{H})$ is a Hilbert space. In fact, it is a Banach $*$-algebra with respect to the inherited norm $\|S\|_{\mathcal{L}^2} = (S,S)_{\mathcal{L}^2}^{1/2}$. One has

$$\begin{aligned} (S_2, S_1)_{\mathcal{L}^2} &= \overline{(S_1, S_2)_{\mathcal{L}^2}} \\ (RS_1, S_2)_{\mathcal{L}^2} &= (S_1, R^* S_2)_{\mathcal{L}^2} \\ (S_1 R, S_2)_{\mathcal{L}^2} &= (S_1, S_2 R^*)_{\mathcal{L}^2} \\ \|RS\|_{\mathcal{L}^2} &\leqslant \|R\|\, \|S\|_{\mathcal{L}^2} \end{aligned} \tag{2.25}$$

$\forall\, R \in \mathcal{L}(\mathcal{H})$ and $\forall\, S, S_1, S_2 \in \mathcal{L}^2(\mathcal{H})$. Further,

$$\|T\| \leqslant \|T\|_{\mathcal{L}^2} \leqslant \|T\|_{\mathcal{L}^1} \qquad \forall\, T \in \mathcal{L}^1(\mathcal{H}) \tag{2.26}$$

and

$$\{\text{finite rank operators}\} \subset \mathcal{L}^1(\mathcal{H}) \subset \mathcal{L}^2(\mathcal{H}) \subset \mathrm{Com}(\mathcal{H}) \tag{2.27}$$

where each possible inclusion of any space into any other space of the chain is dense with respect to the topology of the larger space.

③ Let $S \in \mathcal{L}(\mathcal{H})$. For any two orthonormal basis $\{\varphi_j\}_j$, $\{\psi_j\}_j$ of $\mathcal{H}$, the sequences

$$\{\|S\varphi_j\|^2\}_j\,, \quad \{\|S\psi_j\|^2\}_j\,, \quad \{|(\varphi_i, S\psi_j)|^2\}_{i,j}$$

are simultaneously summable or not. $S \in \mathcal{L}^2(\mathcal{H})$ iff they are summable; if so,

$$\|S\|_{\mathcal{L}^2}^2 = \sum_j \|S\varphi_j\|^2 = \sum_j \|S\psi_j\|^2 = \sum_{i,j} |(\varphi_i, S\psi_j)|^2\,. \tag{2.28}$$

④ Let $S \in \mathrm{Com}(\mathcal{H})$ and let $S = \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j|$ be its canonical form. Then $S \in \mathcal{L}^2(\mathcal{H})$ iff $\sum_j \lambda_j^2 < +\infty$. If so,

$$\|S\|_{\mathcal{L}^2}^2 = \mathrm{Tr}[S^* S] = \sum_j \lambda_j^2\,. \tag{2.29}$$

In particular, $\forall\, \varphi, \varphi', \psi, \psi' \in \mathcal{H}$,

$$\begin{aligned} (\,|\varphi\rangle\langle\psi|\,,\,|\varphi'\rangle\langle\psi'|\,)_{\mathcal{L}^2} &= (\varphi, \varphi')\cdot(\psi', \psi) \\ \|\,|\varphi\rangle\langle\psi|\,\|_{\mathcal{L}^2} &= \|\varphi\|\,\|\psi\|\,. \end{aligned} \tag{2.30}$$

Definition–Theorem 2.2.2 (integral kernels). Let $(M, \mathrm{d}\mu)$ be a measure space.

① Let $K \in L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$. Then the integral $\int_M K(x,y)\varphi(y)\,\mathrm{d}y$ is defined for μ-almost all x and represents a function in $L^2(M, \mathrm{d}\mu)$.

② Let $\mathcal{H} = L^2(M, \mathrm{d}\mu)$ and $S \in \mathcal{L}(\mathcal{H})$. Then $S \in \mathcal{L}^2(\mathcal{H})$ iff $\exists\, K_S \in L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$ such that

$$(S\varphi)(x) = \int_M K_S(x,y)\varphi(y)\,\mathrm{d}\mu(y) \qquad \forall\, \varphi \in L^2(M, \mathrm{d}\mu)\,. \tag{2.31}$$

If so, K_S is said the INTEGRAL KERNEL of S and

$$\|S\|_{\mathcal{L}^2(\mathcal{H})} = \|K_S\|_{L^2(M\times M, \mathrm{d}\mu\otimes\mathrm{d}\mu)}\,. \tag{2.32}$$

Also,

$$K_{cS+c'S'} = cK_S + c'K_{S'} \qquad \forall\, c, c' \in \mathbb{C}\,, \quad \forall\, S, S' \in \mathcal{L}^2(\mathcal{H}) \tag{2.33}$$

and the map

$$\begin{array}{rcl} L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu) & \longrightarrow & \mathcal{L}^2(\mathcal{H}) \\ K_S & \longmapsto & S \end{array} \tag{2.34}$$

is an isomorphism of Banach spaces.

③ If $K_1, K_2 \in L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$ and $K(x,y) := \int_M K_1(x,z)K_2(z,y)\,\mathrm{d}\mu(y)$, then $K \in L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$ as well. $L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$, when endowed with the $\bullet$-product

$$(K_1 \bullet K_2)(x,y) := \int_M K_1(x,z)K_2(z,y)\,\mathrm{d}\mu(z)\,, \tag{2.35}$$

is a Banach algebra. It is also a Banach-$*$ algebra, when endowed with the involution such that $K(x,y) \mapsto \overline{K(y,x)}$.

④ If $S, S_1, S_2 \in \mathcal{L}^2(\mathcal{H})$, then

$$\begin{array}{rcll} K_{S_1S_2} & = & K_{S_1} \bullet K_{S_2} & (\mu\otimes\mu)\text{-a.e.} \\ K_{S^*}(x,y) & = & \overline{K_S(y,x)} & (\mu\otimes\mu)\text{-a.e.} \end{array} \tag{2.36}$$

Thus, (2.34) is also a $*$-isomorphism of Banach $*$-algebras. In particular,

$$S = S^* \quad \Leftrightarrow \quad K_S(y,x) = \overline{K_S(x,y)} \qquad (\mu\otimes\mu)\text{-a.e.} \tag{2.37}$$

⑤ Let $\varphi, \psi \in \mathcal{H}$: then the trace class operator $|\varphi\rangle\langle\psi|$ has kernel $\varphi(x)\overline{\psi(y)}$. More generally, let $T \in \mathcal{L}^1(\mathcal{H})$, with canonical form $T = \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j|$, where $\lambda_j > 0$, $\sum_j \lambda_j = \|T\|_{\mathcal{L}^1(\mathcal{H})}$, and $\{\varphi_j\}_j$ and $\{\psi_j\}_j$ orthonormal systems of $\mathcal{H}$. Its kernel is

$$K_T(x,y) = \sum_j \lambda_j \varphi_j(x)\overline{\psi_j(y)}\,, \tag{2.38}$$

the expansion above converging in $L^2(M \times M, \mathrm{d}\mu \otimes \mathrm{d}\mu)$.

When the Hilbert space is realized as a L^2-space, it is useful to get the trace of a trace class operator directly from its kernel, instead of summing its singular values. This is customarily done by integrating the kernel along the diagonal, modulo the technicalities discussed here below.

One starts from the following classical result.

Theorem 2.2.3. Let $\mathcal{H} = L^2(M, \mathrm{d}\mu)$ as above and $T \in \mathcal{L}^1(\mathcal{H})$.

① Let $S_1, S_2 \in \mathcal{L}^2(\mathcal{H})$ such that $T = S_1 S_2$. In terms of the corresponding kernels, $K_T = K_{S_1} \bullet K_{S_2}$ $(\mu \otimes \mu)$-a.e. Then ([50])

$$\begin{aligned} \mathrm{Tr}[T] &= \int_M (K_{S_1} \bullet K_{S_2})(x,x)\, \mathrm{d}\mu(x) \\ &= \int_{M\times M} K_{S_1}(x,z) K_{S_2}(z,x)\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(z)\,. \end{aligned} \tag{2.39}$$

② In particular, if K_T is continuous,

$$\mathrm{Tr}[T] = \int_M K_T(x,x)\, \mathrm{d}\mu(x)\,. \tag{2.40}$$

(Sometimes (2.40) goes under the name of Mercer's theorem, at least when $M = [0,1]$ with the Lebesgue measure [107].)

One would be tempted to plug $K_T(x,x) = (K_{S_1} \bullet K_{S_2})(x,x)$ into (2.39), thus proving (2.40) in general. Yet, by (2.36), $K_T(x,y) = (K_{S_1} \bullet K_{S_2})(x,y)$ holds $(\mu \otimes \mu)$-a.e. in $M \times M$ and, in the absence of continuity for K_T, there is no guarantee that K_T and $K_{S_1} \bullet K_{S_2}$ agree on the diagonal $\{(x,x) \in M \times M\}$, which has zero $(\mu \otimes \mu)$-measure. A similar argument is the following. Write T in the canonical form $T = \sum_j \lambda_j |\varphi_j\rangle\langle\psi_j|$. In terms of kernels, $K_T(x,y) = \sum_j \lambda_j \varphi_j(x)\overline{\psi_j(y)}$ $(\mu \otimes \mu)$-a.e. One has

$$\mathrm{Tr}[T] = \sum_j (\psi_j, T\psi_j) = \sum_j \lambda_j (\psi_j, \varphi_j) = \int_M \sum_j \lambda_j \varphi_j(x)\overline{\psi_j(x)}\, \mathrm{d}\mu(x) \tag{2.41}$$

and again, in the absence of continuity for K_T, one cannot plug $K_T(x,x) = \sum_j \lambda_j \varphi_j(x)\overline{\psi_j(x)}$ in, in order to prove (2.40) in general.

The way out goes through the following remark. Since T is traceable,

$$\int_M \sum_j \lambda_j |\varphi(x)|^2 \mathrm{d}x = \int_M \sum_j \lambda_j |\psi(x)|^2 \mathrm{d}x = \sum_j \lambda_j < +\infty\,,$$

so the series $\sum_j \lambda_j |\varphi(x)|^2$ and $\sum_j \lambda_j |\psi(x)|^2$ are finite μ-a.e. Due to

$$\sum_j \big|\, \lambda_j \varphi_j(x)\overline{\psi_j(y)}\,\big| \leqslant \Big(\sum_j \lambda_j |\varphi(x)|^2\Big)^{1/2} \Big(\sum_j \lambda_j |\psi(x)|^2\Big)^{1/2},$$

series (2.38) converges absolutely $(\mu \otimes \mu)$-a.e. and $\sum_j \lambda_j \varphi_j(x)\overline{\psi_j(x)}$ converges absolutely μ-a.e. This pointwise limit, in general, is *not* $K_T(x,x)$, since $\{(x,x) \in M \times M\}$ has zero $(\mu \otimes \mu)$-measure: denote it by $\widetilde{K}_T(x,x)$. It turns out that $\widetilde{K}_T(x,x)$ is obtained from $K_T(x,y)$ by an averaging process over cubes centered on the diagonal in $M \times M$ and letting the side of the cubes go to zero. This can be regarded as a way of selecting a "smooth" pointwise representative for the corresponding kernel. Thus, (2.41) gives $\mathrm{Tr}[T] = \int_M \widetilde{K}_T(x,x)\, \mathrm{d}\mu(x)$. This analysis has been carried out in [22, 23]. In the concrete case $(M, \mathrm{d}\mu) = (\mathbb{R}^n, \mathrm{d}x)$ the result is as follows.

Theorem 2.2.4. Let $\mathcal{H} = L^2(\mathbb{R}^n)$ for some positive integer n. Let $T \in \mathcal{L}^1(\mathcal{H})$, with integral kernel $K_T \in L^2(\mathbb{R}^{2n})$. Let $C_r = [-r, r]^n$ be the n-dimensional cube of side $2r$ centred at the origin of $\mathbb{R}^n$ and let $|C_r|$ be its Lebesgue measure.

① The limit

$$\widetilde{K}_T(x,y) := \lim_{r\to 0} \frac{1}{|C_r|^2} \int_{C_r\times C_r} K_T(x+\xi, y+\eta)\, \mathrm{d}\xi\, \mathrm{d}\eta \tag{2.42}$$

exists $(\mathrm{d}x \otimes \mathrm{d}x)$-a.e. and

$$\widetilde{K}_T(x,y) = K_T(x,y) \qquad (\mathrm{d}x \otimes \mathrm{d}x)\text{-a.e.} \tag{2.43}$$

$\widetilde{K}_T$ is uniquely determined by K_T and it agrees with K_T at each point (x,y) of continuity of the latter.

② $\widetilde{K}_T(x,x)$ exists $\mathrm{d}x$-a.e.

③ T has trace

$$\mathrm{Tr}[T] = \int_M \widetilde{K}_T(x,x)\, \mathrm{d}x\,. \tag{2.44}$$

④ Let $S_1, S_2 \in \mathcal{L}^2(\mathcal{H})$ such that $T = S_1 S_2$. Then

$$\widetilde{K}_T(x,x) = (K_{S_1} \bullet K_{S_2})(x,x) \qquad \mathrm{d}x\text{-a.e.} \tag{2.45}$$

⑤ If, in addition, $T > \mathbb{O}$ and $T = \sum_j \lambda_j |\varphi_j\rangle\langle\varphi_j|$ is its canonical form, then

$$\widetilde{K}_T(x,x) = \sum_j \lambda_j |\varphi_j(x)|^2 \qquad \mathrm{d}x\text{-a.e.} \tag{2.46}$$

2.3 Partial trace and reduced density matrices

Let $\mathcal{H}$ and $\mathcal{K}$ be two separable Hilbert spaces. In the treatment of many-body systems one systematically makes use of the operation that sends density matrices on $\mathcal{H}\otimes\mathcal{K}$ to density matrices on $\mathcal{H}$, whose definition and properties are collected in this section.

Definition–Theorem 2.3.1 (partial trace). Let $\mathcal{H}$ and $\mathcal{K}$ be two separable Hilbert spaces. The PARTIAL TRACE with respect to $\mathcal{K}$ is the linear map

$$\begin{aligned} \mathcal{L}^1(\mathcal{H}\otimes\mathcal{K}) &\longrightarrow \mathcal{L}^1(\mathcal{H}) \\ T &\longmapsto \mathrm{Tr}_{\mathcal{K}}[\,T\,] \end{aligned} \tag{2.47}$$

such that, for an orthonormal basis $\{\xi_j\}_j$ of $\mathcal{K}$,

$$(\varphi, \mathrm{Tr}_{\mathcal{K}}[\,T\,]\psi)_{\mathcal{H}} = \sum_j (\varphi\otimes\xi_j, T\,\psi\otimes\xi_j)_{\mathcal{H}\otimes\mathcal{K}} \qquad \forall\,\varphi,\psi\in\mathcal{H}\,. \tag{2.48}$$

One proves that if (2.48) holds, this happens independently of the basis $\{\xi_j\}_j$. *Equivalently*, $\mathrm{Tr}_{\mathcal{K}}[\,T\,]$ is characterized by

$$\mathrm{Tr}\big[\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\cdot A\,\big] = \mathrm{Tr}\big[\,T\cdot(A\otimes\mathbb{1}_{\mathcal{K}})\,]\,\big] \qquad \forall\, A\in\mathcal{L}(\mathcal{H}) \tag{2.49}$$

where trace in the l.h.s. is on $\mathcal{L}^1(\mathcal{H})$, while trace in the r.h.s. is on $\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})$. Properties:

$$\mathrm{Tr}[\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\,] = \mathrm{Tr}[\,T\,] \qquad \forall\, T\in\mathcal{L}^1(\mathcal{H}) \tag{2.50}$$

$$T \geqslant \mathbb{O} \;\Rightarrow\; \mathrm{Tr}_{\mathcal{K}}[\,T\,] \geqslant \mathbb{O} \qquad \forall\, T\in\mathcal{L}^1(\mathcal{H}) \tag{2.51}$$

$$\mathrm{Tr}_{\mathcal{K}}[\,T_1\otimes T_2\,] = T_1\cdot\mathrm{Tr}[\,T_2\,] \qquad \forall\, T_1\in\mathcal{L}^1(\mathcal{H})\,,\ \forall\, T_2\in\mathcal{L}^1(\mathcal{K}) \tag{2.52}$$

In particular, by (2.50) and (2.51),

$$\gamma \text{ is a density matrix on } \mathcal{H}\otimes\mathcal{K} \;\Rightarrow\; \mathrm{Tr}_{\mathcal{K}}[\,\gamma\,] \text{ is a density matrix on } \mathcal{H}\,. \tag{2.53}$$

In passing, we quote an alternative, algorithmic definition of partial trace, which turns out to be equivalent to the previous one and is customarily used for finite rank operators (matrices).

Theorem 2.3.2 (**partial trace in terms of matrix elements**). Let $\{e_i\}_i$ and $\{f_j\}_j$ be any two orthonormal basis of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ respectively, so that $\{e_i \otimes f_j\}_{i,j}$ is an orthonormal basis of $\mathcal{H} \otimes \mathcal{K}$. Let $T \in \mathcal{L}^1(\mathcal{H} \otimes \mathcal{K})$ and denote its matrix elements by

$$t_{i',j';i,j} := (e_{i'} \otimes f_{j'}, T\, e_i \otimes f_j)_{\mathcal{H}\otimes\mathcal{K}}\,. \tag{2.54}$$

Let $\widetilde{T} := \mathrm{Tr}_{\mathcal{K}}[\,T\,]$ and denote its matrix elements by

$$\tilde{t}_{i',i} := (e_{i'}, \mathrm{Tr}_{\mathcal{K}}[\,T\,]\, e_i)_{\mathcal{H}}\,. \tag{2.55}$$

Consider the operators $T_{i',i} \in \mathcal{L}(\mathcal{K})$ defined by

$$(f_{j'}, T_{i',i}\, f_j)_{\mathcal{K}} := t_{i',j';i,j}\,. \tag{2.56}$$

Then $T_{i',i} \in \mathcal{L}^1(\mathcal{K})$ and

$$\tilde{t}_{i',i} = \mathrm{Tr}[\,T_{i',i}\,] \tag{2.57}$$

(the last trace being computed on $\mathcal{K}$). All this is independent of the choice of the basis.

When $\dim \mathcal{H} = n < \infty$ and $\dim \mathcal{K} = m < \infty$, last theorem reads as follows. Fixed the orthonormal basis $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^m$, T is the $nm \times nm$ matrix written in block form as

$$T = \begin{pmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,n} \\ T_{2,1} & T_{2,2} & \cdots & T_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n,1} & T_{n,2} & \cdots & T_{n,n} \end{pmatrix} \tag{2.58}$$

in the basis $\{e_1 \otimes f_1, e_1 \otimes f_2, \ldots, e_1 \otimes f_m, e_2 \otimes f_1, \ldots, e_n \otimes f_{m-1}, e_n \otimes f_m\}$ of $\mathcal{H} \otimes \mathcal{K}$, where the blocks $T_{i',i}$'s are in turn the $m \times m$ matrices

$$T_{i',i} = \begin{pmatrix} t_{i',1;i,1} & t_{i',1;i,2} & \cdots & t_{i',1;i,m} \\ t_{i',2;i,1} & t_{i',2;i,2} & \cdots & t_{i',2;i,m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i',m;i,1} & t_{i',m;i,2} & \cdots & t_{i',m;i,m} \end{pmatrix}. \tag{2.59}$$

Then in the basis $\{e_i\}_{i=1}^n$, $\mathrm{Tr}_{\mathcal{K}}[\,T\,]$ is the $n \times n$ matrix

$$\mathrm{Tr}_{\mathcal{K}}[\,T\,] = \begin{pmatrix} \tilde{t}_{1,1} & \tilde{t}_{1,2} & \cdots & \tilde{t}_{1,n} \\ \tilde{t}_{2,1} & \tilde{t}_{2,2} & \cdots & \tilde{t}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{t}_{n,1} & \tilde{t}_{n,2} & \cdots & \tilde{t}_{n,n} \end{pmatrix} \qquad \text{with} \qquad \tilde{t}_{i',i} = \mathrm{Tr}[\,T_{i',i}\,] = \sum_{\ell=1}^{m} t_{i',\ell;i,\ell}\,. \tag{2.60}$$

When the Hilbert space under consideration is realized as $L^2(\mathbb{R}^n)$, the operation of partial trace takes a peculiar form in terms of kernels. Let us describe it in the case of density matrices, being the objects of interest in the sequel. In fact, (2.53) makes the following definition well-posed.

Definition 2.3.3. Let $\mathcal{H}$ and $\mathcal{K}$ be two separable Hilbert spaces and let γ be a density matrix on $\mathcal{H}\otimes\mathcal{K}$. The density matrix $\gamma^{(\mathcal{H})} := \mathrm{Tr}_{\mathcal{K}}[\,\gamma\,]$ is called REDUCED DENSITY MATRIX (or MARGINAL) of γ relative to $\mathcal{H}$ taken tracing out with respect to $\mathcal{K}$.

Theorem 2.3.4. Let $\mathcal{H} = L^2(\mathbb{R}^n)$ and $\mathcal{K} = L^2(\mathbb{R}^m)$ for some positive integers n, m, where the Lebesgue measure is understood. Let γ be a density matrix on $\mathcal{H}\otimes\mathcal{K} \cong L^2(\mathbb{R}^{n+m})$. Denote its kernel by $\gamma(x,x';y,y')$, where x and y run in $\mathbb{R}^n$ and x' and y' run in $\mathbb{R}^m$. Then $\gamma^{(\mathcal{H})} = \mathrm{Tr}_{\mathcal{K}}[\,\gamma\,]$ has kernel

$$\gamma^{(\mathcal{H})}(x,y) = \int_{\mathbb{R}^m} \gamma(x,z';y,z')\,\mathrm{d}z' \tag{2.61}$$

the integral being defined with the same technique as in theorem 2.2.4.

Equation (2.61) is actually how the reduced density matrix was introduced first by Husimi [60].

To conclude this section, we point out that this construction is realized precisely in view of the key operation

$$\mathcal{L}(\mathcal{H}) \ni \mathcal{O}^* = \mathcal{O} \longmapsto \mathrm{Tr}[\,\gamma\,\mathcal{O}\,] \in \mathbb{R}\,, \tag{2.62}$$

that in Quantum Mechanics is interpreted as the evaluation of the observable $\mathcal{O}$ (a bounded selfadjoint operator on $\mathcal{H}$) in the state γ (a density matrix on $\mathcal{H}$) – see the discussion at the beginning of Sec. 3.1.

When $\mathcal{H} = L^2(\mathbb{R}^n)$, then $\mathrm{Tr}[\,\gamma\,\mathcal{O}\,]$ is given tracing the kernel of $\gamma\,\mathcal{O} \in \mathcal{L}^1(\mathcal{H})$, according to theorems 2.2.3, 2.2.4. In particular, if $\mathcal{O} \in \mathcal{L}^2(\mathcal{H})$, with kernel $\mathcal{O}(\,\cdot\,,\,\cdot\,) \in L^2(\mathbb{R}^{2n})$, then a straightforward application of (2.35) and (2.40) gives

$$\mathrm{Tr}[\,\gamma\,\mathcal{O}\,] = \int_{\mathbb{R}^{2n}} \gamma(x,y)\mathcal{O}(y,x)\,\mathrm{d}x\,\mathrm{d}y\,. \tag{2.63}$$

If γ is the pure state $\gamma = |\psi\rangle\langle\psi|$, this becomes

$$\mathrm{Tr}[\,|\psi\rangle\langle\psi|\,\mathcal{O}\,] = (\psi, \mathcal{O}\,\psi) = \int_{\mathbb{R}^{2n}} \overline{\psi(x)}\mathcal{O}(x,y)\psi(y)\,\mathrm{d}x\,\mathrm{d}y\,. \tag{2.64}$$

Now, a typical physical situation is when the Hilbert space is isomorphic to $\mathcal{H}\otimes\mathcal{K}$ with $\mathcal{H} = L^2(\mathbb{R}^n)$, $\mathcal{K} = L^2(\mathbb{R}^m)$ and the observables of interest act non-trivially only on the subspace $\mathcal{H}\otimes\mathbb{1}_{\mathcal{K}}$, i.e., have the form $\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}$ with $\mathcal{O} = \mathcal{O}^* \in \mathcal{L}(\mathcal{H})$. Notice that such a $\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}$ cannot be Hilbert-Schmidt and, hence, cannot have a $L^2(\mathbb{R}^{2(n+m)})$-kernel, because

$$\mathrm{Tr}[\,(\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}})^*(\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}})\,] = \mathrm{Tr}[\,\mathcal{O}^*\mathcal{O}\,]\cdot\mathrm{Tr}[\mathbb{1}_{\mathcal{K}}] = +\infty \tag{2.65}$$

even if $\mathcal{O}$ is Hilbert-Schmidt. Yet, such a $\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}$ is recognised to have a kernel-like action whenever $\mathcal{O} \in \mathcal{L}^2(\mathcal{H})$. Indeed, since on vectors $\varphi\otimes\varphi' \in \mathcal{H}\otimes\mathcal{K}$ it acts as

$$(\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}})(\varphi\otimes\varphi') = \mathcal{O}\varphi\otimes\varphi'\,,$$

then it is easily seen to act on a generic wave function $\psi \in \mathcal{H}\otimes\mathcal{K} \cong L^2(\mathbb{R}^{n+m})$ as

$$\psi(x;x') \longmapsto \int_{\substack{y\in\mathbb{R}^n\\ y'\in\mathbb{R}^m}} \mathcal{O}(x,y)\,\delta(x'-y')\,\psi(y;y')\,\mathrm{d}y\,\mathrm{d}y' \tag{2.66}$$

namely, with integral kernel

$$K_{\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}}(x,x';y,y') = \mathcal{O}(x,y)\,\delta(x'-y')\,. \tag{2.67}$$

Thus, by general formula (2.63), the expectation of the observable $\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}$ in the state γ on $\mathcal{H}\otimes\mathcal{K}$ is given in terms of kernels by

$$\begin{aligned}
\mathrm{Tr}[\,\gamma\,(\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}})\,] &= \int_{\substack{x,y\in\mathbb{R}^n\\ x',y'\in\mathbb{R}^m}} \gamma(x,x';y,y')\,K_{\mathcal{O}\otimes\mathbb{1}_{\mathcal{K}}}(y,y';x,x')\,\mathrm{d}x\,\mathrm{d}x'\mathrm{d}y\,\mathrm{d}y' \\
&= \int_{\substack{x,y\in\mathbb{R}^n\\ x',y'\in\mathbb{R}^m}} \gamma(x,x';y,y')\,\mathcal{O}(y,x)\,\delta(y'-x')\,\mathrm{d}x\,\mathrm{d}x'\mathrm{d}y\,\mathrm{d}y' \\
&= \int_{x,y\in\mathbb{R}^n} \gamma^{(\mathcal{H})}(x,y)\,\mathcal{O}(y,x)\,\mathrm{d}x\,\mathrm{d}y \\
&= \mathrm{Tr}[\,\gamma^{(\mathcal{H})}\mathcal{O}\,]
\end{aligned} \tag{2.68}$$

where

$$\gamma^{(\mathcal{H})}(x,y) := \int_{\mathbb{R}^m} \gamma(x,z';y,z')\,\mathrm{d}z' \tag{2.69}$$

is the kernel of the reduced density matrix $\gamma^{(\mathcal{H})}$ – compare (2.68) with (2.49).

Of course, theorem 2.3.4 and the subsequent considerations can be easily restated and hold true also when $\mathcal{H} = L^2(\Omega)$ and $\mathcal{K} = L^2(\Lambda)$ for some domains $\Omega\subset\mathbb{R}^n$ and $\Lambda\subset L^2(\mathbb{R}^m)$.

2.4 Distances of states and of reduced density matrices

This is a detour section where we prepare some results the application of which, in the context of B.E.C., will be not discussed until Sec. 7.2. Due to their generality, it is preferable to present them next to the previous sections which deals with the general theory. The purpose is to compare distances among states with distances among the corresponding marginals.

First, we focus on *pure* states. This is to point out that, although there is a one-to-one correspondence between pure states and rays of the Hilbert space, yet some care is needed when, to measure distances among two pure states, one chooses trace-norm for density matrices or vector-norm for the representatives of the rays.

Theorem 2.4.1. Let $\mathcal{H}$ be a separable Hilbert space. Let $\psi_1,\psi_2\in\mathcal{H}$ with $\|\psi_1\| = \|\psi_2\| = 1$ and set $\gamma_i := |\psi_i\rangle\langle\psi_i|$ for $i=1,2$. Then

$$\|\,\psi_1-\psi_2\,\|^2_{\mathcal{H}} = 2\big(1-\mathfrak{Re}(\psi_1,\psi_2)\big) \tag{2.70}$$

$$\|\,\gamma_1-\gamma_2\,\|^2_{\mathcal{L}^1(\mathcal{H})} = 4\big(1-|(\psi_1,\psi_2)|^2\big) \tag{2.71}$$

$$\|\,\gamma_1-\gamma_2\,\|_{\mathcal{L}^1(\mathcal{H})} \leqslant 2\,\|\,\psi_1-\psi_2\,\|_{\mathcal{H}} \tag{2.72}$$

and the constant in (2.72) is optimal.

Proof. (2.70) is obvious.

To prove (2.71) notice first that $\gamma_1-\gamma_2$ acts non-trivially only on the subspace generated by ψ_1,ψ_2. So set $\mathcal{H}_2 := \mathrm{Span}_{\mathbb{C}}\{\psi_1,\psi_2\}$, whence $\mathcal{H} = \mathcal{H}_2\oplus\mathcal{H}_2^{\perp}$, $\gamma_1-\gamma_2 = (\gamma_1-\gamma_2)\big|_{\mathcal{H}_2}$, $(\gamma_1-\gamma_2)\big|_{\mathcal{H}_2^{\perp}} = \mathbb{O}$. Moreover, notice that

$$(\gamma_1-\gamma_2)^2 = \big(1-|(\psi_1,\psi_2)|^2\big)\,\mathbb{1}\,, \tag{2.73}$$

thus, $|\gamma_1 - \gamma_2| = \sqrt{(\gamma_1 - \gamma_2)^2}$ in diagonal form reads

$$\Big|(\gamma_1 - \gamma_2)\big|_{\mathcal{H}_2}\Big| = \sqrt{1 - |(\psi_1, \psi_2)|^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{2.74}$$

This can be seen by direct computation. Indeed

$$\begin{aligned} (\gamma_1 - \gamma_2)^2 &= (\, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \,) \cdot (\, |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| \,) \\ &= |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| - (\psi_1, \psi_2)|\psi_1\rangle\langle\psi_2| - (\psi_2, \psi_1)|\psi_2\rangle\langle\psi_1| \,, \end{aligned} \tag{2.75}$$

then, in case $|(\psi_1, \psi_2)| < 1$ (while (2.74) is trivially true if $|(\psi_1, \psi_2)| = 1$), one expands $\psi \in \mathcal{H}$ as

$$\psi = a_1\, \psi_1 + a_2\, \psi_2 + \sum_{i \geqslant 3} a_i\, \psi_i \tag{2.76}$$

where $\{\psi_3, \psi_4, \dots\}$ is an orthonormal basis of $\mathcal{H}_2^{\perp}$. Applying (2.75) to (2.76) one gets

$$(\gamma_1 - \gamma_2)^2\, \psi = \big(1 - |(\psi_1, \psi_2)|^2\big)\, \psi \,. \tag{2.77}$$

Alternatively, orthonormalisation of the basis $\{\psi_1, \psi_2\}$ of $\mathcal{H}_2$ gives

$$\begin{aligned} \varphi_1 &= \psi_1 \\ \varphi_2 &= \frac{\psi_2 - (\psi_1, \psi_2)\, \psi_2}{\|\psi_2 - (\psi_1, \psi_2)\, \psi_2\|} \end{aligned} \tag{2.78}$$

and in the new basis $\{\varphi_1, \varphi_2\}$ the two projections read

$$\begin{aligned} \gamma_1\big|_{\mathcal{H}_2} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ \gamma_2\big|_{\mathcal{H}_2} &= \begin{pmatrix} |(\psi_1, \psi_2)|^2 & (\psi_1, \psi_2)\sqrt{1 - |(\psi_1, \psi_2)|^2} \\ \overline{(\psi_1, \psi_2)}\sqrt{1 - |(\psi_1, \psi_2)|^2} & 1 - |(\psi_1, \psi_2)| \end{pmatrix} . \end{aligned} \tag{2.79}$$

Computing the eigenvalues λ_1, λ_2 of $(\gamma_1 - \gamma_2)^2$ in this basis gives

$$\lambda_1 = \lambda_2 = 1 - |(\psi_1, \psi_2)|^2 \,, \tag{2.80}$$

thus, $|\gamma_1 - \gamma_2| = \sqrt{(\gamma_1 - \gamma_2)^2}$ in diagonal form is exactly (2.74). As a consequence,

$$\| \, \gamma_1 - \gamma_2 \, \|_{\mathcal{L}^1(\mathcal{H})} = \mathrm{Tr}|\gamma_1 - \gamma_2| = 2\sqrt{1 - |(\psi_1, \psi_2)|^2} \,, \tag{2.81}$$

which proves (2.71).

To prove (2.72), look for an optimal $\alpha > 0$ such that

$$\| \, \gamma_1 - \gamma_2 \, \|_{\mathcal{L}^1(\mathcal{H})} \leqslant \alpha \, \| \, \psi_1 - \psi_2 \, \| \,. \tag{2.82}$$

By plugging (2.70) and (2.71) into (2.82) squared, and setting $x := \mathfrak{Re}(\psi_1, \psi_2)$ and $y := \mathfrak{Im}(\psi_1, \psi_2)$, one has

$$f_\alpha(x, y) := x^2 - \frac{\alpha^2}{2} x + \frac{\alpha^2}{2} - 1 + y^2 \geqslant 0 \tag{2.83}$$

to be minimized on the circle $\mathcal{C} = \{x^2 + y^2 \leqslant 1\}$ of $\mathbb{R}^2$. Since

$$\left\{ \begin{array}{l} 0 = \partial_x f_\alpha(x,y) = 2x - \frac{\alpha^2}{2} \\ 0 = \partial_y f_\alpha(x,y) = 2y \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l} x = \frac{\alpha^2}{4} \\ y = 0 \end{array} \right. \tag{2.84}$$

then, if $\alpha < 2$, f_α reaches its minimum in $(\frac{\alpha^2}{4}, 0) \in \mathcal{C}$; yet, $f_\alpha(\frac{\alpha^2}{4}, 0) = -1 + \frac{\alpha^2}{4} < 0$, then there exist norm-one vectors $\widetilde{\psi}_1, \widetilde{\psi}_2 \in \mathcal{H}$ such that the corresponding projections violates (2.82), that is, $\| \widetilde{\gamma}_1 - \widetilde{\gamma}_2 \|_{\mathcal{L}^1(\mathcal{H})} > \alpha \| \widetilde{\psi}_1 - \widetilde{\psi}_2 \|$. On the other hand, if $\alpha \geqslant 2$ then $(\frac{\alpha^2}{4}, 0)$ does not belong to the interior of $\mathcal{C}$, and f_α must achieve its minimum in the boundary $\partial\mathcal{C}$, actually in the point $(1, 0)$:

$$f_\alpha(x,y)\big|_{\partial\mathcal{C}} = f_\alpha(x, \pm\sqrt{1-x^2}) = \frac{\alpha^2}{2}(1-x) \geqslant f_\alpha(1,0) = 0\,. \tag{2.85}$$

So the optimal α in order (2.83) to be satisfied $\forall\,(x,y) \in \mathcal{C}$ is $\alpha = 2$. □

Notice, in passing, that a non optimal bound like (2.72) could be easily obtained in the following direct way. First, since $\forall\,\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$

$$\begin{aligned} \| (\psi_1,\varphi)\,\psi_1 - (\psi_2,\varphi)\,\psi_2 \| &\leqslant \| (\psi_1,\varphi)\,\psi_1 - (\psi_2,\varphi)\,\psi_1 \| + \| (\psi_2,\varphi)\,\psi_1 - (\psi_2,\varphi)\,\psi_2 \| \\ &= |\,(\psi_1,\varphi) - (\psi_2,\varphi)\,| \cdot \|\psi_1\| + |\,(\psi_2,\varphi)\,| \cdot \| \psi_1 - \psi_2 \| \\ &\leqslant 2 \| \psi_1 - \psi_2 \|\,, \end{aligned} \tag{2.86}$$

then

$$\| \gamma_1 - \gamma_2 \|_{\mathcal{L}(\mathcal{H})} = \sup_{\|\varphi\|=1} \| (\psi_1,\varphi)\,\psi_1 - (\psi_2,\varphi)\,\psi_2 \| \leqslant 2 \| \psi_1 - \psi_2 \|\,. \tag{2.87}$$

Next observe that

$$\begin{aligned} \| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})} &= \sqrt{\lambda_1} + \sqrt{\lambda_2} \,\leqslant\, 2\,\max\{\sqrt{\lambda_1}, \sqrt{\lambda_2}\} \,=\, 2\,\sqrt{\max\{\lambda_1,\lambda_2\}} \\ &= 2\,\sqrt{\| (\gamma_1 - \gamma_2)^2 \|_{\mathcal{L}(\mathcal{H})}} \,\leqslant\, 2\,\sqrt{\| \gamma_1 - \gamma_2 \|^2_{\mathcal{L}(\mathcal{H})}} \\ &= 2 \| \gamma_1 - \gamma_2 \|_{\mathcal{L}(\mathcal{H})} \end{aligned} \tag{2.88}$$

whence

$$\| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})} \leqslant 4 \| \psi_1 - \psi_2 \|\,. \tag{2.89}$$

Yet (2.89) provides a non-sharp constant.

By theorem 2.4.1, if two unit vectors in $\mathcal{H}$ are close, so are the corresponding projections. The converse is false. For, if e.g. $\psi_2 = -\psi_1$ then $\| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})} = 0$ while $\| \psi_1 - \psi_2 \| = \sqrt{2}$. Actually close rank-one projections (i.e., pure states) correspond to close *rays*, instead of representatives, each unit vector in $\mathcal{H}$ determining a ray up to a phase that cancels in the corresponding projection. This is the content of the following definition 2.4.2 and theorem 2.4.3.

Definition–Theorem 2.4.2 (distance of rays). Let $\mathcal{H}$ be a separable Hilbert space and denote by $\{\psi\}$ a ray, where $\psi \in \mathcal{H}$ is any norm-one representative of it. In terms of norm-one representatives, the two rays coincide and one writes $\{\psi_1\} = \{\psi_2\}$, iff ψ_1 and ψ_2 are the same vector in $\mathcal{H}$ up to a phase. Define the distance between two rays to be the quantity

$$d\big(\{\psi_1\},\{\psi_2\}\big) := \min_{\delta\in[0,2\pi]} \big\| \psi_1 - e^{i\delta}\psi_2 \big\|_{\mathcal{H}}\,. \tag{2.90}$$

The min is well-posed, since the function $\delta \mapsto \| \psi_1 - e^{i\delta}\psi_2 \|_{\mathcal{H}}$ is clearly continuous on the compact $[0, 2\pi]$. This is a true distance making the set of rays on $\mathcal{H}$ a metric space:

$$\begin{array}{rcl} d\left(\{\psi_1\},\{\psi_2\}\right) & = & 0 \qquad \Leftrightarrow \quad \{\psi_1\} = \{\psi_2\} \\ d\left(\{\psi_1\},\{\psi_2\}\right) & = & d\left(\{\psi_2\},\{\psi_1\}\right) \\ d\left(\{\psi_1\},\{\psi_2\}\right) & \leqslant & d\left(\{\psi_1\},\{\psi_3\}\right) + d\left(\{\psi_3\},\{\psi_2\}\right). \end{array} \tag{2.91}$$

Theorem 2.4.3 (**close pure state projections ⇒ close rays**)**.** Let $\mathcal{H}$ be a separable Hilbert space, let $\psi_1, \psi_2 \in \mathcal{H}$ with $\|\psi_1\| = \|\psi_2\| = 1$, representing the rays $\{\psi_1\}, \{\psi_2\}$, and set $\gamma_i := |\psi_i\rangle\langle\psi_i|$ for $i = 1, 2$. Then

$$d\left(\{\psi_1\},\{\psi_2\}\right) \leqslant \frac{1}{\sqrt{2}} \| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})} \tag{2.92}$$

the constant being optimal.

Proof. One looks for an optimal $\beta > 0$ such that

$$d\left(\{\psi_1\},\{\psi_2\}\right) \leqslant \beta \| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})}. \tag{2.93}$$

Notice preliminarily that with no loss of generality one can change ψ_2 by a phase $e^{i\eta}$ so that (ψ_1, ψ_2) is *real*: indeed both $\|\gamma_1 - \gamma_2\|_{\mathcal{L}^1(\mathcal{H})}$ and $d\left(\{\psi_1\},\{\psi_2\}\right)$ are left unchanged by (2.71) and (2.90), respectively. So, by (2.70),

$$\begin{aligned} \left[\, d\left(\{\psi_1\},\{\psi_2\}\right) \right]^2 &= \min_{\delta\in[0,2\pi]} \left\| \psi_1 - e^{i\delta}\psi_2 \right\|_{\mathcal{H}}^2 \\ &= 2 \min_{\delta\in[0,2\pi]} \left(1 - \mathfrak{Re}\left[e^{i\delta}(\psi_1, \psi_2)\right]\right) \\ &= 2 \min_{\delta\in[0,2\pi]} (1 - x\cos\delta) \\ &= 2\,(1 - x\,\mathrm{sign}(x)) \end{aligned} \tag{2.94}$$

while, by (2.71),

$$\| \gamma_1 - \gamma_2 \|_{\mathcal{L}^1(\mathcal{H})}^2 = 4\,(1 - x^2) \tag{2.95}$$

with $x := \mathfrak{Re}(\psi_1, \psi_2) = (\psi_1, \psi_2) \in [-1, 1]$. It is then clear that

$$\begin{array}{c} 2\,(1 - x\,\mathrm{sign}(x)) \leqslant 4\,\beta^2\,(1 - x^2) \\ \forall\, x \in [-1, 1] \end{array} \qquad \Rightarrow \qquad \beta \geqslant \frac{1}{\sqrt{2}} \tag{2.96}$$

and $\beta = 1/\sqrt{2}$ is the sharp constant. □

Next, we come to the main point: *partial trace of density matrices are close if the density matrices themselves are*, the converse being false. In Sec. 7.2 this will be used to claim that if two N-body states are close, then the same holds for the corresponding one-particle reduced density matrices. *Distance* between the two states one starts with is meant to be the trace-norm of the difference of density matrices; if, in particular, one deals with pure states, then one indifferently means closeness of vectors, rays or density matrices in the spirit of theorems 2.4.1 and 2.4.3.

Theorem 2.4.4. Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces. Then, $\forall\, T \in \mathcal{L}^1(\mathcal{H} \otimes \mathcal{K})$,

$$\left\| \mathrm{Tr}_{\mathcal{K}}[\, T\,] \right\|_{\mathcal{L}^1(\mathcal{H})} \leqslant \| T \|_{\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})}. \tag{2.97}$$

Proof. From one side, recall from theorem 2.1.1, point ⑤, that by the canonical isometric isomorphism $(\mathcal{L}^1(\mathcal{H}), \|\cdot\|_{\mathcal{L}^1}) = (\mathrm{Com}(\mathcal{H}), \|\cdot\|_{\mathcal{L}})^*$ one has

$$\big\|\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\,\big\|_{\mathcal{L}^1(\mathcal{H})} = \sup_{\substack{C\in\mathrm{Com}(\mathcal{H})\\ C\neq\mathbb{0}}} \frac{\big|\,\mathrm{Tr}\,[\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\cdot C\,]\,\big|}{\|C\|_{\mathcal{L}(\mathcal{H})}}\,. \tag{2.98}$$

From the other side, by characterization (2.49) of partial trace, and by (2.6),

$$\big|\,\mathrm{Tr}\,[\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\cdot C\,]\,\big| = \big|\,\mathrm{Tr}\,[\,T\cdot(C\otimes\mathbb{1}_{\mathcal{K}})\,]\,\big| \leqslant \|C\|_{\mathcal{L}(\mathcal{H})}\cdot\|\,T\,\|_{\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})}\,. \tag{2.99}$$

So (2.97) is proved. □

One can alternatively prove the key inequality (2.97) by controlling partial trace in terms of matrix elements, according to theorem 2.3.2. This gives an explicit indication of how the inequality can be strict – see (2.103).

Alternative proof of theorem 2.4.4. Diagonalize T so that in its block form (2.58) each block (2.59) is

$$T_{i',i} = \delta_{i',i}\begin{pmatrix} \lambda_1^{(i)} & 0 & 0 & \cdots \\ 0 & \lambda_2^{(i)} & 0 & \cdots \\ 0 & 0 & \lambda_3^{(i)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.100}$$

and

$$\|\,T\,\|_{\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})} = \sum_i\sum_j \big|\,\lambda_j^{(i)}\,\big|\,. \tag{2.101}$$

By (2.60)

$$\mathrm{Tr}_{\mathcal{K}}[\,T\,] = \begin{pmatrix} \mu_1 & 0 & \cdots \\ 0 & \mu_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}\,, \qquad \mu_i = \mathrm{Tr}[\,T_{i,i}\,] = \sum_j \lambda_j^{(i)}\,, \tag{2.102}$$

and

$$\big\|\,\mathrm{Tr}_{\mathcal{K}}[\,T\,]\,\big\|_{\mathcal{L}^1(\mathcal{H})} = \sum_i |\mu_i| \leqslant \sum_i\sum_j \big|\,\lambda_j^{(i)}\,\big| = \|\,T\,\|_{\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})} \qquad \square \tag{2.103}$$

Our main application of theorem 2.4.4 (see theorem 7.2.3) will be in the form

$$\big\|\,\mathrm{Tr}_{\mathcal{K}}[\,\gamma_1\,] - \mathrm{Tr}_{\mathcal{K}}[\,\gamma_2\,]\,\big\|_{\mathcal{L}^1(\mathcal{H})} \leqslant \big\|\,\gamma_1-\gamma_2\,\big\|_{\mathcal{L}^1(\mathcal{H}\otimes\mathcal{K})} \tag{2.104}$$

for some states γ_1, γ_2 on $\mathcal{H}\otimes\mathcal{K}$. (2.104) follows from (2.97) by linearity of $\mathrm{Tr}_{\mathcal{K}}$. In case of *pure* states it will be useful to use (2.104) in the form

$$\big\|\,\mathrm{Tr}_{\mathcal{K}}[\,\gamma_1\,] - \mathrm{Tr}_{\mathcal{K}}[\,\gamma_2\,]\,\big\|_{\mathcal{L}^1(\mathcal{H})} \leqslant 2\,\|\,\psi_1-\psi_2\,\|_{\mathcal{H}\otimes\mathcal{K}} \tag{2.105}$$

To conclude this section, let us quote a useful inequality for the sequel, which allows to control the distance between two pure states in terms of the distance between the first one and an unnormalised version of the second.

Theorem 2.4.5. Let $\psi, \varphi \in \mathcal{H}$, with $\|\psi\| = 1$ and $\|\psi - \varphi\| < 1$. Then

$$\left\| \psi - \frac{\varphi}{\|\varphi\|} \right\| \leqslant \frac{2\,\|\psi - \varphi\|}{1 - \|\psi - \varphi\|}\,. \tag{2.106}$$

Proof. Since $\big|\, \|\psi\| - \|\varphi\| \,\big| \leqslant \|\psi - \varphi\|$ and, consequently, $1 - \|\psi - \varphi\| \leqslant \|\varphi\| \leqslant 1 + \|\psi - \varphi\|$, then

$$\begin{aligned} \left\| \psi - \frac{\varphi}{\|\varphi\|} \right\| &= \frac{\big\| \, \|\varphi\|\, \psi - \varphi \, \big\|}{\|\varphi\|} = \frac{\big\| \, (\, \|\varphi\| - 1\,)\, \psi + (\psi - \varphi) \, \big\|}{\|\varphi\|} \\ &\leqslant \frac{\big|\, \|\varphi\| - \|\psi\| \,\big| \cdot \|\psi\| + \|\psi - \varphi\|}{1 - \|\psi - \varphi\|} \leqslant \frac{2\,\|\psi - \varphi\|}{1 - \|\psi - \varphi\|}\,. \end{aligned} \tag{2.107}$$

□

Chapter 3

Formalisation of Bose-Einstein condensation

The mathematical investigation on B.E.C. this work is dealing with relies on a suitable rigorous definition of such a physical phenomenon, that is, a formalisation which is mathematically well-posed, has a precise physical interpretation in terms of first principles, and is experimentally acceptable in the sense that it provides a clear signature of condensation to be checked in the laboratory. Depending on the point one stresses (mathematics, theoretical interpretation, or experimental perspective), a complex of definitions of B.E.C. has been established in the literature since Bose and Einstein's "theoretical discover".

Following Ref. [83, 84], this chapter focuses on the appropriate mathematical tools leading to a well-posed and physically meaningful formalisation of B.E.C., while the experimental scenario, which in the last 12 years has become more and more rich, has been already briefly sketched in the warm up of Chap. 1. Here the discussion is based on a definition of B.E.C. which has been essentially well-established since half a century, but which still deserves an appropriate interpretation.

Along this line, Sec. 3.1 and 3.2 provide the mathematical framework for the description of large many-body boson systems. The *bosonic sector* of a Hilbert space will be introduced and described with suitable quantum labels. The need and the meaning of a limit of infinite number of particle will be discussed, bridging the mathematical treatment with its physical interpretation. In particular, stress will be put on the *scaling* nature of any physically reasonable and mathematically well-posed limit of infinite systems.

The choice here is to place all the discussion within the standard quantum mechanical picture, that is, without appealing to the formalism of second quantization. Both being formalisms, the essence of the Physics behind is untouched. Moreover, both are standard textbooks subjects. Admittedly, the second quantization would definitely avoid some unnecessary complications in the notation, for instance, the emergence of combinatorial factors or the use of labels for vectors in the many-body Hilbert space (the use of the number operator $N(\varphi) = a^*(\varphi)a(\varphi)$ for one-particle state φ already does the job). On the other side, most of the mathematical results and techniques presented in this chapter and in the following ones do not involve second quantization or, at least, most of them has been originally stated and derived in the Schrödinger formalism, and this motivates the present choice.

In Sec. 3.3 and 3.4 we will revisit the classical Penrose and Osanger's definition

within our current formalism for an *asymptotic analysis*, i.e., in the limit of infinitely many particles. Emphasis will be given to the crucial role that reduced density matrices have in expressing condensation: in fact, the distribution of particles one observes in the experiments is nothing but the diagonal part of the kernel of the one-body reduced density matrix. To control (asymptotically) 100% B.E.C., such one-body marginal has to approximate the projection onto the one-body condensate wave function: *convergence* can be set in several natural but distinct topologies and we accounts for the *equivalence* of a number of them.

Then in Sec. 3.5 we report two recent proofs of B.E.C. in the sense of the preceding discussion, in the Gross-Pitaevskiĭ scaling limit for a dilute system of weakly interacting spinless bosons.

3.1 Many-body quantum systems of identical bosons

Notoriously, the mathematical apparatus outlined in Chap. 2 is the standard tool at the basis of the description of a physical system in Quantum Mechanics with a finite number of degrees of freedom. *States* of the systems are density matrices on some Hilbert space $\mathcal{H}$: pure states are rays of $\mathcal{H}$, while mixed states are density matrices with rank $\geqslant 2$. *Observables* are (possibly unbounded) selfadjoint operators on $\mathcal{H}$ and the expectation value of an observable $\mathcal{O}$ in the state γ is the real number $\mathrm{Tr}[\,\gamma\,\mathcal{O}\,]$. *Dynamics* is given by a strongly continuous one-parameter group $\{e^{-itH/\hbar}\}_{t\in\mathbb{R}}$ of unitary operators on $\mathcal{H}$ generated by the selfadjoint H, the Hamiltonian of the system: a state γ evolves into $\gamma_t = e^{-itH/\hbar}\,\gamma\,e^{itH/\hbar}$. In particular, a pure state ψ evolves into $\psi_t = e^{-itH/\hbar}\,\psi$ (the infinitesimal version of which, is the Schrödinger equation), ψ being any norm 1 vector in $\mathcal{H}$ representing the ray $\{\psi\}$.

Here the interest is towards special MANY-BODY QUANTUM SYSTEMS: systems of N undistinguishable quantum particles. Thus, the building block is the Hilbert space $\mathcal{H}$ of a single particle, while the N-body Hilbert space is $\mathcal{H}_N := \mathcal{H}^{\otimes N}$. Actually the Spin-Statistics theorem, which is a postulate in this non-relativistic setting, prescribes that boson states in $\mathcal{H}_N$ are symmetric with respect to permutation of the particles, while fermion states are antisymmetric. In this section we will recall and develop the tools for the study of the bosonic case, which is the one we are going to consider in the sequel concerning Bose-Einstein condensation.

Notation. By $\mathcal{H}$ one will denote the single-particle Hilbert space. In view of our next discussion on *spinless* particles, we will systematically choose $\mathcal{H} = L^2(\Omega)$ for some $\Omega \subset \mathbb{R}^d$: here Ω is the region of the d-dimensional Euclidean space where each particle is allowed to stay ($d = 1, 2, 3$, for the cases of interest), be it the region of the laboratory where an alkali gas is trapped and cooled or the region of space occupied by a star of bosons. Ω will be chosen to be suitably regular, possibly the whole $\mathbb{R}^d$. Thus, each quantum particle is described by an infinite-dimensional Hilbert space. Yet the very same results that we are going to discuss do hold, with a slightly simplified notation, even in the finite dimensional case $\mathcal{H} \cong \mathbb{C}^n$. By $\mathcal{H}_N = \mathcal{H}^{\otimes N} \cong L^2(\Omega^N)$ one will denote the N-body Hilbert space, where $\Omega^N = \Omega \times \cdots \times \Omega \subset \mathbb{R}^{Nd}$ is the N-th Cartesian product. Starting from any orthonormal basis of $\mathcal{H}$, say $\{\varphi_j\}_j$, one builds the orthonormal basis

$$\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \ldots, j_N} \tag{3.1}$$

of $\mathcal{H}_N$. For shortness, vectors of such a particular collection will be referred to as the BASIS VECTORS of $\mathcal{H}_N$, with no further specification. In this notation, the basis vector

$$\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N} \tag{3.2}$$

of $\mathcal{H}_N$ is the state of N particles, labeled by $k = 1, \ldots, N$, where the k-th particle is in the one-body state φ_{j_k}. Although in general this is a physically forbidden state, since it violates undistiguishability, yet it has a precise mathematical meaning. Alternative form to (3.2) is the self-explanatory notation $|j_1, j_2, \ldots, j_N\rangle$, which is familiar in the physical literature. All this is to say that in (3.2) the particle *ordering* is kept into account. As an element of $L^2(\Omega^N)$, (3.2) is the wave function

$$(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})(x_1, x_2, \ldots, x_N) = \varphi_{j_1}(x_1)\varphi_{j_2}(x_2) \cdots \varphi_{j_N}(x_N) \tag{3.3}$$

where each variable $x_k \in \mathbb{R}^d$ runs in Ω and $\varphi_{j_k}(\,\cdot\,) \in L^2(\Omega)$. The generic element of $\mathcal{H}_N$ will be denoted by Ψ_N. Also, denote by $\mathfrak{S}_N$ the symmetric group of order N, i.e., the group of permutations of N symbols.

Our aim is to focus on the subspace of the total Hilbert space containing the only physically admissible states for a bosonic system.

Definition–Theorem 3.1.1 (symmetrization – I). Let $\sigma \in \mathfrak{S}_N$. The PERMUTATION OPERATOR relative to σ is the operator $P_\sigma : \mathcal{H}_N \longrightarrow \mathcal{H}_N$ such that

$$(P_\sigma \Psi_N)(x_1, x_2, \ldots, x_N) := \Psi_N(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(N)})\,, \tag{3.4}$$

so that $\sigma \mapsto P_\sigma$ is the representation of the permutation group $\mathfrak{S}_N$ on $\mathcal{H}_N$. On the basis wave functions, P_σ acts as

$$\begin{aligned} \big(P_\sigma(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})\big)(x_1, x_2, \ldots, x_N) = \\ = \varphi_{j_1}(x_{\sigma(1)})\varphi_{j_2}(x_{\sigma(2)}) \cdots \varphi_{j_N}(x_{\sigma(N)})\,, \end{aligned} \tag{3.5}$$

i.e., in terms of abstract basis vectors,

$$P_\sigma(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}) = \varphi_{j_{\sigma^{-1}(1)}} \otimes \varphi_{j_{\sigma^{-1}(2)}} \otimes \cdots \otimes \varphi_{j_{\sigma^{-1}(N)}}\,. \tag{3.6}$$

Notice that both in (3.5) and in (3.6) the following is meant: after the action of P_σ on the basis vector, the k-th particle, that was in the state φ_{j_k}, is now in the state $\varphi_{j_{\sigma^{-1}(k)}}$, while the state φ_{j_k} is now occupied by the $\sigma(k)$-th particle. It is easily seen that P_σ is unitary $\forall\, \sigma \in \mathfrak{S}_N$. Define also the SYMMETRIZER on $\mathcal{H}_N$ to be the operator

$$\begin{aligned} &\mathcal{S} : \mathcal{H}_N \longrightarrow \mathcal{H}_N \\ &\mathcal{S} := \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} P_\sigma\,. \end{aligned} \tag{3.7}$$

One sees that

$$\mathcal{S} = \mathcal{S}^* = \mathcal{S}^2 \tag{3.8}$$

so $\mathcal{S}$ is an orthogonal projection. The Hilbert subspace it projects onto is said the SYMMETRIC or BOSONIC SECTOR of $\mathcal{H}_N$, denoted by $\mathcal{H}_{N,\,\mathrm{sym}}$ or, in particular, by $L^2_{\mathrm{sym}}(\Omega^N)$:

$$\mathcal{H}_{N,\,\mathrm{sym}} := \mathcal{S}\,\mathcal{H}_N\,. \tag{3.9}$$

It is a postulate in this context that physically admissible *pure* states for a system of N undistinguishable bosons are invariant under $\mathcal{S}$, i.e., are norm-one vectors in $\mathcal{H}_{N,\,\mathrm{sym}}$, i.e., are eigenvectors of $\mathcal{S}$ with eigenvalue 1:

$$\mathcal{S}\,\Psi_N = \Psi_N \qquad \text{(admissible pure states for the bosonic system)}\,. \tag{3.10}$$

More generally, physically admissible (possibly *mixed*) states are density matrices operators on $\mathcal{H}_N$ which act non-trivially only on $\mathcal{H}_{N,\,\mathrm{sym}}$:

$$\begin{aligned} \mathcal{H}_N &= \mathcal{S}\,\mathcal{H}_N \oplus (\mathbb{1}-\mathcal{S})\,\mathcal{H}_N \\ &\equiv \mathcal{H}_{N,\,\mathrm{sym}} \oplus \mathcal{H}_{N,\,\mathrm{asym}} \\ \gamma_N(\mathcal{H}_{N,\,\mathrm{sym}}) &\subset \mathcal{H}_{N,\,\mathrm{sym}} \\ \gamma_N\big|_{\mathcal{H}_{N,\,\mathrm{asym}}} &= \mathbb{O}\,. \end{aligned} \tag{3.11}$$

When $\mathcal{H}_N = L^2(\Omega^N)$, the bosonic sector is nothing but the subspace of permutationally symmetric wave functions, i.e., wave functions such that

$$\begin{gathered} \Psi_N(x_1, x_2, \dots, x_N) = \Psi_N(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)}) \\ \forall\,\sigma \in \mathfrak{S}_N\,, \qquad \forall\,\Psi_N \in L^2_{\mathrm{sym}}(\Omega^N)\,, \end{gathered} \tag{3.12}$$

and bosonic states are density matrices such that their kernels are symmetric under permutation of each set of variables,

$$\begin{gathered} \gamma_N(x_1, \dots, x_N; y_1, \dots, x_N) = \gamma_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}; y_{\sigma(1)}, \dots, y_{\sigma(N)}) \\ \forall\,\sigma \in \mathfrak{S}_N\,, \qquad \forall \text{ bosonic state } \gamma_N\,. \end{gathered} \tag{3.13}$$

As a consequence of the aforementioned postulate and of first principles of Quantum Mechanics, physical observables do not connect the symmetric (bosonic) and antisymmetric (fermionic) sectors (for, otherwise one could provide a physical measure – a projection – that would prepare a superposition of a bosonic + fermionic state, starting from a bosonic state only). Thus, observables one shall be dealing with, i.e., bosonic observables, are operators $\mathcal{O}_N \in \mathcal{L}(\mathcal{H}_N)$ which act non-trivially only on $\mathcal{H}_{N,\,\mathrm{sym}}$ as γ_N in (3.10). In particular, when $\mathcal{O}_N \in \mathcal{L}^2(L^2(\Omega^N))$, namely, when $\mathcal{O}_N$ has an integral kernel, then it satisfies (3.13) as well. The Hamiltonian too of a system of N identical bosons shares the same symmetry

$$[\,\mathcal{S}, e^{-itH_N}] = \mathbb{O} \qquad \forall\, t \in \mathbb{R}. \tag{3.14}$$

Hence, any bosonic initial state evolves inside the bosonic sector at any later time.

The basis $\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \dots, j_N}$ of $\mathcal{H}_N$ turns out not to be convenient for describing the physical subspace $\mathcal{H}_{N,\,\mathrm{sym}}$. Among them, only basis vectors that have the form of a N-th tensor power are physical. We will exhibit soon a convenient basis of $\mathcal{H}_{N,\,\mathrm{sym}}$ in equation (3.22). To this aim, first notice that $\mathcal{H}_{N,\,\mathrm{sym}}$ is certainly generated by $\{\mathcal{S}(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})\}_{j_1, j_2, \dots, j_N}$, so one needs to manage vectors of this kind.

First we adopt the following self-explanatory and rather useful notation.

Definition 3.1.2 (**quantum labels for N-th tensor product vectors**)**.** A basis vector $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ of $\mathcal{H}_N$ is said to have QUANTUM LABELS

$$|i_1, n_1; i_2, n_2; \dots; i_r, n_r\rangle \tag{3.15}$$

(also written $|\varphi_1, n_1; \varphi_2, n_2; \dots; \varphi_r, n_r\rangle$), for some $r = 1, \dots, N$ and some positive integers $n_1, \dots, n_r$ such that $\sum_{i=1}^r n_i = N$ and $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_r \geqslant 1$, iff the following holds:

① there are r distinct indices $i_1, \dots, i_r$ in $\{j_1, j_2, \dots, j_N\}$ (namely, there are r distinct vectors $\varphi_{i_1}, \dots, \varphi_{i_r}$ among $\varphi_{j_1}, \varphi_{j_2}, \dots, \varphi_{j_N}$),

② there are, in $\{j_1, j_2, \dots, j_N\}$, n_1 indices equal to i_1, n_2 indices equal to i_2, … , and n_r indices equal to i_r.

Equivalently, there is a suitable permutation operator P_σ such that

$$P_\sigma(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}) = \varphi_{i_1}^{\otimes n_1} \otimes \cdots \otimes \varphi_{i_r}^{\otimes n_r}$$

for r distinct $\varphi_{i_1}, \dots, \varphi_{i_r}$ out of $\varphi_{j_1}, \varphi_{j_2}, \dots, \varphi_{j_N}$.

The well known spirit of this definition is simply that quantum labels (3.15) shall be all that one is allowed to say on a physically admissible pure state: due to undistinguishability, it is not possible to say *which* particles of the N-body state labeled by $|i_1, n_1; \dots; i_r, n_r\rangle$ are, e.g., in the one-body state φ_{i_1}, but simply that there are n_1 out of N in that state φ_{i_1}. Thus, for instance, the three vectors in $\mathcal{H}_3$

$$\xi \otimes \xi \otimes \eta\,, \quad \xi \otimes \eta \otimes \xi\,, \quad \eta \otimes \xi \otimes \xi$$

(with $\|\xi\| = \|\eta\| = 1$ and $(\xi, \eta) = 0$ in $\mathcal{H}$), are mathematically distinct objects, and are all the possible basis vectors with quantum labels $|\xi, 2; \eta, 1\rangle$. Also, not even one is physically admissible, i.e., belongs to $\mathcal{H}_{3,\,\mathrm{sym}}$, while the only physically meaningful pure state one can build with them is (apart a normalisation)

$$\begin{aligned}\mathcal{S}(\xi \otimes \xi \otimes \eta) = \mathcal{S}(\xi \otimes \eta \otimes \xi) = \mathcal{S}(\eta \otimes \xi \otimes \xi) = \\ = \frac{1}{3}(\xi \otimes \xi \otimes \eta + \xi \otimes \eta \otimes \xi + \eta \otimes \xi \otimes \xi)\,.\end{aligned}$$

More generally one remarks the following.

Theorem 3.1.3. ① Fix an integer r between 1 and N, fix r positive integers $n_1, \dots, n_r$ with $\sum_{i=1}^r n_r = N$, and pick r distinct vectors $\varphi_{i_1}, \dots, \varphi_{i_r}$ in the basis $\{\varphi_j\}_j$ of $\mathcal{H}$. Then the number of distinct vectors in the basis $\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \dots, j_N}$ of $\mathcal{H}_N$ which have the same quantum labels $|\varphi_{i_1}, n_1; \dots; \varphi_{i_r}, n_r\rangle$ is

$$\frac{N!}{n_1!\, n_2!\, \cdots n_r!}$$

and, unless $r = 1$ (whence $n_1 = N$), none of them is in $\mathcal{H}_{N,\,\mathrm{sym}}$.

② All distinct basis vectors of $\mathcal{H}_N$ with the same quantum labels are orthonormal in $\mathcal{H}_N$ and have the same image in $\mathcal{H}_{N,\,\mathrm{sym}}$ under the action of $\mathcal{S}$.

③ Any two basis vectors of $\mathcal{H}_N$ with distinct quantum labels are orthonormal and give orthogonal images in $\mathcal{H}_{N,\,\mathrm{sym}}$ under the action of $\mathcal{S}$.

④ Partition the basis $\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \dots, j_N}$ of $\mathcal{H}_N$ into classes of vectors with the same quantum labels. Pick a vector in each class, take their image under $\mathcal{S}$ and renormalise them to 1. This way, one gets an orthonormal basis of $\mathcal{H}_{N,\,\mathrm{sym}}$.

⑤ Let $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ have quantum labels $|i_1, n_1; i_2, n_2; \dots; i_r, n_r\rangle$. Then $\mathcal{S}(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})$ is a linear combination of all the $N!/(n_1!\, n_2!\, \cdots n_r!)$ distinct basis vectors with the same quantum labels, each of these being weighted with $(n_1!\, n_2! \cdots n_r!)$.

⑥ Let $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ have quantum labels $|i_1, n_1; i_2, n_2; \ldots; i_r, n_r\rangle$. Then

$$\big\| \mathcal{S}(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}) \big\|^2_{\mathcal{H}_N} = \frac{n_1!\, n_2! \cdots n_r!}{N!}\,. \tag{3.16}$$

So it is natural to asopt the following definition (the well-posedness of which is guaranteed by theorem 3.1.3).

Definition 3.1.4 (**quantum labels in the bosonic sector**)**.** Each orthonormal basis of $\mathcal{H}_{N,\,\mathrm{sym}}$ built as in point ④ of theorem 3.1.3 is said a CANONICAL BASIS of $\mathcal{H}_{N,\,\mathrm{sym}}$ associated to the basis $\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \ldots, j_N}$ of $\mathcal{H}_N$. Each vector Ψ_N of such a canonical basis is said to have quantum labels $|\varphi_1, n_1; \ldots; \varphi_r, n_r\rangle$ iff

$$\Psi_N = \mathcal{S}\,(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})$$

where $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ is labeled by $|\varphi_1, n_1; \ldots; \varphi_r, n_r\rangle$, or equivalently, iff Ψ_N is the image, under $\mathcal{S}$, of any basis vector of $\mathcal{H}_N$ with those quantum labels, for instance

$$\Psi_N = \mathcal{S}\,(\varphi_{i_1}^{\otimes n_1} \otimes \cdots \varphi_{i_r}^{\otimes n_r})\,. \tag{3.17}$$

The standard (and physically ubiquitous) example one has in mind concerns the one-body Hilbert space $\mathcal{H} \cong \mathbb{C}^2$, where an orthonormal basis $\{e_1, e_2\}$ is chosen. Then

$$\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$$

is an orthonormal basis of $\mathcal{H}_2 = \mathcal{H}^{\otimes 2}$ and, via $\mathcal{S}$ (and after renormalisation), one builds the canonical basis

$$\{e_1 \otimes e_1, \frac{1}{\sqrt{2}}(e_1 \otimes e_2 + e_2 \otimes e_1), e_2 \otimes e_2\}$$

of $\mathcal{H}_{N,\,\mathrm{sym}}$, often referred to as *triplet* basis. The states of this triplet basis have quantum labels $|e_1, 2\rangle$, $|e_1, 1; e_2, 1\rangle$ and $|e_2, 2\rangle$, respectively. The vector $\frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1)$ of $(\mathbb{1} - \mathcal{S})\mathcal{H}_2$ completes the triplet basis to an orthonormal basis of the whole $\mathcal{H}_2$ and is usually referred to as *singlet* state.

Incidentally, renormalisation to 1 of each $\mathcal{S}\Psi_N$ (provided that $\Psi_N \notin \ker \mathcal{S}$) is a standard operation that deserves a specific notation.

Definition–Theorem 3.1.5 (**symmetrization – II**)**.** The SYMMETRIZED of any $\Psi_N \in \mathcal{H}_N$ such that $\Psi_N \notin \ker \mathcal{S}$ is the norm-one vector

$$\big(\Psi_N\big)_{\mathrm{sym}} := \frac{\mathcal{S}\Psi_N}{\|\mathcal{S}\Psi_N\|} \in \mathcal{H}_{N,\,\mathrm{sym}}\,. \tag{3.18}$$

Any vector of the basis $\{\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\}_{j_1, j_2, \ldots, j_N}$ of $\mathcal{H}_N$ has a nonzero image under $\mathcal{S}$. Let $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ have quantum labels $|i_1, n_1; i_2, n_2; \ldots; i_r, n_r\rangle$. Then, according to (3.16),

$$\big(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\big)_{\mathrm{sym}} = \sqrt{\frac{N!}{n_1!\, n_2! \cdots n_r!}}\; \mathcal{S}(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})\,. \tag{3.19}$$

Last equation can be simplified, by point ⑤ in theorem 3.1.3, as

$$\big(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\big)_{\mathrm{sym}} = \sqrt{\frac{n_1!\, n_2! \cdots n_r!}{N!}} \sum_{\sigma \in \mathfrak{S}'_N} P_\sigma(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N})\,. \tag{3.20}$$

where $\mathfrak{S}'_N \subset \mathfrak{S}_N$ consists of all the $N!/(n_1!\, n_2! \cdots n_r!)$ distinct permutations of the N-ple

$$\big(\underbrace{i_1, \dots i_1}_{(n_1)}, \underbrace{i_2, \dots i_2}_{(n_2)}, \dots, \underbrace{i_r, \dots i_r}_{(n_r)}\big)\,.$$

The symmetrized of any basis vector $\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$ of $\mathcal{H}_N$ with quantum labels $|i_1, n_1; i_2, n_2; \dots; i_r, n_r\rangle$ will be denoted also by

$$\begin{aligned} |i_1, n_1; i_2, n_2; \dots; i_r, n_r\rangle_{\mathrm{sym}} &:= \big(\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\big)_{\mathrm{sym}} \\ &= \big(\varphi_{i_1}^{\otimes n_1} \otimes \varphi_{i_2}^{\otimes n_2} \otimes \cdots \otimes \varphi_{i_r}^{\otimes n_r}\big)_{\mathrm{sym}}\,. \end{aligned} \tag{3.21}$$

For example, when $\|\xi\| = \|\eta\| = 1$ and $(\xi, \eta) = 0$,

$$\begin{aligned} \mathcal{S}(\xi \otimes \xi \otimes \eta) =& \frac{1}{6}(\xi \otimes \xi \otimes \eta + \xi \otimes \xi \otimes \eta + \xi \otimes \eta \otimes \xi \\ &+ \xi \otimes \eta \otimes \xi + \eta \otimes \xi \otimes \xi + \eta \otimes \xi \otimes \xi) \\ =& \frac{1}{3}(\xi \otimes \xi \otimes \eta + \xi \otimes \eta \otimes \xi + \eta \otimes \xi \otimes \xi)\,, \\ \big(\xi \otimes \xi \otimes \eta\big)_{\mathrm{sym}} =& \frac{1}{\sqrt{3}}(\xi \otimes \xi \otimes \eta + \xi \otimes \eta \otimes \xi + \eta \otimes \xi \otimes \xi)\,. \end{aligned}$$

Summarizing, one chooses an orthonormal basis $\{\varphi_j\}_j$ of the single-particle Hilbert space $\mathcal{H}$ – and this has to be done in the most convenient way for the Physics of the problem – which determines the orthonormal product basis (3.1) of the N-body space $\mathcal{H}_N$. Then, partitioning (3.1) into classes of basis vectors with the same quantum labels, and taking the $(\,\cdot\,)_{\mathrm{sym}}$ in each class, one gets a canonical orthonormal basis of the bosonic sector $\mathcal{H}_{N,\,\mathrm{sym}}$. Thus, varying

- ✓ $r = 1, \dots, N$,
- ✓ $\varphi_{i_1}, \dots, \varphi_{i_r}$ distinct in $\{\varphi_j\}_j$,
- ✓ $(n_1, \dots, n_r)$ such that $n_1 \geqslant \cdots \geqslant n_r \geqslant 1$ in $\mathbb{N}$ and $\sum_{i=1}^r n_i = N$,

one runs over all distinct quantum labels $|\varphi_{i_1}, i_1; \dots; \varphi_{i_r}, i_r\rangle$ which label such a basis of $\mathcal{H}_{N,\,\mathrm{sym}}$. Written in terms of this basis, a generic N-body bosonic pure state reads

$$\begin{aligned} \Psi_N &= \sum_{|\varphi_{i_1}, n_1; \dots; \varphi_{i_r}, n_r\rangle} \Big(\big(\varphi_{i_1}^{\otimes n_1} \otimes \cdots \otimes \varphi_{i_r}^{\otimes n_r}\big)_{\mathrm{sym}}\,,\, \Psi_N \Big)_{\mathcal{H}_N} \big(\varphi_{i_1}^{\otimes n_1} \otimes \cdots \otimes \varphi_{i_r}^{\otimes n_r}\big)_{\mathrm{sym}} \\ &= \sum_{\substack{r \\ i_1, \dots i_r \\ n_1, \dots, n_r}} |\varphi_{i_1}, n_1; \dots, \varphi_{i_r}, n_r\rangle_{\mathrm{sym}\ \mathrm{sym}}\langle \varphi_{i_1}, n_1; \dots, \varphi_{i_r}, n_r | \Psi_N\rangle \\ &= \sum_{\substack{\{\text{quantum} \\ \text{labels}\}}} C_{\substack{\text{quantum} \\ \text{labels}}}\, |\varphi_{i_1}, n_1; \dots; \varphi_{i_r}, n_r\rangle \end{aligned} \tag{3.22}$$

(the last line being just a suggestive condensed form).

Combining this knowledge of a basis of $\mathcal{H}_{N,\,\mathrm{sym}}$ with prescription (3.11) for bosonic density matrices, it follows that the generic (possibly mixed) N-body bosonic state is a density matrix

$$\gamma_N = \sum_j \lambda_j |\Phi_N^{(j)}\rangle\langle\Phi_N^{(j)}|\,, \qquad \lambda_j > 0\,, \qquad \sum_j \lambda_j = 1 \tag{3.23}$$

for some orthonormal basis $\{\Phi_N^{(j)}\}_j$ of $\mathcal{H}_{N,\mathrm{sym}}$, which in general is *not* the canonical basis entering (3.22). Expression (3.23) is nothing but the concrete realization of the diagonal form (2.21) in this picture.

3.2 Limit of infinite systems

It is an unfortunate but true fact that essentially always many-body problems are too hard to be solved explicitly, due to their overwhelming complexity. A standard perspective is to study the problem in some kind of infinite limit for the "large" system. To give logical consistency to the present chapter, this section provides details of such an approach in view of the following two sections dealing with gases of bosons in a possibly condensed phase. This anticipates a more complete discussion on the main concept of *scaling limit* which is the topic of the next chapter.

For concreteness, let $\mathcal{S}_0$ be a system of N_0 undistinguishable bosons, possibly interacting and confined in some region of the d-dimensional space. Associated to it, one has a Hilbert space $\mathcal{H}_{N_0} = \mathcal{H}^{\otimes N_0}$, $\mathcal{H}$ being the Hilbert space by which a single particle is described, and a Hamiltonian H_{N_0} which in general is not the sum of single-particle Hamiltonians, due to the presence of some interaction.

The customary mathematical idealization is to model the "true" physical $\mathcal{S}_0$ by an increasing sequence $\{\mathcal{S}_N\}_N$ of larger and larger N-body systems that share with $\mathcal{S}_0$ some prescribed features, and to take the limit $N \to \infty$ (in fact, in current experiments with Bose-Einstein condensates, N ranges from some hundreds to 10^{11} – see Table 1.1). More precisely, *it has to be emphasized that each $\mathcal{S}_N$ is* not *simply the same $\mathcal{S}_0$, but with a larger N instead of N_0, and that unless some prescriptions on the Hamiltonian H_N of $\mathcal{S}_N$ are taken, the limit procedure is ambiguous.*

What one assumes is that each $\mathcal{S}_N$ is made by N particles of the same kind as in S_{N_0}, confined and coupled each other as particles are in S_0,[1] but with coupling strength and size of the confinement depending on N in such a way that

① as $N = N_0$, the system $\mathcal{S}_N$ is exactly $\mathcal{S}_0$, i.e., $H_N|_{N=N_0} = H_{N_0}$ on $\mathcal{H}_N = \mathcal{H}_{N_0}$,

② for any N, or asymptotically as $N \to \infty$, *certain* physical quantities of $\mathcal{S}_N$ in the state γ_N are the same as the corresponding ones for $\mathcal{S}_0$ in the state γ_{N_0}.

For instance, a typical case (discussed in Sec. 4.3) is a gas of spinless undistinguishable bosons where, in the limit $N \to \infty$, the ratio between interaction and kinetic energy in the ground state is kept constant: such a ratio would not be constant with N if one simply took $\mathcal{S}_N$ to be exactly $\mathcal{S}_0$ with N instead of N_0, namely, without rescaling the potential.

This way, each $\mathcal{S}_N$ resembles by construction the original $\mathcal{S}_0$ at least for all those physical quantities that are kept fixed with N.

Suppose now that one is interested in some physical quantity f (i.e., the expectation of some observable) of the original system $\mathcal{S}_0$ in some state, e.g., its ground state. Normally f is not accessible exactly, yet being in principle a well-defined function $f(N, L, a, \dots)$ of the labels of the quantum state under consideration and of the relevant physical

[1] This, in principle, can change the shape of the trap: this is the case when the scaling rate is faster along one spatial dimension than the others, with the aim to obtain a lower dimensional model. We will be not discussing this possibility, so for us by scaling the trap mantains the same aspect ratio as for S_0.

parameters of the Hamiltonian, such as the number N of particles, the size L of the system (e.g., the typical linear size of the trap where the gas is confined in), the s-wave scattering length a of the interparticle interaction, etc. To obtain f, prescription ② above fixes in principle how $L, a, \dots$ have to *scale* with N (say $L_N, a_N, \dots$), as $N \to \infty$, thus making the limit for f unambiguous in the form $\lim_{N\to\infty} f(N, L_N, a_N, \dots)$.

The collection of all such prescriptions in H_N identifies the so called SCALING LIMIT under which, as $N \to \infty$, quantities of interest of the form $f(N, L, a, \dots)$ are investigated, to get explicitly their asymptotics. Thus, a given scaling is nothing but a curve going to infinity in the direction of monotonically increasing N in the domain $\{(N, L, a, \dots)\}$ of the function f (see Fig. 4.1), such that at each point the corresponding system $\mathcal{S}_N$ exhibits a number of features in common with the original $\mathcal{S}_0$. If in the scaling limit one gets

$$\lim_{N\to\infty} f(N, L_N, a_N, \dots) = \mathcal{F}\,,$$

then such $\mathcal{F}$ is regarded as an approximation of the value $f(N_0, L_0, a_0, \dots)$ of the true physical system under consideration, since the asymptotics has been obtained under the same physical regime of $\mathcal{S}_0$ (see Sec. 4.3 for details).

So the N_0-body physical system is investigated via a sequence $\{\mathcal{S}_N\}_N$ of larger and larger auxiliary systems each of which is described by a Hamiltonian H_N acting on $\mathcal{H}_N$. The original $\mathcal{S}_0$ in the state γ_{N_0} is regarded as a sequence of states $\{\gamma_N\}_N$.

Our standard setting will be

$$\mathcal{H}_N = L^2(\mathbb{R}^{Nd}) \cong L^2(\mathbb{R}^d)^{\otimes N} \tag{3.24}$$

where $d = 1, 2, 3$ is the dimension of the system. If $\Omega \subset \mathbb{R}^d$ is the region where the original system is confined in, then each N-body system is confined in a region Ω_N possibly depending on N,[2] that is, the size of the domain of integration may scale with N.

Each of the γ_N is well-defined, although their limit as $N \to \infty$ is ill-posed. In fact, the sequence of increasing systems $\mathcal{S}_N$ cannot be thought to "converge" to the original $\mathcal{S}_0$. More precisely, H_N acts on a larger and larger Hilbert space $\mathcal{H}_N$ which eventually becomes a far too large space to control the $N = \infty$ limit system (in the infinite tensor product $\mathcal{H}^{\otimes\infty}$ separability is lost and distinct tensor power vectors $\varphi^{\otimes\infty}$ are always orthogonal).

Nevertheless, observables of interest, when the "large" system is modelled as infinite, usually are observables only on finite parts of it and *it is at this very point that the tool of the reduced density matrices finds its application.*

Observables on a fixed subsystem of k particles out of N_0 will be lifted to the corresponding observables on the same fixed subsystem of each $\mathcal{S}_N$. Boson symmetry makes all choices of k particles equivalent, but from the mathematical viewpoint we will distinguish the variables ordering, so that the subsystem of k particles will be meant to be the subsystem of the *first* k.

More precisely, for some fixed integer k with $1 \leqslant k < N$, one regards $\mathcal{H}_N$ as

$$\mathcal{H}_N \cong \mathcal{H}_k \otimes \mathcal{H}_{N-k} \cong L^2(\mathbb{R}^{kd}) \otimes L^2(\mathbb{R}^{(N-k)d}) \tag{3.25}$$

[2]Notice the distinct notation: each $\Omega_N \subset \mathbb{R}^d$ is an element of the sequence $\{\Omega_N\}_N$ of enlarging domains in $\mathbb{R}^d$, while by Ω^N one means the Cartesian product $\Omega \times \Omega \times \cdots \times \Omega \subset \mathbb{R}^{Nd}$.

and one considers observables of the form $\mathcal{O} \otimes \mathbb{1}_{N-k}$ with $\mathcal{O} = \mathcal{O}^* \in \mathcal{L}(\mathcal{H}_k)$. The expectation of $\mathcal{O} \otimes \mathbb{1}_{N_0-k}$ in the state γ_{N_0} is thus investigated through the sequence of expectations

$$\mathrm{Tr}[\,\gamma_N(\mathcal{O} \otimes \mathbb{1}_{N-k})\,] = \mathrm{Tr}[\,\gamma_N^{(k)}\mathcal{O}\,] \tag{3.26}$$

where $\gamma_N^{(k)}$ is the reduced density matrix of the k-particle subsystem. Then one takes an appropriate scaling limit $N \to \infty$ as discussed above. Notice that dealing with observables on some fixed subsystem only allows one to overcome the ambiguity of the limit in $\{\gamma_N\}_N$: indeed the sequence $\{\gamma_N^{(k)}\}_N$ is all inside $\mathcal{L}^1(\mathcal{H}_k)$, hence taking $N \to \infty$ makes sense, provided that a suitable topology for the limit is specified.

Natural questions are

($\mathcal{Q}_1$) whether a limit exists for a reasonable, physically meaningful set of observables on the k-particle subspace $\mathcal{H}_k \otimes \mathbb{1}_{N-k}$ and whether it is of the form $\mathrm{Tr}[\,\gamma_\infty^{(k)}\mathcal{O}\,]$ for some density matrix $\gamma_\infty^{(k)}$ on $\mathcal{H}_k$,

($\mathcal{Q}_2$) whether this depends or not on k.

Usually ($\mathcal{Q}_1$) is addressed to $\forall\,\mathcal{O} \in \mathrm{Com}(\mathcal{H}_k)$: if in this case it has an affirmative answer, this is equivalent to say that $\gamma_N^{(k)} \to \gamma_\infty^{(k)}$ weakly-$*$ in $\mathcal{L}^1(\mathcal{H}_k)$, by point ⑤ in theorem 2.1.1:

$$\gamma_N^{(k)} \xrightarrow[N\to\infty]{\text{weakly-}* \text{ in } \mathcal{L}^1(\mathcal{H}_k)} \gamma_\infty^{(k)} \qquad \Longleftrightarrow \qquad \begin{cases} \lim\limits_{N\to\infty} \mathrm{Tr}[\,(\gamma_N^{(k)} - \gamma_\infty^{(k)})\,C\,] = 0 \\ \\ \forall\, C \in \mathrm{Com}(\mathcal{H}_k)\,. \end{cases} \tag{3.27}$$

The quantum mechanical interpretation is that the state $\gamma_\infty^{(k)}$ is an approximation of the "true" $\gamma_{N_0}^{(k)}$: the larger N_0, the better the approximation.

Question ($\mathcal{Q}_2$) deserves a special care when Bose-Einstein condensation is investigated in the limit of infinite system and it will be discussed for our purposes in Sec. 7.1. In fact, as we are going to remark, control of marginals at distinct k's gives information on the interparticle correlations, which, in turn, is central in the study of several B.E.C. features.

In conclusion, the twofold essence of this scaling approach is

- ✓ to regard the original physical system under consideration, $\mathcal{S}_0$, to be so large that only observables on finite portions of it are of interest, so that one deals with reduced density matrices only,

- ✓ to perform a $N \to \infty$ limit through an appropriate scaling, that is, keeping some relevant physical quantities constant with N: $\mathcal{S}_0$ is mimed with an enlarging sequence $\{\mathcal{S}_N\}_N$, the possible limit $\gamma_N^{(k)} \to \gamma_\infty^{(k)}$ is investigated at fixed k; within the scaling in use, $\gamma_\infty^{(k)}$ is an approximation of the true $\gamma_{N_0}^{(k)}$.

3.3 Bose-Einstein condensation for finite size systems

According to the physical scenario depicted in Chap. 1, a physically meaningful and reasonably general mathematical formalisation of Bose Einstein condensation has to account for the *macroscopic occupation of the same single-particle state*, the characteristic feature of such a phenomenon.

The experimental evidence of that relies, e.g., on the typical peak in the observed particle space distribution (or momentum distribution) of a suitably cooled and trapped alkali vapour (see, e.g., Fig. 1.8). From its shape one argues (we are coming back to this point later, see the introductory considerations of Chap. 6) that with some approximation the system is in a factorised state, all factors being the same one-body wave function. For a non-interacting system it corresponds to a bound state of the trap for a single particle. For interacting systems it is a modification of it – for instance instead of being the ground state of the trap, it is the minimizer of a suitable functional. The density profile is thus proportional to the density of each one-particle factor, called the *condensate wave function*. Yet for interacting particles, only N-body states have physical meaning, and one has to give meaning to a notion of occupation of a single-particle state in a many-body state.

So, occupation number is the crucial point when formalising B.E.C. In this section we are going to discuss it with respect to *finite* size systems, while in the next one we will focus on the same problem in the context of the limit of *infinite* system. Here and in the following the key remark will be that *occupation number is a concept that can be made physically and mathematically meaningful at the level of one-particle reduced density matrices, instead of at the level of N-body states.*

In fact,

- ✓ in the limit of infinite system, the tool of fixed-k-particle reduced density matrices emerges as the natural one, while the N-body state looses its meaning as $N \to \infty$, as we have already discussed in the previous section.

- ✓ Even when dealing with B.E.C. for finite size systems, the many-body wave function is neither easily accessible, due to the complexity of the problem, nor easily observable. Experimentally one observes, e.g., a spatial density distribution $n(x)$, that is, the average number of particles in some neighborhood of given point. By undistinguishability

$$n(x) = N \int_{\Omega^{N-1}} \left|\Psi_N(x, x_2, \dots, x_N)\right|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N = N\, \gamma_N^{(1)}(x, x)$$

  (indeed $\int_\Omega n(x)\,\mathrm{d}x = N$), $\Omega \subset \mathbb{R}^d$ being the region where the system is confined in. Thus, one is actually dealing with the one-particle reduced density matrix $\gamma_N^{(1)}$ of the N-body state γ_N. This is another way to say that physical observables involved are essentially one-particle observables.

So one needs to relate the standard description of the many-body states to properties of the marginals, in order to manage a satisfactory formalisation of condensation and, in particular, of the concept of macroscopic occupation of the same single-particle state.

For pure states (3.21), i.e., vectors of the canonical basis of the many-body bosonic sector $\mathcal{H}_{N,\mathrm{sym}}$, the notion of occupation numbers can be *indifferently* set at the level of N-body states or 1-marginals. In the first case, the statement that in the state

$$|\varphi_{i_1}, n_1; \varphi_{i_2}, n_2; \dots; \varphi_{i_r}, n_r\rangle_{\mathrm{sym}}$$

there is a fraction n_{i_1}/N of particles in the same one-particle state φ_{i_1}, etc., is definitely unambiguous. The same information is present at the one-body level, due to the following theorem.

Theorem 3.3.1 (**occupation numbers in a N-body pure state, [83]**). Let $\gamma_N = |\Psi_N\rangle\langle\Psi_N|$ be the N-body pure state with definite quantum labels

$$\Psi_N = |\varphi_{i_1}, n_1; \varphi_{i_2}, n_2; \ldots; \varphi_{i_r}, n_r\rangle_{\mathrm{sym}}\,. \tag{3.28}$$

Then the corresponding one-particle reduced density matrix is the *finite rank* operator

$$\gamma_N^{(1)} = \sum_{k=1}^{r} \frac{n_k}{N}\, |\varphi_{i_k}\rangle\langle\varphi_{i_k}|\,. \tag{3.29}$$

In particular, if $|\varphi_{i_1}, n_1; \varphi_{i_2}, n_2; \ldots; \varphi_{i_r}, n_r\rangle_{\mathrm{sym}}$ and $|\varphi_{i_1}, m_1; \varphi_{i_2}, m_2; \ldots; \varphi_{i_r}, m_r\rangle_{\mathrm{sym}}$ have the same one-particle reduced density matrix then $n_1 = m_1, \ldots, n_r = m_r$.

Proof. Recall (see equation (3.20) and theorem 3.1.3) that $|\varphi_{i_1}, n_1; \varphi_{i_2}, n_2; \ldots; \varphi_{i_r}, n_r\rangle_{\mathrm{sym}}$ is a linear combination of $N!/(n_1!\, n_2! \cdots n_r!)$ orthonormal basis vectors of the form (3.2), each weighted with the factor $((n_1!\, n_2! \cdots n_r!)/N!)^{1/2}$: these are the vectors

$$\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}$$

of $\mathcal{H}_N$ for all the possible distinct N-ples of indices $(j_1, j_2, \ldots, j_N)$ obtained by permutation of the assigned N-ple

$$\big(\underbrace{i_1, \ldots i_1}_{(n_1)}, \underbrace{i_2, \ldots i_2}_{(n_2)}, \ldots, \underbrace{i_r, \ldots i_r}_{(n_r)}\big)\,.$$

When constructing $\gamma_N = |\Psi_N\rangle\langle\Psi_N|$ one then gets diagonal terms of the form

$$\frac{n_1!\, n_2! \cdots n_r!}{N!}\, |\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\rangle\langle\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}| \tag{3.30}$$

and off-diagonal terms of the form

$$\frac{n_1!\, n_2! \cdots n_r!}{N!}\, |\varphi_{j_1} \otimes \varphi_{j_2} \otimes \cdots \otimes \varphi_{j_N}\rangle\langle\varphi_{j'_1} \otimes \varphi_{j'_2} \otimes \cdots \otimes \varphi_{j'_N}| \tag{3.31}$$

with

$$\begin{aligned} &(j_1, j_2, \ldots, j_N) \neq (j'_1, j'_2, \ldots, j'_N) \qquad \text{(labels of off-diagonal terms)} \\ &\begin{matrix} j_1, j_2, \ldots, j_N, \\ j'_1, j'_2, \ldots, j'_N \end{matrix} \;\in\; \big\{ \underbrace{i_1, \ldots i_1}_{(n_1)}, \underbrace{i_2, \ldots i_2}_{(n_2)}, \ldots, \underbrace{i_r, \ldots i_r}_{(n_r)} \big\}\,. \end{aligned} \tag{3.32}$$

Denote by $\mathrm{Tr}_{[N-1]} : \mathcal{H}^{\otimes N} \to \mathcal{H}$ the partial trace over the last $N-1$ particles (i.e., variables). One then sees that only diagonal terms contribute when partial trace is taken, for necessarily $(j_2, \ldots, j_N) \neq (j'_2, \ldots, j'_N)$ for any off-diagonal term. Indeed if $j_1 = j'_1$ then $(j_2, \ldots, j_N) \neq (j'_2, \ldots, j'_N)$ in order (3.32) to hold, while if $j_1 \neq j'_1$ then the equality $(j_2, \ldots, j_N) = (j'_2, \ldots, j'_N)$ would contradict that the collections

$$\{j_1, j_2, \ldots, j_N\} \quad \text{and} \quad \{j'_1, j'_2, \ldots, j'_N\}$$

are made by the same N integers. Then, when evaluated, e.g., in terms of kernels, $\mathrm{Tr}_{[N-1]}$ in (3.31) gives products of integrals of the form $\int_\Omega \varphi_{j_k}(x_k)\overline{\varphi_{j'_k}(x_k)}\,\mathrm{d}x_k$, $k = 2, \ldots, N$ and there must be at least one k such that $j_k \neq j'_k$ whence, by orthonormality,

$$\int_\Omega \varphi_{j_k}(x_k)\overline{\varphi_{j'_k}(x_k)}\,\mathrm{d}x_k = 0\,. \tag{3.33}$$

Thus, only diagonal terms (3.30) contribute to partial trace and each of them gives

$$\begin{aligned}\mathrm{Tr}_{[N-1]}\Big[\,\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\otimes\varphi_{j_2}\otimes\cdots\otimes\varphi_{j_N}\rangle\langle\varphi_{j_1}\otimes\varphi_{j_2}\otimes\cdots\otimes\varphi_{j_N}|\,\Big]\\ =\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\rangle\langle\varphi_{j_1}|\prod_{k=2}^{N}\int_\Omega|\varphi_{j_k}(x_k)|^2\,\mathrm{d}x_k\\ =\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\rangle\langle\varphi_{j_1}|\,.\end{aligned}\tag{3.34}$$

Actually, $|\varphi_{j_1}\rangle\langle\varphi_{j_1}|$ can be one of the r distinct projections $|\varphi_{i_1}\rangle\langle\varphi_{i_1}|,\dots,|\varphi_{i_r}\rangle\langle\varphi_{i_r}|$ and it remains to count, in the linear combination (3.20) giving $|\varphi_{i_1},n_1;\varphi_{i_2},n_2;\dots;\varphi_{i_r},n_r\rangle_{\mathrm{sym}}$, the number ν_{i_1} of terms with φ_{i_1} in first position, the number ν_{i_2} of terms with φ_{i_2} in first position, etc., for then

$$\gamma_N^{(1)}=\sum_{k=1}^{r}\nu_{i_k}\,\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{i_k}\rangle\langle\varphi_{i_k}|\,.\tag{3.35}$$

To do this counting, consider again the $N!/(n_1!\,n_2!\cdots n_r!)$ distinct ways to place $n_1\geqslant 1$ copies of $\varphi_{i_1},\dots$, and $n_r\geqslant 1$ copies of φ_{i_r} in an ordered set of $N=n_1+\cdots+n_r$ boxes. Fixing, e.g., one copy of φ_1 in first position, one ends up with all distinct dispositions of $n_1-1\geqslant 0$ copies of $\varphi_{i_1},\dots$, and $n_r\geqslant 1$ copies of φ_{i_r} in an ordered set of $N-1$ boxes: their number is $(N-1)!/((n_1-1)!\,n_2!\cdots n_r!)$. Thus, in general,

$$\nu_{i_k}=\frac{(N-1)!}{n_1!\cdots(n_k-1)!\cdots n_r!}\,,\qquad k=1,\dots,r\,.\tag{3.36}$$

In conclusion, (3.35) and (3.36) give (3.29). □

On the other hand, for a generic N-body bosonic pure or mixed state – see (3.22) and (3.23) respectively – one should deal with some kind of weighted means of occupation numbers in terms of states with definite quantum labels. For instance, it is hard to say at first glance what "macroscopic occupation of single-particle states" might be in the (perfectly admissible) 8-body bosonic pure state

$$\Psi_8=\frac{1}{\sqrt{6}}\big(\,|\varphi,6;\xi,2\rangle_{\mathrm{sym}}+|\varphi,5;\xi,2;\eta,1\rangle_{\mathrm{sym}}+2\,|\varphi,4;\eta,4\rangle_{\mathrm{sym}}\,\big)$$

(with φ,ξ,η orthonormal in the one-body Hilbert space). One might say, e.g., that the one-body states φ, ξ and η have relative occupation in Ψ_8 given by

$$\frac{\frac{6}{8}+\frac{5}{8}+2\cdot\frac{4}{8}}{1+1+2}=\frac{19}{32}\,,\quad\frac{\frac{2}{8}+\frac{2}{8}}{1+1+2}=\frac{4}{32}\,,\quad\frac{\frac{1}{8}+2\cdot\frac{4}{8}}{1+1+2}=\frac{9}{32}$$

respectively. Set $\gamma_8:=|\Psi_8\rangle\langle\Psi_8|$: after a somewhat lengthy computation one finds

$$\gamma_8^{(1)}=\frac{27}{48}|\varphi\rangle\langle\varphi|+\frac{4}{48}|\xi\rangle\langle\xi|+\frac{17}{48}|\eta\rangle\langle\eta|+\frac{1}{8\sqrt{6}}|\varphi\rangle\langle\eta|+\frac{1}{8\sqrt{6}}|\eta\rangle\langle\varphi|$$

that is, a different (although similar) set of candidate occupation numbers. In both cases these numbers do not reconstruct Ψ_8 uniquely, of course, although in both cases one should agree, e.g., that the one-body state φ is "occupied" for a fraction which is slightly more than one half. Actually, when diagonalizing $\gamma_8^{(1)}$, one finds that the largest

eigenvalue is $\lambda_0 = \frac{22+\sqrt{31}}{48} \approx 0.57$ (the others being $\lambda_1 = \frac{22-\sqrt{31}}{48}$ and $\lambda_2 = \frac{4}{48}$) and it corresponds to the (normalised) eigenvector $(\frac{1}{2} + \frac{5}{2\sqrt{31}})^{1/2}\varphi + (\frac{1}{2} - \frac{5}{2\sqrt{31}})^{1/2}\eta$, that is, approximately, $0.97\varphi + 0.22\eta$.

When addressing the same question of occupancy in the one-particle reduced density matrix $\gamma_N^{(1)}$, instead of in the full γ_N, one can always have the decomposition (2.21)

$$\gamma_N^{(1)} = \sum_j \lambda_j |\varphi_j\rangle\langle\varphi_j| \,, \qquad \lambda_j > 0 \,, \qquad \sum_j \lambda_j = 1 \,,$$

where $\{\varphi_j\}_j$ is an orthonormal system of the one-particle Hilbert space $\mathcal{H}$. Each nonzero eigenvalue λ_j deserves the role of occupation number in the one-body state $\{\varphi_j\}_j$.

The above considerations serve as a motivation to the standard mathematical definition of B.E.C. for generic interacting N-body boson systems, which we are now discussing. In it, emphasis is given to the one-particle reduced density matrix. Such a definition has been given for the first time by Penrose and Onsager [95] half a century ago. In the current notation it reads as follows.

Definition 3.3.2 (**B.E.C. after Penrose-Osanger**)**.** Let γ_N be the state of a system of N identical and possibly interacting bosons and let the corresponding one-particle reduced density matrix $\gamma_N^{(1)}$ have diagonal form

$$\gamma_N^{(1)} = \sum_{j=0}^{M} \lambda_j |\varphi_j\rangle\langle\varphi_j| \,, \qquad \lambda_j \geqslant \lambda_{j+1} > 0 \,, \qquad \sum_{j=0}^{M} \lambda_j = 1 \,, \tag{3.37}$$

where $M \leqslant \infty$ and $\{\varphi_j\}_{j=0}^{M}$ is an orthonormal system of the one-particle Hilbert space $\mathcal{H}$. The N-body state γ_N is said to exhibit BOSE-EINSTEIN CONDENSATION iff the largest eigenvalue λ_0 is such that

$$\lambda_0 \sim 1 \tag{3.38}$$

with an approximation that has to be specified for each specific system under consideration. (The phrase "λ_0 almost 1" is somewhat vague in a situation where there is no limit in the number of particles, but in practice this does not usually lead to difficulty.)

Definition 3.3.3. One has SIMPLE CONDENSATION if one and only one eigenvalue of $\gamma_N^{(1)}$ is of order 1, all the others being negligible, and FRAGMENTED CONDENSATION if more that one eigenvalue is of order 1, e.g., λ_0, $\lambda_1 \sim 1$.

Thus, B.E.C. is defined to be a property *of the state* under consideration: no reference is made to the Hamiltonian of the system, that is, assigned a bosonic system with its Hamiltonian, there may be or there may be not states of condensation.

A natural problem then arises on whether B.E.C. can be *proved* in some state, once H_N is given, for instance in the ground state. We are discussing this point (*existence* of condensation) in Sec. 3.5 for a typical dilute gas of trapped and pairwise interacting bosons.

Another issue is the persistence of condensation in an originally condensed state which evolves in time: we are coming back to this point (*time-stability* of condensation) in Chap. 5.

Since a rigorous treatment of these issues requires a limit of infinitely many particle, in the next section we will restate the definition of B.E.C. for infinite systems.

3.4 B.E.C. in the limit of infinite systems

So far we have considered finite systems of bosons and the operational definition 3.3.2 of B.E.C. Actually one would like to appeal to some kind of limit $N \to \infty$ at least for two main reasons: (*i*) in practice the number of particles is so large to justify the assumption of infinite size; (*ii*) to *prove* condensation for interacting systems requires to solve a formidably difficult many-body problem in order to access $\gamma_N^{(1)}$, while there might be hopes to get asymptotic results when N goes to infinity.

Concerning the latter, recall that for interacting systems the ground state already poses a challenging problem which is still largely unsolved. B.E.C. itself has, so far, never been proved for many-body Hamiltonians with genuine interactions, except for the special case of hard-core bosons on a lattice at half-filling (see, e.g., [4]).

Discussion in Sec. 3.2 shows that $N \to \infty$ is not just a thermodynamic limit. Instead a more careful treatment is needed: some scaling has to be specified. After this is done, operational definition 3.3.2 takes the following asymptotic form.

Definition 3.4.1 (**asymptotic B.E.C.**). Let $\{\mathcal{S}_N\}_N$ be a sequence of systems, each of N identical bosons, with Hamiltonian H_N on the Hilbert space $\mathcal{H}_N = \mathcal{H}^{\otimes N}$ (where $\mathcal{H}$ is the one-particle Hilbert space). Let each $\mathcal{S}_N$ be in the state γ_N. The sequence $\{\mathcal{S}_N\}_N$ is assumed to encode the prescriptions under which limit $N \to \infty$ is taken in a N_0-body system $\mathcal{S}_0$ in the state γ_{N_0}, according to the discussion of Sec. 3.2. Let

$$\gamma_N^{(1)} = \sum_{j=0}^{M_N} \lambda_j^{[N]} \, |\varphi_j^{[N]}\rangle\langle\varphi_j^{[N]}| \,, \qquad \lambda_j^{[N]} \geqslant \lambda_{j+1}^{[N]} > 0 \,, \qquad \sum_{j=0}^{M_N} \lambda_j^{[N]} = 1 \,, \tag{3.39}$$

be the canonical diagonal form of the one-particle reduced density matrix $\gamma_N^{(1)}$ relative to each γ_N, where $M_N \leqslant \infty$ and $\{\varphi_j^{[N]}\}_{j=0}^{M_N}$ is an orthonormal system of $\mathcal{H}$. Then the system $\mathcal{S}_0$ in the state γ_{N_0} is said to exhibit

- ULTRA-WEAK ASYMPTOTIC BOSE-EINSTEIN CONDENSATION iff

$$\exists \liminf_{N\to\infty} \lambda_0^{[N]} =: \lambda_0 \tag{3.40}$$

  and $1 - \lambda_0$ is conveniently small, with an agreement depending on the specific system $\mathcal{S}_0$ under consideration.

- WEAK ASYMPTOTIC BOSE-EINSTEIN CONDENSATION iff

$$\exists \lim_{N\to\infty} \lambda_0^{[N]} =: \lambda_0 \tag{3.41}$$

  and $1 - \lambda_0$ is conveniently small, with an agreement depending on the specific system $\mathcal{S}_0$ under consideration.

- STRONG ASYMPTOTIC SIMPLE BOSE-EINSTEIN CONDENSATION iff

$$\exists \lim_{N\to\infty} \gamma_N^{(1)} =: \gamma_\infty^{(1)} \tag{3.42}$$

  in some topology for density matrices, where $\gamma_\infty^{(1)}$ is still a density matrix with diagonal form

$$\gamma_\infty^{(1)} = \sum_j \lambda_j |\varphi_j\rangle\langle\varphi_j| \tag{3.43}$$

and the largest eigenvalues λ_0 is conveniently close to 1, with an agreement depending on the specific system $\mathcal{S}_0$ under consideration. In other terms

$$\gamma_\infty^{(1)} \sim \lambda_0 |\varphi_0\rangle\langle\varphi_0| \,. \tag{3.44}$$

In this case λ_0 is said the FRACTION of the condensate, while $1 - \lambda_0$ is said its DEPLETION, and the corresponding one-body eigenstate φ_0 is said the CONDENSATE WAVE FUNCTION.

- STRONG ASYMPTOTIC FRAGMENTED BOSE-EINSTEIN CONDENSATION iff (3.42) holds and in (3.43) the two largest eigenvalues λ_0, λ_1 (or more) are of order 1, all the others being negligible, with an agreement depending on the specific system $\mathcal{S}_0$ under consideration. In this case the condensate is said to be fragmented into the condensate wave functions φ_0 and φ_1.

- COMPLETE (OR 100%) ASYMPTOTIC BOSE-EINSTEIN CONDENSATION in the condensate wave function φ iff $\exists\, \varphi \in \mathcal{H}$, with $\|\varphi\| = 1$, such that

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi\rangle\langle\varphi| \tag{3.45}$$

in some topology for density matrices.

In this definition we have included also two "weak" notions of condensation that sometimes one can find in the literature, but we will be dealing only with the customary "strong" notion, i.e., controlling condensation in terms of reduced density matrices. Of course, strong ⇒ weak ⇒ ultra-weak.

Unlike the finite-size case, asymptotic B.E.C. in definition 3.4.1 is not only a property of the state, but also of the way how limit $N \to \infty$ is taken. This is not often mentioned and has to be recalled to prevent thinking that asymptotic results are exactly features of the original $\mathcal{S}_0$ itself, instead of the auxiliary sequence $\{\mathcal{S}_N\}_N$. This is a quite usual misunderstanding when physicists of the Condensed Matter community read results provided by the Mathematical Physics community. Instead, asymptotic results have to be interpreted together with a control of the error terms depending on N and these, in turn, reflect how good the choice of the scaling is to mime $\mathcal{S}_0$ with $\{\mathcal{S}_N\}_N$. In general, unfortunately, a control of the error terms is currently a major open problem and, concerning rigorous results, one has often to content oneself with asymptotic results only.

A typical example is the case of the so called Gross-Pitaevskiĭ scaling limit, which is systematically used to study highly diluted gases of very weakly interacting bosons. By definition (see Sec. 4.3), this is a depletionless limit, in the sense that it cannot account for depletion. Indeed in this scaling the quantity ρa^3 is vanishing with N, where ρ is a (suitably defined) mean density of the gas and a is the s-wave scattering length of the interparticle interaction. Since from Bogolubov's theory, according to (1.26), one knows that depletion goes as $(\rho a^3)^{1/2}$, one cannot find depletion in this limit. This does *not* mean that such a scaling cannot describe B.E.C. with depletion, but only that asymptotic results intrinsically do not account for it. Depletion is zero in the limit and (small and) positive before the limit.

We come now to the *topology* by which one controls the limit $\gamma_N^{(1)} \to \gamma_\infty^{(1)}$. We have already argumented that a natural choice is the weak-$*$ topology in the trace class

$\mathcal{L}^1(\mathcal{H})$: this is exactly the convergence of the expectation values of all compact one-body observables – see equation (3.27). This is actually the most frequently adopted topology in the literature.

One may choose to control convergence in the much finer trace norm topology, but one discovers that the two limits are equivalent in this special case, where a sequence of density matrices converges to a density matrix. Indeed the following holds.

Theorem 3.4.2 (**Weakly converging density matrices converge in trace norm**)**.** Let $\{\gamma_N\}_N$ be a sequence of density matrices[3] and let γ be a density matrix, all on some Hilbert space $\mathcal{H}$. Then $\gamma_N \rightharpoonup \gamma$ weakly-$*$ in $\mathcal{L}^1(\mathcal{H})$ iff $\|\gamma_N - \gamma\|_{\mathcal{L}^1} \to 0$.

Theorem 3.4.2 is an old fact. It is a consequence of a general mechanism of convergence in the ideals of compact operators, first proved by Simon in 1979 ([101], theorem 2.20). As far as the trace ideal is concerned, Simon's theorem implies trace norm convergence, as $N \to \infty$, whenever $A_N \rightharpoonup A$, $A_N^* \rightharpoonup A^*$, $|A_N| \rightharpoonup |A|$ weakly, and $\|A_N\|_{\mathcal{L}^1} \to \|A\|_{\mathcal{L}^1}$. This is clearly the case for our sequence $\{\gamma_N\}_N$, all density matrices being positive and with trace equal to one. An improved version of this theorem has been achieved two years after by Arazy [7] and Simon [102], after it had been conjectured in [101]: they showed that it suffices that $A_N \rightharpoonup A$ weakly and that $\|A_N\|_{\mathcal{L}^1} \to \|A\|_{\mathcal{L}^1}$ to conclude that $\|A_N - A\|_{\mathcal{L}^1} \to 0$. In this spirit, we choose to refer to the original weaker Simon's theorem and we give detail of such a convergence mechanism in our case with density matrices.

Proof of theorem 3.4.2. Convergence in trace norm implies trace class weak-$*$ convergence. Indeed, by (2.6),

$$\mathrm{Tr}[(\gamma_N - \gamma)C] \leqslant \|C\| \, \|\gamma_N - \gamma\|_{\mathcal{L}^1} \xrightarrow{N\to\infty} 0 \tag{3.46}$$

for every $C \in \mathrm{Com}(\mathcal{H})$, whence

$$\mathrm{Tr}|(\gamma_N - \gamma)C| \xrightarrow{N\to\infty} 0\,. \tag{3.47}$$

Let us now prove the converse. First, trace class weak-$*$ convergence implies weak operator convergence. Indeed, by assumption, (3.47) holds $\forall\, C \in \mathrm{Com}(\mathcal{H})$ and the special choice $C = |\psi\rangle\langle\psi|$, $\forall\, \psi \in \mathcal{H}$, gives

$$\big|\,(\psi, (\gamma_N - \gamma)\psi)\,\big| = \big|\,\mathrm{Tr}[\,(\gamma_N - \gamma)|\psi\rangle\langle\psi|\,]\,\big| \leqslant \mathrm{Tr}\big|\,(\gamma_N - \gamma)|\psi\rangle\langle\psi|\,\big| \xrightarrow{N\to\infty} 0\,, \tag{3.48}$$

whence, by polarization,

$$(\psi, (\gamma_N - \gamma)\varphi) \xrightarrow{N\to\infty} 0 \qquad \forall\, \psi, \varphi \in \mathcal{H}\,. \tag{3.49}$$

Next, choose any $\varepsilon > 0$. Since $\mathrm{Tr}[\,\gamma\,] = 1$, one can choose a finite-dimensional orthogonal projection P onto the span of a sufficiently large number of eigenvectors of γ such that

$$\|\,P\gamma P\,\|_{\mathcal{L}^1} \geqslant 1 - \varepsilon\,. \tag{3.50}$$

[3]Here we dropped the index k: all the γ_N's act on $\mathcal{H}$. For our purposes the statement is: if $\gamma_N^{(1)} \rightharpoonup |\varphi\rangle\langle\varphi|$ weakly-$*$ in the trace class, then $\gamma_N^{(1)} \to |\varphi\rangle\langle\varphi|$ in the trace norm.

Notice that $P\gamma P \geqslant \mathbb{O}$, whence $\| P\gamma P \|_{\mathcal{L}^1} = \mathrm{Tr}[\, P\gamma P\,]$. Also, P is a positive compact (in particular: finite rank) operator. By trace class weak-$*$ convergence, $\exists\, N_\varepsilon$ such that, as $N \geqslant N_\varepsilon$, $|\,\mathrm{Tr}[(\gamma_N - \gamma)P]\,| \leqslant \varepsilon$, so

$$\begin{aligned} \| P\gamma_N P \|_{\mathcal{L}^1} &= \mathrm{Tr}|P\gamma_N P| = \mathrm{Tr}[\gamma_N P] \\ &= \mathrm{Tr}[(\gamma_N - \gamma)P] + \mathrm{Tr}[\gamma P] \\ &\geqslant -\big|\,\mathrm{Tr}[(\gamma_N - \gamma)P]\,\big| + \| P\gamma P \|_{\mathcal{L}^1} \\ &\geqslant -\varepsilon + (1-\varepsilon) = 1 - 2\varepsilon\,. \end{aligned} \tag{3.51}$$

Define the orthogonal projection $Q := \mathbb{1} - P$. Then $Q\gamma_N Q \geqslant \mathbb{O}$ and, $\forall\, N \geqslant N_\varepsilon$,

$$\begin{aligned} \| Q\gamma_N Q \|_{\mathcal{L}^1} &= \mathrm{Tr}[\gamma_N Q] \\ &= \mathrm{Tr}[\gamma_N] - \mathrm{Tr}[\gamma_N P] \\ &= 1 - \| P\gamma_N P \|_{\mathcal{L}^1} \\ &\leqslant 1 - (1 - 2\varepsilon) = 2\varepsilon\,. \end{aligned} \tag{3.52}$$

One analogously has

$$\| Q\,\gamma\,Q \|_{\mathcal{L}^1} \leqslant \varepsilon\,. \tag{3.53}$$

Thus,

$$\begin{aligned} \| \gamma_N - \gamma \|_{\mathcal{L}^1} =& \| (P+Q)(\gamma_N - \gamma)(P+Q) \|_{\mathcal{L}^1} \\ \leqslant\ & \| P(\gamma_N - \gamma)P \|_{\mathcal{L}^1} + \| Q(\gamma_N - \gamma)Q \|_{\mathcal{L}^1} \\ &+ \| P\,\gamma\,Q \|_{\mathcal{L}^1} + \| Q\,\gamma\,P \|_{\mathcal{L}^1} + \| P\,\gamma_N\,Q \|_{\mathcal{L}^1} + \| Q\,\gamma_N\,P \|_{\mathcal{L}^1}\,. \end{aligned} \tag{3.54}$$

Due to weak operator convergence (3.49),

$$\lim_{N\to\infty} \| P(\gamma_N - \gamma)P \|_{\mathcal{L}^1} = \lim_{N\to\infty} \| Q(\gamma_N - \gamma)Q \|_{\mathcal{L}^1} = 0\,, \tag{3.55}$$

so it remains to control that last four terms in the r.h.s. of (3.54) are small too.

This is immediately done as follows. Take for instance $\| P\,\gamma_N\,Q \|_{\mathcal{L}^1}$:

$$\| P\,\gamma_N\,Q \|_{\mathcal{L}^1} \leqslant \|P\| \cdot \| \gamma_N\,Q \|_{\mathcal{L}^1} \leqslant \sqrt{2\varepsilon} \tag{3.56}$$

due to (2.6) and to the following Schwartz inequality

$$\begin{aligned} \| \gamma_N\,Q \|_{\mathcal{L}^1} &\leqslant \| \sqrt{\gamma_N} \|_{\mathcal{L}^2} \cdot \| \sqrt{\gamma_N}\,Q \|_{\mathcal{L}^2} \\ &= \| \sqrt{\gamma_N} \|_{\mathcal{L}^2} \cdot \mathrm{Tr}[\, Q\gamma_N Q\,]^{1/2} \\ &\leqslant \sqrt{2\varepsilon} \end{aligned} \tag{3.57}$$

(where again it is crucial that density matrices are positive trace-one operators).

All other remaining terms in (3.54) are bounded analogously, so that

$$\| \gamma_N - \gamma \|_{\mathcal{L}^1} \leqslant \| P(\gamma_N - \gamma)P \|_{\mathcal{L}^1} + \| Q(\gamma_N - \gamma)Q \|_{\mathcal{L}^1} + \mathrm{const} \cdot \sqrt{\varepsilon}\,. \tag{3.58}$$

By (3.55) and since ε is arbitrary, then

$$\lim_{N\to\infty} \| \gamma_N - \gamma \|_{\mathcal{L}^1} = 0 \tag{3.59}$$

that is, convergence in trace norm. □

Since Hilbert-Schmidt convergence implies weak operator convergence, one has also the following.

Theorem 3.4.3 (**Hilbert-Schmidt or trace class norm convergence for density matrices is the same**)**.** Let $\{\gamma_N\}_N$ be a sequence of density matrices and let γ be a density matrix, all on some Hilbert space $\mathcal{H}$. Then $\|\gamma_N - \gamma\|_{\mathcal{L}^2} \to 0$ iff $\|\gamma_N - \gamma\|_{\mathcal{L}^1} \to 0$.

Here is a summary of equivalent ways of controlling factorisation of marginals (that is, $\gamma_N^{(1)} \to |\varphi\rangle\langle\varphi|$) for 100% asymptotic B.E.C.

Theorem 3.4.4 (**factorisation of marginals for 100% asymptotic B.E.C.**)**.** Let $\{\gamma_N^{(1)}\}_N$ be the sequence of one-particle reduced density matrices corresponding to the sequence $\{\gamma_N\}_N$ of N-body bosonic states, each on the Hilbert space $\mathcal{H}_N = \mathcal{H}^{\otimes N}$, where $\mathcal{H}$ is some one-particle Hilbert space. Let $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$. Then the following are equivalent:

i) $\gamma_N^{(1)} \xrightarrow{N\to\infty} |\varphi\rangle\langle\varphi|$ weakly-$*$ in trace class

ii) $\gamma_N^{(1)} \xrightarrow{N\to\infty} |\varphi\rangle\langle\varphi|$ in the Hilbert-Schmidt norm

iii) $\gamma_N^{(1)} \xrightarrow{N\to\infty} |\varphi\rangle\langle\varphi|$ in the trace class norm

iv) $(\varphi, \gamma_N^{(1)}\varphi) \xrightarrow{N\to\infty} 1$.

The same clearly can be restated also with respect to convergence $\gamma_N^{(k)} \xrightarrow{N\to\infty} |\varphi\rangle\langle\varphi|^{\otimes k}$, where k is a generic positive integer.

Proof. $(i) \Rightarrow (iv)$ because tracing the difference $\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|$ against the particular compact $|\varphi\rangle\langle\varphi|$ gives

$$\mathrm{Tr}[\,(\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|)|\varphi\rangle\langle\varphi|\,] = (\varphi, \gamma_N^{(1)}\varphi) - 1\,. \tag{3.60}$$

$(iv) \Rightarrow (ii)$ because

$$\begin{aligned} \big\|\,\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|\,\big\|^2_{\mathcal{L}^2(\mathcal{H})} &= \mathrm{Tr}[\,(\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|)^2\,] \\ &= \mathrm{Tr}[\,(\gamma_N^{(1)})^2\,] - 2\,\mathrm{Tr}[\,\gamma_N^{(1)}\,|\varphi\rangle\langle\varphi|\,] + \mathrm{Tr}[\,(|\varphi\rangle\langle\varphi|)^2\,] \\ &\leqslant 2\,\big(1 - (\varphi, \gamma_N^{(1)}\varphi)\big)\,. \end{aligned} \tag{3.61}$$

$(ii) \Leftrightarrow (iii)$ by theorem 3.4.3. $(i) \Leftrightarrow (iii)$ by theorem 3.4.2. Actually, in the present case, the limit point γ is the one-dimensional projection $|\varphi\rangle\langle\varphi|$, so in the proof of theorems 3.4.2 and 3.4.3 one can choose $P = |\varphi\rangle\langle\varphi| = \gamma$, thus improving the inequalities to $\|\,P\gamma P\,\|_{\mathcal{L}^1} = 1$, $\|\,P\gamma_N P\,\|_{\mathcal{L}^1} \geqslant 1 - \varepsilon$, $\|\,Q\gamma_N Q\,\|_{\mathcal{L}^1} \leqslant \varepsilon$, and $\|\,Q\gamma Q\,\|_{\mathcal{L}^1} = 0$. □

Other topologies than the trace class weak-$*$ one are present in the literature. This is typical when one is analysing of the *dynamics* of the N-body system. In this case one studies objects like $\gamma_{N,t}^{(1)}$, that is, *time-dependent* marginals corresponding to the state $\gamma_{N,t}$ of the system which is evolving in time and one wants to know whether in some approximation $\gamma_{N,t}^{(1)} \approx |\varphi_t\rangle\langle\varphi_t|$ for N large enough: this way, one gets information about the *evolution of the condensate wave function* φ_t.

So here the objects of interest are density matrices with kernels that are functions both of time and of space, i.e., $(t, x, y) \mapsto \gamma_{N,t}^{(1)}(x, y)$. To manipulate them in the limit $N \to \infty$ it is technically convenient to regard them as elements in suitable Sobolev spaces and this broadens the landscape of the possible topologies used to control the limit. This scenario will be presented in Sec. 5.4.

3.5 Proof of B.E.C. in a scaling limit: two examples. Stationary Gross-Pitaevskiĭ theory.

In this section we report two recent proofs of B.E.C. in the sense of definition 3.4.1, that is, asymptotic B.E.C. in a scaling limit for an infinite system of interacting bosons. We cite Ref. [75] as a comprehensive review. The first one is the case of a gas in a box with periodic or Neumann boundary conditions.

Theorem 3.5.1 (B.E.C. for dilute homogeneous gases – Lieb and Seiringer (2002), [74]). Let N_0 undistinguishable spinless bosons of mass m in a three-dimensional box Ω of side L with either periodic or Neumann boundary conditions be paired by a two-body, non-negative, spherically symmetric potential V with finite s-wave scattering length a, according to the Hamiltonian

$$H_{N_0} = -\frac{\hbar^2}{2m}\sum_{i=1}^{N_0}\Delta_{x_i} + \sum_{1\leqslant i<j\leqslant N_0} V(x_i - x_j) \tag{3.62}$$

(here each $x_i \in \Omega \subset \mathbb{R}^3$). Take the *Gross-Pitaevskiĭ scaling* [4] in the Hamiltonian, that is, $N_0 \mapsto N$, $L \mapsto L_N$, and $a \mapsto a_N$, such that

$$\frac{N a_N}{L_N} = \frac{N_0\, a}{L} = \text{const}\,. \tag{3.63}$$

This means that one considers systems with Hamiltonians

$$H_N = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_{x_i} + \sum_{1\leqslant i<j\leqslant N} V_N(x_i - x_j) \tag{3.64}$$

acting on $\mathcal{H}_N = L^2((\Omega_N)^N)$, with the same boundary conditions, where Ω_N is the cube of side L_N, and

$$V_N(x) := \Big(\frac{a}{a_N}\Big)^2 V\Big(\frac{a}{a_N}x\Big) \tag{3.65}$$

so that the scaled interaction V_N has scattering length a_N. Let Ψ_N be the ground state of H_N, and let $\gamma_N^{(1)}$ be the corresponding one-particle reduced density matrix. Then

$$\lim_{N\to\infty}\frac{1}{L_N^3}\int\limits_{\Omega_N\times\Omega_N}\gamma_N^{(1)}(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1\,. \tag{3.66}$$

Let us add some explanatory comments to theorem 3.5.1.

- ✓ First, the standard rigorous definition of scattering length can be found in appendix A of [80] and it implies immediately that V_N has scattering length a_N.
- ✓ Scaling (3.63) is a limit of infinite dilution and with a constant ratio between kinetic and interaction energy. Among all infinite equivalent realizations of (3.63), one can fix it as a "thermodynamic limit" $\rho_N := N/L_N^3 = N_0/L^3 = \text{const}$, whence necessarily $a_N \sim N^{-2/3}$ (due to the identity $a_N = (Na_N/L_N)(N/L_N^3)^{-1/3}N^{-2/3}$). Actually it is more manageable to fix $L_N = L$ and this is what we will refer to.

[4] a detailed presentation of which is in Sec. 4.3

So the scaling takes the form $Na_N = N_0\, a = \text{const}$. Thus, we will have in mind to fill the fixed box Ω with more and more particles, yet getting a more and more dilute system, because $\rho_N a_N^3 \to 0$, that is, the mean interparticle distance $\rho_N^{-1/3}$ is infinitely larger than a_N. (All this is discussed in detail in Sec. 4.3.)

✓ Limit (3.66) is equivalent to

$$\lim_{N\to\infty} (\varphi_0, \gamma_N^{(1)} \varphi_0) = 1\,, \qquad \varphi_0(x) \equiv \frac{1}{L^{3/2}}\,, \tag{3.67}$$

which in turn is equivalent to

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi_0\rangle\langle\varphi_0| \tag{3.68}$$

in the trace norm (see theorem 3.4.4). So one has 100% asymptotic condensation in the many-body ground state. Such φ_0, i.e., the condensate wave function, is recognised to be the ground state wave function of the free particle in the box Ω.

✓ By (3.68), $\gamma_N^{(1)}(x,y) \to L^{-3}$ at least in $L^2(\Omega\times\Omega)$-sense. So the kernel of the one-body reduced density matrix is approximately constant. In particular, it tends to be a function of the difference $x - y$ only, that is, to be translational invariant. This feature justifies the name one customarily refers to (3.66), that is, the OFF-DIAGONAL LONG RANGE ORDER in the kernel of $\gamma_N^{(1)}$. Such a notion of B.E.C. too has been introduced first by Penrose and Osanger [95].

An analogous result holds for inhomogeneous systems, e.g., a dilute gas of bosons confined by some trapping potential. To state it properly, one has to define a suitable notion of *mean density* of particles and of the *size* of the trap. Also, the condensate wave function cannot be expected to be of the form (3.67), due to spatial non uniformity. We enclose all these preliminaries – actually the rigorous basis of the Gross-Pitaevskiĭ theory – in the following statement.

Definition–Theorem 3.5.2 (GP theory – Lieb, Seiringer, and Yngvason (1999), [77]). Let N_0 undistinguishable spinless bosons of mass m be confined in a three-dimensional trapping potential $U \in L^\infty_{\text{loc}}(\mathbb{R}^3)$ such that $U(x) \to +\infty$ as $|x| \to +\infty$. Define the characteristic length L of the trap as

$$L := \sqrt{\frac{2\mu}{\hbar\omega}} \tag{3.69}$$

where

$$\mu := \frac{\hbar^2}{2m} \tag{3.70}$$

and $\hbar\omega$ is the ground state energy of $-\mu\Delta + U$. Also, let the particles be paired by a two-body, non-negative, spherically symmetric potential V with finite s-wave scattering length a, according to the Hamiltonian

$$H_{N_0} = \sum_{i=1}^{N_0} \Big(-\frac{\hbar^2}{2m}\Delta_{x_i} + U(x_i)\Big) + \sum_{1\leqslant i<j\leqslant N_0} V(x_i - x_j) \tag{3.71}$$

acting on $L^2(\mathbb{R}^{3N_0})$ – each $x_i \in \mathbb{R}^3$.

① Define the GROSS-PITAEVSKIĬ ENERGY FUNCTIONAL

$$\mathcal{E}^{\mathrm{GP}}[\varphi] := \int_{\mathbb{R}^3} \big(\mu|\nabla\varphi|^2 + U|\varphi|^2 + 4\pi\mu a|\varphi|^4\big)\,\mathrm{d}x \tag{3.72}$$

on the natural domain $\mathcal{D}$ of the normalised one-body wave functions on which it makes sense,[5] i.e.,

$$\mathcal{D} := \big\{\varphi \in H^1(\mathbb{R}^3) \,:\, U|\varphi|^2 \in L^1(\mathbb{R}^3), \|\varphi\|_{L^2} = 1\big\}\,. \tag{3.73}$$

It admits a unique (up to a phase) strictly positive minimizer φ^{GP}, which satisfies the TIME-INDEPENDENT GROSS-PITAEVSKIĬ EQUATION

$$-\mu\Delta\varphi^{\mathrm{GP}}(x) + U(x)\varphi^{\mathrm{GP}}(x) + 8\pi\mu a\varphi^{\mathrm{GP}}(x)^3 = \mu^{\mathrm{GP}}\varphi^{\mathrm{GP}}(x)\,, \tag{3.74}$$

the GP chemical potential μ^{GP} (i.e., the Lagrange multiplier) being fixed by normalisation. Denote the minimum of the functional as the GROSS-PITAEVSKIĬ ENERGY

$$E^{\mathrm{GP}} := \mathcal{E}^{\mathrm{GP}}[\varphi^{\mathrm{GP}}]\,. \tag{3.75}$$

Such φ^{GP} turns out to be also in $C^1(\mathbb{R}^3)\cap L^\infty(\mathbb{R}^3)$, with an exponential falloff of the kind $\varphi^{\mathrm{GP}}(x) \leqslant M_\alpha e^{-\alpha|x|}$ for any $\alpha > 0$ and an appropriate $M_\alpha > 0$; furthermore, $U \in C^\infty(\mathbb{R}^3) \Rightarrow \varphi^{\mathrm{GP}} \in C^\infty(\mathbb{R}^3)$, and if U is spherically symmetric and monotone increasing with $|x|$, then φ^{GP} is spherically symmetric and monotone decreasing.

② Take the Gross-Pitaevskiĭ scaling in the Hamiltonian with fixed L and $a \mapsto a_N$ such that

$$Na_N = N_0\, a = \mathrm{const}\,. \tag{3.76}$$

This means that one considers systems with Hamiltonians

$$H_N = \sum_{i=1}^{N}\Big(-\frac{\hbar^2}{2m}\Delta_{x_i} + U(x_i)\Big) + \sum_{1\leqslant i<j\leqslant N} V_N(x_i - x_j) \tag{3.77}$$

acting on $\mathcal{H}_N = L^2(\mathbb{R}^{3N})$, with

$$V_N(x) = \Big(\frac{N}{N_0}\Big)^2 V\Big(\frac{N}{N_0}x\Big) \tag{3.78}$$

so that the scaled interaction V_N has scattering length $a_N = aN_0/N$.

③ The gas is inhomogeneous. In view of describing its non uniform density as $N\varphi^{\mathrm{GP}}(x)^2$, define the GROSS-PITAEVSKIĬ DENSITY as its average

$$\bar\rho_N := \frac{1}{N}\int_{\mathbb{R}^3}\big(N(\varphi^{\mathrm{GP}}(x))^2\big)^2\,\mathrm{d}x = N\big\|\varphi^{\mathrm{GP}}\big\|_4^4\,. \tag{3.79}$$

Notice that $\bar\rho_N \sim N$ and $\bar\rho_N a_N^3 \sim N^{-2}$ as in the corresponding homogeneous case.

④ Let $\Psi_N^{\mathrm{g.s.}}$ be the ground state of H_N and let $E_N^{\mathrm{g.s.}}$ be the ground state energy. Then

$$\lim_{N\to\infty}\frac{E_N^{\mathrm{g.s.}}}{N} = E^{\mathrm{GP}}\,, \tag{3.80}$$

that is, the GP energy is the asymptotics of the many-body ground state energy per particle.

[5] in fact, $H^1(\mathbb{R}^3) \hookrightarrow L^4(\mathbb{R}^3)$

One can now formulate the second example of proof of asymptotic B.E.C. we announced: the case of complete condensation in the ground state of a trapped gas of bosons.

Theorem 3.5.3 (B.E.C. for dilute trapped gases – Lieb and Seiringer (2002), [74]). Assume the model and the scaling of theorem 3.5.2. Let $\gamma_N^{(1)}$ be the one-particle reduced density matrix corresponding to the ground state $\Psi_N^{\text{g.s.}}$ of the N-body system. Then

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi^{\text{GP}}\rangle\langle\varphi^{\text{GP}}| \tag{3.81}$$

in the trace norm.

Thus, also in this more realistic case, one has complete condensation in the ground state, and the condensate wave function is the minimizer of the GP functional (3.72), that is, a solution of the GP equation (3.74). Actually the GP energy functional in a box Ω with periodic or Neumann boundary conditions is minimized by the wave function φ_0 in (3.67), so that (3.81) reads as (3.68): in fact, we have split for convenience the unique work [74] into the two theorems 3.5.1 and 3.5.3.

By (3.81),

$$\gamma_N^{(1)}(x,y) \to \varphi^{\text{GP}}(x)\overline{\varphi^{\text{GP}}(y)} \tag{3.82}$$

at least in the $L^2(\mathbb{R}^3\times\mathbb{R}^3)$-sense. Consequently $\gamma_N^{(1)}(x,y)$ remains finite up to distances $|x-y|$ fixed by the extension of the function φ^{GP}. This is again that typical signature of B.E.C. which goes under the name of off-diagonal long range order for the kernel of the one-body marginal.

In standard textbooks treatment of B.E.C., off-diagonal long range order *in space* is formally equivalent, in Fourier transform, to the presence of a singular term in *momentum* distribution (see, e.g., Sec. 2.1 of [97]): the contribution of such a $\delta(p)$ term is thus interpreted as a macroscopic occupation of the single-particle state with momentum $p=0$. The analysis in [74] provides a rigorous statement in this sense, in the following form: under the assumptions of theorem 3.5.3,

$$\lim_{N\to\infty} \widehat{\rho}_N(p) = |\widehat{\varphi}^{\text{GP}}(p)|^2 \tag{3.83}$$

in L^1, where $\widehat{\varphi}^{\text{GP}}$ is the Fourier transform of φ^{GP} and

$$\widehat{\rho}_N(p) := \int_{\mathbb{R}^3\times\mathbb{R}^3} \gamma_N^{(1)}(x,y)\exp[\,ip(x-y)/\hbar]\,\mathrm{d}x\,\mathrm{d}y \tag{3.84}$$

denotes the one-body momentum density in the ground state.

Chapter 4

Role of scaling limits in the rigorous analysis of B.E.C.

4.1 Motivations

Asymptotic analysis of condensation in many-body dilute Bose gases in the limit $N \to \infty$ relies on the fact that (1) in practice the number of particles is so large ($N \sim 100 \div 10^{11}$) to justify the assumption of infinite size, and (2) condensation for interacting systems requires to solve a formidably difficult many-body problem in order to access $\gamma_N^{(1)}$, while there are techniques to get its asymptotics.

In Sec. 3.2 it has been pointed out that the nature of this limit is that of a *scaling limit*: the true system $\mathcal{S}_0$ in the state γ_{N_0} and with Hamiltonian H_{N_0} is regarded as an enlarging sequence $\{\mathcal{S}_N\}_N$ of N-body systems, each in the state γ_N and with Hamiltonian H_N, made by N particles of the same kind as in S_{N_0}, confined and coupled each other as particles are in S_0, but with coupling strength and size of the confinement depending on N in such a way that as $N = N_0$, $\mathcal{S}_N$ is exactly $\mathcal{S}_0$, i.e., $H_N|_{N=N_0} = H_{N_0}$ on $\mathcal{H}_N = \mathcal{H}_{N_0}$, and for any N, or asymptotically as $N \to \infty$, *certain* physical quantities of $\mathcal{S}_N$ in the state γ_N are the same as the corresponding ones for $\mathcal{S}_0$ in the state γ_{N_0}.

Thus, one fixes some quantities $\mathcal{Q}(N, L, a, \dots)$, i.e., the expectation values of some observables on $\mathcal{S}_N$ in the state γ_N (e.g. the ground state energy), to be constant with N: they depend on parameters labelling the quantum state under consideration and the Hamiltonian, such as the number N of particles, the size L of the system, the s-wave scattering length a of the interparticle interaction, etc. The prescription that these $\mathcal{Q}$'s are kept constant in the limit fixes in principle how $L, a, \dots$ have to *scale* with N (say $L_N, a_N, \dots$), thus making the limit for f unambiguous in the form $\lim_{N\to\infty} f(N, L_N, a_N, \dots)$.

In this chapter, following Ref. [84], the role of two particular scaling limits is analysed and commented, which turns out to be crucial in the rigorous study of B.E.C. First, we consider the so called Gross-Pitaevskiĭ (GP) limit. Its name itself has been customarily adopted due to one of the main results achievable in this limit, namely, the rigorous derivation of the Gross-Pitaevskiĭ equation, that for half a century has been known to be at the basis of the description of the condensate wave function in the diluted regime.

To state it properly, a preliminary discussion is needed on the interaction and kinetic average energies in a dilute Bose system, which is placed in Sec. 4.2. The definition of

the GP scaling and the mathematical and physical scenario behind it are then discussed in Sec. 4.3. Remarkable rigorous results achieved through this scaling have been already reported in Sec. 3.5. Such a discussion is meant to emphasize and clarify the *centrality* of the GP scaling in the study of B.E.C.

An alternative (and a modification) to the GP scaling, is the so called THOMAS-FERMI (TF) LIMIT: it is discussed in Sec. 4.4 in connection with its GP counterpart. The whole treatment is in three dimensions, as the most natural choice when thinking to the standard realizations of B.E.C. in the lab. The corresponding lower-dimensional scenario is then sketched in Sec. 4.6.

4.2 Ground state energy asymptotics for cold dilute Bose gases

The GP scaling has been introduced and eventually rigorously defined to be the "dynamical" limit of ultra-high dilution. That is, one wants that kinetic and interaction energies are still comparable as $N \to \infty$ – the limit is said *dynamical* because one can investigate the dynamics also in the limit – and the dilution becomes infinite. To make this precise, a control of the energy asymptotics is needed and it is not accidental that a systematic use of such a scaling has been made only after the 1997 rigorous solution of the ground state energy problem for cold dilute Bose gases – a task which covered the whole second half of last century.

To fix the setting, let H be the Hamiltonian for N spinless undistinguishable bosons of mass m enclosed in a three-dimensional cubic box Ω of side L, with suitable boundary conditions to be specified, interacting by a spherically symmetric two-body repulsive (i.e., non-negative) potential V, that is

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_{x_i} + \sum_{1\leqslant i<j\leqslant N} V(x_i - x_j) \tag{4.1}$$

($x_i \in \mathbb{R}^3$), acting on the bosonic sector (see (3.9)) of $L^2(\Omega^{\otimes N}, \mathrm{d}x_1 \cdots \mathrm{d}x_N)$, as is appropriate for the statistics in use. Also, recall by (3.70) the position $\mu := \frac{\hbar^2}{2m}$.

Attempts to derive rigorously the ground state energy per particle $E^{\mathrm{g.s.}}/N$ of (4.1) – absolute and bosonic ground state energies coincide, H being permutationally symmetric – in the physical regime of interest of high dilution and low temperature have been performed intensively between the late 1940s and the early 1960s.

Bogolubov's works [14, 15], perturbative derivations (based on the delta pseudopotentials introduced by Fermi [47], Breit [21], and Blatt and Weisskopf [13]) both in the simplified case of a Bose gas of hard spheres, due to Huang, Lee and Yang [58, 69, 70] and Wu [110], and in a more general setting due to Brueckner and Sawada [25], Beliaev [12], Hugenholtz and Pines [59], Girardeau and Arnowitt [53], as well as Lieb's simplified approach [72] (see Ref. [73] for a review of these early developments), all rely on special a priori assumptions about the ground state (as the occurrence of condensation) or on the selection of particular terms from likely divergent perturbation series.

Dyson's 1957 estimate [37] for an upper and a lower bound of $E^{\mathrm{g.s.}}/N$ is the only fully rigorous result of that period, though limited to the hard sphere case and with a lower bound that turns out to be off the mark by a factor $(10\sqrt{2})^{-1}$ with respect to the

expected asymptotics. Also, its techniques have turned out to have a crucial influence even on the most recent mathematical treatments of BEC.

In the developments of these investigations, the non-trivial fact emerged that, as long as the system is sufficiently diluted, the *only* other relevant length beyond L is the s-wave scattering length a of the pair interaction (see appendix A of Ref. [80] for a standard rigorous definition of a), through the dimensionless parameter ρa^3, $\rho := N/L^3$ being the density of the gas. Thus assumptions on V have to be made to guarantee that a exists and is finite: it suffices (see [80]) and it will be assumed hereafter (with possibly more stringent conditions) that $\int_{|x|\geqslant R} V(x)\,\mathrm{d}x < \infty$ for some R.

It was only in 1997 that the problem was closed in a wide generality, due to Lieb and Yngvason [79] who completed the rigorous proof of an estimate of the form

$$\left|\frac{E^{\text{g.s.}}/N}{4\pi\mu\rho a} - 1\right| \leqslant \mathcal{O}(\rho a^3)\,. \tag{4.2}$$

More precisely, after an innocent manipulation of the original results one gets the following.

Theorem 4.2.1 (**Ground state energy upper and lower bounds for dilute Bose gases**)**.** Let H_N be the Hamiltonian of a Bose gas as defined in (4.1) and let $E^{\text{g.s.}}$ be its ground state energy. Then, as an upper bound ([77]),

$$\frac{E^{\text{g.s.}}/N}{4\pi\mu\rho a} \leqslant \left(1 - \frac{1}{N}\right)\left(1 + \text{const}\,\sqrt[3]{\frac{4}{3}\pi\rho a^3\Big(1 - \frac{1}{N}\Big)}\,\right) \tag{4.3}$$

provided that periodic boundary conditions are imposed and that ρa^3 is small enough in the sense $\frac{4}{3}\pi\rho a^3 < (1 - 1/N)^{-1}$, and, as a lower bound ([79]),

$$\frac{E^{\text{g.s.}}/N}{4\pi\mu\rho a} \geqslant 1 - \text{const}\,\sqrt[17]{\frac{4}{3}\pi\rho a^3} \tag{4.4}$$

provided that Neumann boundary conditions are imposed (plus the technical restriction that V has finite range, otherwise the error term is possibly different) and that inequality $c_1 N^{-51/3} < \rho a^3 < c_2$ is fulfilled for some suitable positive constants c_1, c_2.

Hence, due to the optimal choice of the boundary conditions (leading to the lowest possible energy for the lower bound and the highest energy for the upper one), one has

$$\lim_{\substack{N\to\infty \\ \rho a^3\to 0}} \frac{E^{\text{g.s.}}/N}{4\pi\mu\rho a} = 1 \tag{4.5}$$

for all boundary conditions and positive (compactly supported) potentials, with $\rho a^3 \to 0$ not faster than $N^{-51/3}$.

Recall that $\rho a^3 \ll 1$ is an expression of *high dilution*, because $a \ll \rho^{-1/3}$, i.e., the mean interparticle distance $\rho^{-1/3}$ is much larger than the scattering length a of the potential. Thus (4.5) amounts to $E_0/N \approx 4\pi\mu\rho a$, provided that the gas is sufficiently populated and diluted. Since $\lim_{\rho\to 0} E_0 = 0$, the system in its ground state does behave as a gas (the ground state is not a many-body bound state).

Also, as already seen in Sec. 1.1, another length scale enters naturally the description of the system, namely the *healing length* $\ell = (8\pi\rho a)^{-1/2}$ given by an uncertainty principle

$\mu\rho a \sim (\hbar/\ell)^2/m$. In the highly diluted regime $\rho a^3 \ll 1$, one has $\rho^{-1/3}/\ell \sim (\rho a^3)^{1/6} \ll 1$, whence $a \ll \rho^{-1/3} \ll \ell$, that is, the impossibility to localize the particles relative to each other. We remark, in passing, that $\rho^{-1/3} \ll \ell$ is crucial in the proof of the lower bound (4.4), where an intermediate length $\tilde{\ell}$ is fixed such that $\rho^{-1/3} \ll \tilde{\ell} \ll \ell$ and particles in the box Ω are distributed into cubic cells of side $\tilde{\ell}$, which turn out to still be highly populated; furthermore, $\tilde{\ell} \ll (\rho a)^{-1/2}$ is necessary for a large enough spectral gap in Temple's inequality that is at the basis of inequality (4.4).

Next to this, there is the ground state kinetic energy per particle: the standard textbooks computation for non-interacting particles gives

$$E^{\mathrm{kin}}/N = \frac{\pi^2\mu}{L^2} \tag{4.6}$$

which amounts to the energy gap in the box, *the natural energy unit of the problem.*

In the more realistic case of an inhomogeneous gas confined into a trap, an analogous picture can be proved [77] (see theorem 3.5.2), provided that the box side L and the density ρ are substituted with the characteristic length of the trap and a suitably defined mean density, respectively: so, as long as N is large enough and ρa^3 is small enough, one has total energy per particle $\sim \rho a$ and kinetic energy per particle $\sim L^{-2}$, and the ratio between these two energies amounts to $\rho a/L^{-2}$, up to a constant $(4/\pi)$ of order 1.

4.3 The Gross-Pitaevskiĭ scaling scenario

In this section the GP scaling is introduced and its main features are discussed.

The prescriptions (in the sense of Sec. 3.2 and 4.1) to this scaling are as follows.

(*i*) It is a "dynamical" limit, where interaction and kinetic energies are still comparable.

(*ii*) It is a limit of "infinite dilution".

If $\rho a^3 \to 0$ within the rates admitted in the analysis reviewed in Sec. 4.2, then the energies to be compared are asymptotically ρa and L^{-2}, so that the constraint is

$$\frac{\rho a}{1/L^2} = \frac{Na}{L} =: g = \text{const} \qquad (\textbf{GP-lim}) \tag{4.7}$$

or, slightly more generally, $Na/L \to g = \text{const}$. The value of such a constant is fixed by the parameters (N_0, L_0, a_0) of the true physical system under consideration: $g = N_0 a_0/L_0$. Since

$$\rho a^3 = g^3 N^{-2}\,, \tag{4.8}$$

this is actually a limit of infinite dilution as $N \to \infty$. Also, ρa^3 vanishes within the range of the validity of the asymptotics $E^{\mathrm{g.s.}}/N \approx 4\pi\mu\rho a$ reviewed in the previous section. Conversely, notice that $\rho a^3 \sim N^{-2}$ is the *only* vanishing rate of the gas parameter guaranteeing the ratio $\rho a/L^{-2}$ to be constant. To keep the quantity $\mathcal{Q}(N, L, a) := Na/L$ (asymptotically) constant, both L and a have to be scaled accordingly: L_N simply accounts for the rate of the box enlargement with N, while a_N scales by modifying the interaction potential V_N in a way that has to be precised later.

Table 4.1. Typical experimental data of Bose-Einstein condensate of alkali vapours. The gas being inhomogeneous, the quoted mean densities have been computed by the authors by integrating over the whole condensate density profile. The quoted scattering lengths have been determined in the same experiments, but the ^{87}Rb one which is taken from later measurements [109]. Scattering length are expressed in terms of a. u. (Bohr radius) $a_{\rm Bohr}$. The ratio g between total and kinetic ground state energy per particle is computed via (4.8).

		N	$\bar\rho$ (cm^{-3})	a ($a_{\rm Bohr}$)	$\bar\rho a^3$	g
^{87}Rb	Wieman, Cornell *et al* (1995) [6]	$4\cdot10^6$	$2.5\cdot10^{12}$	103 ± 5	$4.0\cdot10^{-7}$	186
^{23}Na	Ketterle *et al* (1995) [34]	$5\cdot10^5$	$4\cdot10^{14}$	92 ± 25	$4.7\cdot10^{-5}$	226
^{41}K	Inguscio *et al* (2001) [89]	10^4	$6\cdot10^{11}$	78 ± 20	$4.2\cdot10^{-8}$	1.6

With no claim of an exhaustive review of the B.E.C. vast experimental landscape, typical values of the ratio $\rho a/L^{-2}$ in experiments with alkali vapours are collected in Table 4.1 , where three historically pivotal B.E.C. realizations are quoted.

By no means fixing the ratio (4.7) amounts to claim that the two energies have the same order of magnitude. Actually, in a condensate the kinetic energy is usually highly dominated by the interaction energy. One needs to keep these energies *comparable*, although their ratio is far from 1, to end up in the limit with a system where the *dynamics* can still be investigated. Instead, when one is not interested in the dynamical properties of the Bose gas, another scaling may turn out to be useful, where the kinetic term in the Hamiltonian is neglected. If this is the case when the ratio $\rho a/L^{-2}$ is sent to infinity and the dilution $\rho a^3\to0$, then the scaling goes under the name of *Thomas-Fermi limit*: we discuss it in Sec. 4.4 in connection with the GP limit.

Let us denote the scaling of the scattering length by $a_N\sim1/\eta_N$, with $\eta_N>0$. In principle η_N is allowed to diverge, be constant, or vanish with N. Then (4.7) fixes the characteristic length, the mean density of the system, and the mean energy per particle of the system, as in the first row of Table 4.2. The typical power-law choice $\eta_N=N^\alpha$ is made explicit in the second line of the same Table and then specified when $\alpha=0,\frac23,1$. In particular, the $\alpha=\frac23$ case is the thermodynamic limit ($\rho=$ const) within the GP scaling; notice that conversely the $N/L^3=$ const prescription is not enough to fix alone the whole scaling as $N\to\infty$. In general, the choice of η_N reflects on the limiting behaviour of the interaction energy per particle – see the right column of Table 4.2 – which is still clearly visible in the limit, being by (4.7) comparable to the energy gap in the box. Anyway, since a change $\eta_N\mapsto\widetilde\eta_N$ is just a rescaling of the length unit, any such realization turns out to be mathematically equivalent before the limit.

Thus, the scaling is summarized as

$$a_N=\frac{a_0}{\eta_N}\,,\quad L_N=\frac{(N/N_0)L_0}{\eta_N}\,,\quad \eta_{N_0}=1. \tag{4.9}$$

Table 4.2. Possible realizations of the GP scaling

a_N	L_N	$\rho_N = N/L_N^3$	$1/L_N^2 \sim \rho_N a_N$
$\sim 1/\eta_N$	$\sim N/\eta_N$	$\sim \eta_N^3/N^2$	$\sim (\eta_N/N)^2$
$\sim N^{-\alpha}$	$\sim N^{1-\alpha}$	$\sim N^{3\alpha-2}$	$\sim N^{2\alpha-2}$
fixed	$\sim N$	$\sim N^{-2}$	$\sim N^{-2}$
$\sim N^{-2/3}$	$\sim N^{1/3}$	fixed	$\sim N^{-2/3}$
N^{-1}	fixed	$\sim N$	fixed

The scaling $L_0 \mapsto L_N$ is implemented by modifying the domain of the Hamiltonian

$$H_N = \sum_{i=1}^{N} \big(-\mu\Delta_{x_i} \big) + \sum_{1 \leqslant i < j \leqslant N} V_N(x_i - x_j) \tag{4.10}$$

to be a suitable dense in $L^2\big((\Omega_N)^{\otimes N}, \mathrm{d}x_1 \cdots \mathrm{d}x_N\big)$, Ω_N being the three-dimensional box of side L_N. Of course, for the analogous inhomogeneous case an appropriate confining potential U_N is taken into account, with characteristic length L_N, and the Hamiltonian is

$$H_N = \sum_{i=1}^{N} \big(-\mu\Delta_{x_i} + U_N(x_i) \big) + \sum_{1 \leqslant i < j \leqslant N} V_N(x_i - x_j) \tag{4.11}$$

acting on $L^2(\mathbb{R}^{3N}, \mathrm{d}x_1 \cdots \mathrm{d}x_N)$.

The scaling $a_0 \mapsto a_N \sim \eta_N^{-1}$ is implemented by some redefinition of the pair potential V_N with respect to the true positive compactly supported spherically symmetric pair potential V with scattering length a_0. If $a_N \to 0$ one has a *short-range* limit and, in order for the interaction not to disappear, a *hard-core* limit has to be realized as well. Due to a straightforward rescaling in the definition of the scattering length, one immediately sees that

$$V_N(x) = \eta_N^2 V(\eta_N x) \tag{4.12}$$

has exactly scattering length $a_N = a_0/\eta_N$, provided that V has scattering length a_0. Assume, in addition, that V is bounded in the L^1 norm: then $\eta_N^3 V(\eta_N x) \to b_0\delta(x)$ in the sense of distributions, and $V_N \approx \eta_N^{-1} b_0 \delta(x)$, that is, V_N is an approximate delta function on the same scale η_N^{-1} of the scattering length; here $b_0 := \int V = \|V\|_1$ gives the first Born approximation of a_0, and the Spruch-Rosenberg inequality [105] states

$$a_0 \leqslant \frac{b_0}{8\pi\mu}\,. \tag{4.13}$$

Let us comment on the most useful choice for η_N. The N-body wave functions which H_N acts on are supported (or essentially supported) in $(\Omega_N)^{\otimes N}$ and a typical task in the B.E.C. analysis is to express them as a product of suitable one-particle wave functions supported (or essentially supported) in Ω_N. As summarized in Table 4.2, depending on whether $N/\eta_N \to 0$ or $N/\eta_N \to \infty$, the box (or the trap) where the gas is confined in shrinks to a point or enlarges to become the whole $\mathbb{R}^3$ and, correspondingly, the mean kinetic energy per particle diverges or vanishes. Thus, in these two mathematical regimes, one deals with one-particle wave functions that are normalised in the L^2 sense,

but with H^1 norm that is necessarily either diverging or vanishing. Similarly, the fixed L^2 normalisation forces the L^∞ norm of these one-particle wave functions to vanish or to diverge respectively.

Then the choice

$$\eta_N = N/N_0 \tag{4.14}$$

emerges to be the most appropriate one for explicit computations, because the one-particle wave function L^2, L^∞, H^1 norms do not scale with N. This corresponds to fix the box (or the trap), filling it with more and more particles. This way, the scaled pair potential has scattering length $a_0 N_0/N$ and is an approximate delta function on that scale, weighted with a $1/N$ pre-factor sometimes interpreted as a mean field pre-factor – although its origin is *not* a mean field treatment, where conversely the particle pairing would be described with a long range potential with vanishing strength.

Let us also comment on possible alternatives to (4.12). There, the short-range + hard-core limits are performed simultaneously with N, as $N \to \infty$. A different choice of reasonable generality would be to realize these limits with different rates with N, which is the same to say that scattering length and effective range of V_N scale differently. In the literature, this is often done by scaling

$$V_N(x) = \frac{1}{\eta_N} \xi_N^3 V(\xi_N x) \tag{4.15}$$

with $\xi_{N_0} = \eta_{N_0} = 1$ and $\xi_N = \mathcal{O}(\eta_N)$ as $\eta_N, \xi_N \to \infty$. In fact, under the further assumption that V is *bounded*, one can prove that

$$\lim_{N\to\infty} \eta_N a_N = \begin{cases} a_0 & \text{if } \xi_N/\eta_N \to 1 \\ \dfrac{b_0}{8\pi\mu} & \text{if } \xi_N = o(\eta_N)\,. \end{cases} \tag{4.16}$$

Proof of (4.16) is an adaptation of lemma A.1 in Ref. [42] and is postponed to the end of this chapter (Sec. 4.7). Thus, $\xi_N^3 V(\xi_N x)$ is an approximate delta function on the scale ξ_N^{-1} and (4.15) defines a potential $V_N(x) = \eta_N^{-1}\xi_N^3 V(\xi_N x) \approx b_0 \eta_N^{-1}\delta(x)$ whose range $\sim \xi_N^{-1}$ is possibly much larger than its scattering length $\sim \eta_N^{-1}$. When $\eta_N = N/N_0$, the N-dependence that remains to be prescribed in $\xi_N = \mathcal{O}(N)$ can be unrestrictively chosen to be $\xi_N = (N/N_0)^\beta$, with $0 < \beta \leqslant 1$: the pair potential has scattering length $\sim N^{-1}$ and range $\sim N^{-\beta}$ and is an approximate delta function on the scale $N^{-\beta}$, weighted with the $1/N$ "mean field" pre-factor.

Despite the deep mathematical interest of taking the limit when the two-body interaction scales as (4.15), one has to be warned to consider it a proper GP limit. Indeed, by (6.38) and (4.16), one has

$$\frac{N a_N}{L_N} \xrightarrow{N\to\infty} \frac{N_0 b_0}{L_0 8\pi\mu} \geq \frac{N_0 a_0}{L_0} \tag{4.17}$$

that is, (4.7) is macroscopically violated by a finite quantity. What happens with the interaction (4.15), and can be monitored concretely [42, 44] in the realization $\eta_N \sim N$, is that the kinetic energy is *decreased* with respect to the total energy, when compared with the case of the interaction (4.12): any particle feels only part of the potential when scattering against another particle, having V_N a much longer range than the scattering length, and the effective two-body process is its semiclassical approximation. The modification of the N-body wave function when any two particles are close each other turns

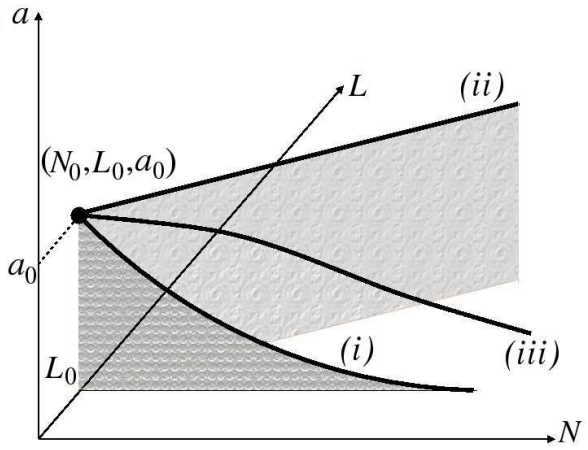


Figure 4.1. Curves of scaling, as $N \to \infty$, in the space of the relevant parameters $(N, L, a, \dots)$. The physical system under investigation lives at the point $(N_0, L_0, a_0, \dots)$. Curves (i) and (ii) represent the GP scaling in the form $a_N \sim N^{-1}$ and $a_N = \text{const}$ respectively (third and fifth rows of Table 4.2). A generic scaling $(N, L, a, \dots) \to \infty$ may be curve (iii).

out to be less pronounced than in the much more singular (4.12) and, consequently, the H^1 norm decreases.

To conclude this section, let us now discuss the topic of the validity of the scenario described so far. Once the (GP) scaling is chosen and a limit result is possibly achieved, the question arises on how well such a result approximates the corresponding physical quantity of the true system under consideration. To this aim, a control is needed of the error terms in the asymptotics of that specific quantity. Unfortunately, a rigorous treatment of the error terms for any of the quantities investigated in the GP limit turns out to be highly non-trivial and at present it is a major open problem. Rigorous results currently known are essentially *asymptotic* results, usually obtained in quite weak or indirect ways, though the task of addressing error terms is receiving more and more interest.

The perspective one should have in checking the validity of the (GP) scaling, *after* a control of the errors has been established, is the following. In the notation of Sec. 4.1, let again f be any of the physical quantities of interest and let $f(N, L, a, \dots)$ denote its dependence on the relevant parameters that in the original system amounts to $N_0, L_0, a_0, \dots$. One seeks information on $f(N_0, L_0, a_0, \dots)$ through the limit

$$\lim_{N\to\infty} f(N, L_N, a_N, \dots)$$

that is, through a special restriction of f to a curve parametrized by N according to the specific scaling in use, GP in our case (4.9). Figure 4.1 gives a pictorial idea of this. Let us assume that an asymptotic result is achieved of the form

$$\lim_{N\to\infty} f(N, L_N, a_N, \dots) = \mathcal{F} \tag{4.18}$$

with an error control of the form

$$\frac{|f(N, L_N, a_N, \dots) - \mathcal{F}|}{|\mathcal{F}|} \leqslant \mathcal{E}(N) = o(1) \tag{4.19}$$

(of course, (4.19) is necessarily formal: $|f - \mathcal{F}|$ can be though as an absolute value for scalars, like the energy, or a norm or a whatever weaker distance for reduced density matrices). Then the problem of to what extent the true $f(N_0, L_0, a_0, \dots)$ can be approximated by $\mathcal{F}$, relies on how satisfactorily small the error $\mathcal{E}(N_0)$ is. Indeed, by construction,

$$f(N, L_N, a_N, \dots)\Big|_{N=N_0} = f(N_0, L_0, a_0, \dots)\,. \tag{4.20}$$

So the overall validity of the GP scaling is ultimately based on the possibility of achieving small relative errors $\mathcal{E}(N_0)$ for the largest number of physical quantities of interest.

Distinct scalings are reasonably expected to be more or less appropriate depending on the different investigated quantities and the GP scaling itself should not be thought to be the absolute best (although it is the scaling through which major results have been achieved, as theorems 3.5.1, 3.5.2, 3.5.3, 5.3.1). Admittedly, the discussion above should have pointed out that at least the GP scaling is central in the study of dilute Bose systems, because it is *the* scaling which guarantees that the kinetic energy does not disappear in the limit (with respect to the total energy) and allows an investigation of the dynamics in the limit.

This leads to the last noticeable feature of this scenario in connection with B.E.C., namely, the GP scaling accounts only for *depletionless condensation*. Notoriously [33, 71], when condensation occurs, due to the interaction a fraction of atoms do not occupy the condensate even at zero temperature because of correlation effects: this *quantum* depletion (to distinguish it from the *thermal* depletion that enhances this phenomenon at positive temperatures), that is, this decrease in the condensate fraction, can be predicted within the Bogolubov theory to be proportional to $(\rho a^3)^{1/2}$. Hence, the GP limit cannot take into account depletion.

As we have already commented, this does *not* mean that it cannot describe B.E.C. with depletion, but only that asymptotic results by construction cannot be expected to account for it. Depletion is zero in the limit and (small and) positive before the limit. If a physical quantity $f(N_0, L_0, a_0, \dots)$ has a sufficiently small error $\mathcal{E}(N_0)$ in the GP limit $\mathcal{F}$, one has to deduce that depletion does not affect $f(N_0, L_0, a_0, \dots)$ strongly and the GP scaling turns out to be an appropriate tool. However, one might expect that other quantities of interest do depend heavily on the depletion and are not conveniently accessible through the same scaling.

4.4 From the GP to the Thomas-Fermi scaling

One may adopt another kind of limit $N \to \infty$ still guaranteeing infinite dilution, but where the ratio between the interaction and the kinetic energy is diverging. In this case the two-body interaction and the box (or trap) length have to scale with N according to the following so-called Thomas-Fermi (TF) limit

$$\begin{array}{rcl} \dfrac{\rho a}{1/L^2} = \dfrac{Na}{L} & \to & \infty \\ \rho a^3 & \to & 0 \end{array} \qquad (\textbf{TF-lim}) \tag{4.21}$$

the generic realization of which, for a homogeneous gas, is

$$a_N = \frac{a_0}{\eta_N}\,, \quad L_N = \frac{(N/N_0)L_0}{\xi_N}\,, \quad \eta_{N_0} = \xi_{N_0} = 1 \tag{4.22}$$

Table 4.3. TF scaling. First row: generic realization for a *homogeneous* Bose gas. In particular: fixed box choice (second row); fixed box + power law choice (third row). Fourth row: realization of the TF scaling for *inhomogeneous* Bose gases in a fixed trap, with mean density (4.30).

a_N	L_N	$\rho_N = N/L_N^3$	$\rho_N a_N^3$	$1/L_N^2$	$\rho_N a_N$
$\sim 1/\eta_N$	$\sim N/\xi_N$	$\sim \xi_N^3/N^2$	$(\xi_N/\eta_N)^3/N^2$	$\sim (\xi_N/N)^2$	$\sim \xi_N^3/(\eta_N N^2)$
$\sim 1/\eta_N$	fixed	$\sim N$	$\sim N/\eta_N^3 \to 0$	fixed	$\sim N/\eta_N \to \infty$
$\sim N^{-\alpha}$	fixed	$\sim N$	$\sim N^{1-3\alpha} \to 0$	fixed	$\sim N^{1-\alpha} \to \infty$
$\sim N^{-\alpha}$	fixed	$\sim N^{s/(s+3)}$	$\sim N^{s/(s+3)-3\alpha} \to 0$	fixed	$\sim N^{s/(s+3)-\alpha} \to \infty$

with the conditions

$$\frac{\xi_N}{\eta_N} \to \infty\,, \quad \frac{\xi_N}{\eta_N N^{2/3}} \to 0\,. \tag{4.23}$$

Indeed

$$\frac{N a_N}{L_N} = \frac{N_0 a_0}{L_0}\,\frac{\xi_N}{\eta_N}$$

is prescribed to diverge and

$$\rho_N a_N^3 = \Big(\frac{N_0 a_0}{L_0}\Big)^3 \Big(\frac{\xi_N}{\eta_N N^{2/3}}\Big)^3$$

is prescribed to vanish.

The generic scalings $L_0 \mapsto L_N$ and $a_0 \mapsto a_N$ in the Hamiltonian H_N are realized as for the GP case. Again, it is natural to fix the realization $\xi_N = N/N_0$, in order to keep the box side fixed and to deal with one-body wave function whose L^2, H^1 and L^∞ norms are not scaling with N (see the discussion at the end of Sec. 4.3). Then, with no loss of generality, one can set $\eta_N = (N/N_0)^\alpha$: then prescriptions (4.23) imply $\frac{1}{3} < \alpha < 1$. A completely analogous scenario holds for inhomogeneous gases, where a suitable mean density has to be adopted. The only difference with the GP treatment of inhomogeneous gases, in this respect, is that such a mean density turns out to scale differently with N (for instance differently than $\bar\rho \sim N$ in the fixed trap GP case), but anyway in such a way that (4.21) is satisfied. Table 4.3 summarizes all that.

One can think the TF scaling as another kind of curve in the $(N, L, a, \dots)$-domain starting from the physical point $(N_0, L_0, a_0, \dots)$ and going to infinity, just like curve (*iii*) in Fig. 4.1. Thus, the same considerations hold concerning the validity of such a scaling, as discussed in Sec. 4.3. In particular, when comparing the GP and the TF scaling keeping the box (or the trap) fixed, the first corresponds to scale the scattering length in such a way that $Na_N = \text{const}$, whereas the second corresponds to scale it much more slowly so that $Na_N \to \infty$.

In the GP limit, the ground state energy of dilute trapped gases is recovered by minimizing the GP energy functional (3.72). In the TF limit, the same role is played by the functional (4.24) obtained by neglecting the kinetic term in the previous one. Of course, the two minima do not coincide for a fixed value of the coupling in front of the quartic term.

Nevertheless, recall that such a coupling is constantly $4\pi\mu a_0 = 4\pi\mu N a_N/N_0$ and that, when both the GP and the TF scaling are performed in the fixed box realization,

the first amounts to $Na_N = \text{const}$, the second to $Na_N \to \infty$. So the interest arises in comparing the two minima (with and without the kinetic term in the functional) letting $Na_N \to \infty$. This goes under the name of *Thomas-Fermi limit of the GP energy* and the result is that the two energies are indeed asymptotically the same.

Definition–Theorem 4.4.1 (TF limit of the GP energy – Lieb, Seiringer and Yngvason [78]). Let $U \in L^\infty_{\text{loc}}(\mathbb{R}^3)$ such that $U(x) \to +\infty$ as $|x| \to +\infty$. Also, let U be locally Hölder continuous and asymptotically equal to some function $\mathcal{U}$ that is homogeneous of some order $s > 0$, i.e., $\mathcal{U}(\lambda x) = \lambda^s \mathcal{U}(x)\ \forall \lambda > 0$.[1]

① Let $\nu \geqslant 0$; after suppressing the kinetic term in the GP functional (3.72), define the Thomas-Fermi energy functional

$$\mathcal{E}^{\text{TF}}_\nu[\gamma] := \int_{\mathbb{R}^3} \big(U\gamma + 4\pi\mu\nu\gamma^2\big)\mathrm{d}x \tag{4.24}$$

on the domain $\mathcal{D}$ of the normalised one-body "densities"

$$\mathcal{D} := \left\{ \begin{array}{c} \gamma \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) : \\ \gamma(x) \geqslant 0,\ U\gamma \in L^1(\mathbb{R}^3), \\ \|\gamma\|_{L^1} = 1 \end{array} \right\}. \tag{4.25}$$

Notice that when ν is taken to be any $\nu_N = Na_N/N_0$, then the $\mathcal{E}^{\text{TF}}_{\nu_N}$'s are a sequence of functionals such that $\mathcal{E}^{\text{TF}}_{\nu_{N_0}}$ has exactly the form of $\mathcal{E}^{\text{GP}}$ without the gradient term.

② For each ν, $\mathcal{E}^{\text{TF}}_\nu$ admits a unique minimizer γ^{TF}_ν given explicitly by

$$\gamma^{\text{TF}}_\nu(x) = \frac{\big[\mu^{\text{TF}}_\nu - U(x)\big]_+}{8\pi\mu\nu} \tag{4.26}$$

where $[\,\cdot\,]_+$ is the positive part and the the TF chemical potential μ^{TF}_ν (i.e., the Lagrange multiplier) is fixed by normalisation. Denote the minimum of each functional as the Thomas-Fermi energy

$$E^{\text{TF}}_\nu := \mathcal{E}^{\text{TF}}_\nu[\gamma^{\text{TF}}_\nu]\,. \tag{4.27}$$

③ The chemical potential has the explicit expression below and is given by a variational principle:

$$\begin{aligned} \mu^{\text{TF}}_\nu &= E^{\text{TF}}_\nu + 4\pi\mu\nu \big\|\,\gamma^{\text{TF}}_N\big\|_2^2 \\ &= \inf_{\mathcal{D}} \int_{\mathbb{R}^3} \big(U\gamma + 8\pi\mu\nu\|\gamma\|_\infty\big)\,\mathrm{d}x\,. \end{aligned}$$

Furthermore, $\nu \mapsto \mu^{\text{TF}}_\nu$ is a concave function that is monotonically increasing to infinity and $\mu^{\text{TF}}_0 = 0$.

④ Compare each $\mathcal{E}^{\text{TF}}_\nu$ with the corresponding family of GP functionals $\mathcal{E}^{\text{GP}}_\nu$ where the coupling is allowed to vary, namely,

$$\mathcal{E}^{\text{GP}}_\nu[\varphi] \;:=\; \int_{\mathbb{R}^3} \big(\mu|\nabla\varphi|^2 + U|\varphi|^2 + 4\pi\mu\nu|\varphi|^4\big)\,\mathrm{d}x \tag{4.28}$$

[1] This is the case, e.g., for a harmonic trap. Such an assumption is for technical convenience, see Ref. [78] for a precise definition of locally Hölder continuity and homogeneity.

on the variational domain (3.73), and with minima E_ν^{GP}. Notice that, as far as $\nu = \nu_N = Na_N/N_0$, (i) one recovers the original GP functional and GP energy when $N = N_0$, (ii) along the GP scaling $Na_N = \text{const} = N_0 a_0$, these $\mathcal{E}_{\nu_N}^{\mathrm{GP}}$'s and $E_{\nu_N}^{\mathrm{GP}}$'s are identically the same.

⑤ With this notation, one can probe the behaviour of the GP functional under other types of scalings, in particular, under the TF scaling that in this setting is realized as $Na_N \to \infty$ ("strong coupling" regime). The result of this *TF limit of the GP energy* is that

$$\lim_{\nu\to\infty} \frac{E_\nu^{\mathrm{TF}}}{E_\nu^{\mathrm{GP}}} = 1\,. \tag{4.29}$$

As a consequence, when performing the TF scaling in the many-body Hamiltonian, the ground state energy per particle is recovered in the limit just by minimizing the energy functional where the kinetic term has been suppressed and the coupling diverges. This can be made rigorous through the same techniques used in the treatment of dilute trapped Bose gases in the GP limit [76] and the result is the following.

Definition–Theorem 4.4.2 (Ground state energy of dilute trapped gases in the TF limit). Let N_0 undistinguishable spinless bosons be confined in a trapping potential U and be paired by a repulsive two-body interaction V. In particular, assume $U \in L^\infty_{\mathrm{loc}}(\mathbb{R}^3)$ such that $U(x) \to +\infty$ as $|x| \to +\infty$. Also, let U be locally Hölder continuous and asymptotically equal to some function $\mathcal{U}$ that is homogeneous of some order $s > 0$. Take $L := \sqrt{2\mu/(\hbar\omega)}$ to be the characteristic length L of the trap as in (3.69). Also, assume $V : \mathbb{R}^3 \to \mathbb{R}$ to be non-negative and spherically symmetric, with finite s-wave scattering length a. The Hamiltonian of the system is then H_{N_0} as in (3.71).

① Scale H_{N_0} to H_N, the N-body Hamiltonian (4.11), according to the TF prescription (4.21) with the standard fixed-trap realization. To do that, a suitable mean density of the inhomogeneous gas has to be defined: the appropriate choice (with the notation of theorem 4.4.1) turns out to be the THOMAS-FERMI DENSITY

$$\bar\rho_N^{\mathrm{TF}} := \frac{1}{N}\int_{\mathbb{R}^3} \left(N\gamma^{\mathrm{TF}}_{\frac{Na_0}{N_0}}(x)\right)^2 \mathrm{d}x = N\left\|\gamma^{\mathrm{TF}}_{\frac{Na_0}{N_0}}\right\|_2^2. \tag{4.30}$$

It can be seen explicitly how it scales with N, using the form (4.26) of the TF minimizer and the assumption that U is asymptotically homogeneous of some positive order s: one finds

$$\bar\rho_N^{\mathrm{TF}} \sim N^{s/(s+3)}\,. \tag{4.31}$$

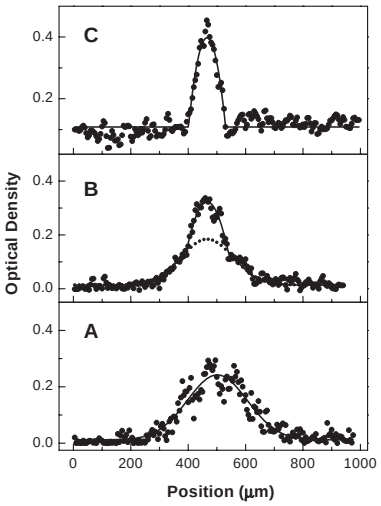
Hence, the scattering length a_N has to scale according to (4.21), in such a way that

$$\begin{array}{ccc} Na_N & \to & \infty \\ \bar\rho_N^{\mathrm{TF}} a_N^3 & \to & 0 \end{array} \tag{4.32}$$

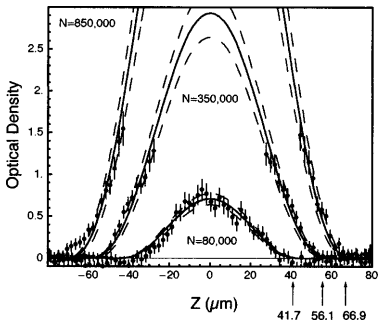
as $N \to \infty$. An unrestrictive power-law choice $a_N \sim N^{-\alpha}$ then imposes $\frac{s}{3(s+3)} < \alpha < 1$. Notice that, since $s > 0$, it necessarily includes the range $\frac{1}{3} < \alpha < 1$ that is characteristic of the TF scaling for a homogeneous gas.

② Let $E_N^{\mathrm{g.s.}}$ be the ground state energy of H_N. In the TF limit $E_N^{\mathrm{g.s.}}$ is recovered as

$$\lim_{N\to\infty} \frac{E_N^{\mathrm{g.s.}}/N}{E_{\nu_N}^{\mathrm{TF}}} = 1\,. \tag{4.33}$$



Density profiles of three samples of ^{41}K after 15 ms of expansion, showing the transition to BEC. **(A)** Thermal sample at $T = 250$ nK. **(B)** Mixed sample at $T = 160$ nK. **(C)** Almost pure condensate. The lines are the best fit with a Gaussian for the thermal component and with an inverted parabola for the condensate component.



Near-resonant images of column density profiles of sodium condensates. The solid curves result from mean-field calculations with $N_0 = 80\,000$, 350 000, and 850 000 atoms in the condensates, respectively, after propagation through the optical system. These number correspond to peak atomic column densities of 442, 1072, and 1826 atoms/μm^2. The asymmetric harmonic-oscillator potential has $\omega_x = \omega_y = 2050$ rad/s and $\omega_z = 170$ rad/s. Arrows give Thomas-Fermi sizes in the z direction. Error bars represent statistical errors. By comparing the theoretical curves to the data we obtain χ^2 values of 0.71, 1.26, and 1.33. The upper and lower dashed curves in each case correspond to N_0 values shifted by $\pm 10\%$.

Figure 4.2. Fit of typical experimental data in the Thomas-Fermi regime. Left: experiment in [89]. Right: experiment in [108].

4.5 GP *vs* TF limit in investigating B.E.C.

It has to be pointed out that, unlike the GP limit, the TF limit does not bring any conclusion on B.E.C., defined in terms of reduced density matrices. Indeed, this tool does not allow to deal with wave functions, but with density profiles only, namely, with the diagonal of the kernel of marginals. It is remarkable, however, that experimentalists often claim that their data on *condensates* fit the density profile of the TF density (4.26), that has the form of the typical "inverted parabola" when the trap is given by a harmonic potential – see, for example, Fig. 4 in Ref. [89], or Fig. 2 in Ref. [108], reproduced in Fig. 4.2.

This happens to be the case because at least in certain experimental regimes the two scalings turn out to be almost equivalent. In current terms, this means the following. If one proceeds from the physical point (N_0, L_0, a_0) along the GP curve (and if the usual technical assumptions on the trapping and interaction potentials are fulfilled), then in the limit $N \to \infty$ the system Bose-condenses in the ground state, the condensate wave function is the corresponding φ^{GP}, and the ground state energy per particle is given by the GP functional evaluated in φ^{GP}. If one moves along the TF curve, then in the limit $N \to \infty$ the ground state energy is provided by the TF functional evaluated in the corresponding γ^{TF}. Some experimental regimes are such that both the two (a priori distinct) limiting results are close enough to the true expectation value of the observable of interest at the physical point (N_0, L_0, a_0) and, hence, are close to each other, so that $|\varphi^{\mathrm{GP}}(x)|^2 \approx \gamma^{\mathrm{TF}}(x)$. Thus, the condensate profile turns out to fit with the TF in good agreement.

Actually, as discussed before in this section, what has to be stressed in this comparison is the crucial role of the quantity $N_0\, a_0/L_0$ for the true physical system. By (4.29), when Na_N/L_N is large, then GP and TF scalings reproduce both the ground state energy. If at the physical point (N_0, L_0, a_0) where the two scalings start from (i.e., the two distinct curves in the diagram of Fig. 4.1) the parameter $N_0\, a_0/L_0$ is already "large enough", then one expects that taking $N \to \infty$ while $g = Na_N/L_N = \text{const}$ (GP lim) or while $Na_N/L_N \to \infty$ (TF lim) gives essentially the *same* description (though only the GP scaling gives access to the full kernel of the reduced density matrices, by which one defines B.E.C.).

Typical experimental values of $N_0\, a_0/L_0$ are listed in the last column of Table 4.1. Notice that in the aforementioned ^{41}K experiment [89], where the density profile of the condensate fits with the TF inverted parabola, one has $N_0\, a_0/L_0 \sim 1.6$; hence, how large this parameter has to be in order that GP and TF scaling give the same description is a quite subtle matter.

Recall that in Table 4.1 such values are computed via (4.8), that is, $N_0\, a_0/L_0 = (\rho_0\, a_0^3\, N_0^2)^{1/3}$: this is because in the same experiment one could measure directly a_0 and N_0, while for ρ_0 one takes a mean value $\bar\rho$ by averaging the experimental particle distribution. To have a check, one can alternatively compute L_0 itself via

$$L_0 \sim \sqrt{\frac{2\mu}{\hbar\omega}}\,, \tag{4.34}$$

that is, through the details of the trapping potential U. In [89] U is actually a magnetic trap consisting of a Ioffe-Prichard potential in quadrupole Ioffe configuration (see [46] for details), which in the current setting reads

$$U(x,y,z) = \frac{m}{2}\big(\omega_{\rm rad}(x^2+y^2) + \omega_{\rm ax} z^2\big)$$

with radial and axial frequencies $\omega_{\rm rad}/(2\pi) = 200\,\text{Hz}$ and $\omega_{\rm ax}/(2\pi) = 16\,\text{Hz}$, and with $m =$ atomic mass of ^{41}K. Then the ground state energy $\hbar\omega$ of $-\mu\Delta + U(\mathbf{r})$ is $\hbar\omega = \hbar\omega_{\rm rad} + \frac{1}{2}\hbar\omega_{\rm ax}$ and by (4.34) the characteristic length of the trap is $L_0 = \sqrt{2\mu/(\hbar\omega)}$ (one neglects here the anisotropy of the trap). So altogether $N_0\, a_0/L_0 \sim 1.54$, comparable to the value of 1.6 listed in Table 4.1.

An agreement between GP and TF density profiles is observed even in a much larger regime of $N_0\, a_0/L_0$: this is the case, for instance, of the experiment with ^{23}Na in [108]: there one has $N_0 = 8\cdot 10^4$, $a_0 = 27.5$ Å, $\omega_{\rm rad} = 2050\,\text{s}^{-1}$, $\omega_{\rm ax} = 170\,\text{s}^{-1}$, $m = m(^{23}\text{Na})$, whence $N_0\, a_0/L_0 \sim 193$. (The scattering length in this case differs considerably from the one quoted in Table 4.1 because of the distinct hyperfine levels of ^{23}Na involved in the condensation phenomenon.)

4.6 Low-dimensional scalings

Despite that the present discussion has been centred in the natural three-dimensional setting, the low-dimensional behaviour of dilute Bose gas in elongated or disc-shaped traps deserves great interest.

The Thomas-Fermi scaling as discussed in Sec. 4.4 has a one- and two-dimensional counterpart too – see Ref. [75, 78, 76]. It has also turned out to be a fundamental tool in the study of trapped rotating condensates (see Ref. [32, 24] and references therein).

In particular, although entering the peculiarities of the low-dimensional settings goes beyond the scope of this chapter, where instead stress has been given to the role of the scalings themselves, we mention that in the two-dimensional case the TF scaling is especially relevant, for it is the natural and the more appropriate alternative to the GP scaling. Indeed (see the discussion in [78]), in the latter the scattering length turns out to decrease exponentially with N and the two-dimensional analog of the fixed quantity g in (4.7) is typically very large, so that a limit $g \to \infty$ as $N \to \infty$ (i.e., the TF limit in two dimensions) is meaningful.

4.7 Appendix: scaling laws for the scattering length

In this appendix, we give the proof of the technical property (4.16), essentially by adapting the original lemma A.1 in Ref. [42]. Restating it in the form (4.35), one can appreciate the substantial content of the original result; that is, *whatever* the realization of a delta-like potential $\xi_N^3 V(\xi_N x) \approx b\,\delta(x)$ is, *provided* that $\xi_N = o(\eta_N)$, then the scattering length of $\eta_N^{-1}\xi_N^3 V(\xi_N x)$ does not scale like a_0/η_N, as in the proper GP case $\xi_N \sim \eta_N$, but instead like $(\frac{b}{8\pi\mu})/\eta_N$, i.e., with the same behaviour but a higher constant. So this is a feature which does not depend on the particular realization $V_N(x) = N^{3\beta-1}V(N^\beta x)$, $0 < \beta < 1$.

Theorem 4.7.1. Let $V : \mathbb{R}^3 \to \mathbb{R}$ be bounded, compactly supported, spherically symmetric, positive, and smooth. Let b be its L^1-norm. Let $V_N(x) \equiv \eta_N^{-1}\xi_N^3 V(\xi_N x)$ with $\eta_N, \xi_N \to \infty$ as $N \to \infty$ in such a way that $\xi_N = o(\eta_N)$. Then, if V_N is regarded as a two-body quantum mechanical potential for particles of mass m and $\mu := \frac{\hbar^2}{2m}$, the s-wave scattering length a_N has the asymptotic behaviour

$$\lim_{N\to\infty} \eta_N a_N = \frac{b}{8\pi\mu}\,. \tag{4.35}$$

Proof. Let $\mathrm{supp}(V) = \mathcal{B}_R$, the ball of radius R centred at the origin. Then $\mathrm{supp}(V_N) = \mathcal{B}_{R/\xi_N}$. One has

$$\begin{aligned} b_N \;\equiv\; \|V_N\|_1 &= \frac{\|V\|_1}{\eta_N} = \frac{b}{\eta_N}\,, \\ \|V_N\|_\infty &= \frac{\xi_N^3}{\eta_N}\|V\|_\infty\,, \end{aligned} \tag{4.36}$$

so that (6.38) reads

$$8\pi\mu\eta_N a_N \leqslant b\,. \tag{4.37}$$

Also recall [80] that

$$8\pi\mu a_N = \int_{\mathbb{R}^3} V_N(x) f_N(x)\mathrm{d}x \tag{4.38}$$

where f_N is the spherically symmetric solution of

$$\begin{cases} \big(-2\mu\Delta + V_N\big) f_N &= 0 \\ f_N(x=\infty) &= 1 \end{cases}\,. \tag{4.39}$$

Let $r = |x|$, $v_N(r) \equiv V_N(x)$ and $f_N(x) \equiv g_N(r)/r$: then $g_N(r) \geqslant 0\ \forall\, r \geqslant 0$,

$$-\mu g_N''(r) + \frac{1}{2} v_N(r) g_N(r) = 0\,,$$

and for $r \geqslant R/\xi_N$ one has $g_N(r) = r - a_N$ (whence $g'(r) = 1$). For $0 \leqslant r \leqslant R/\xi_N$,

$$\begin{aligned}\frac{R}{\xi_N} - a_N - g_N(r) &= g_N(R/\xi_N) - g_N(r)\\ &= \int_r^{R/\xi_N} g_N'(x)\mathrm{d}x\\ &= \int_r^{R/\xi_N} \Big\{ g_N'(R/\xi_N) - \int_x^{R/\xi_N} g_N''(y)\,\mathrm{d}y \Big\}\,\mathrm{d}x\\ &= \Big(\frac{R}{\xi_N} - r\Big) - \int_r^{R/\xi_N}\int_x^{R/\xi_N} \frac{v_N(y)}{2\mu} g_N(y)\,\mathrm{d}y\,\mathrm{d}x\\ &\leqslant \frac{R}{\xi_N} - r\,,\end{aligned}$$

then $g_N(r) \geqslant r - a_N$. So

$$f_N(x) \geqslant \begin{cases} 1 - a_N/|x| & ,\ |x| \geqslant a_N \\ 0 & ,\ |x| \leqslant a_N \end{cases} \tag{4.40}$$

and (4.38) gives

$$\begin{aligned} 8\pi\mu a_N &\geqslant \int\limits_{|x|\geqslant a_N} V_N(x)\Big(1 - \frac{a_N}{|x|}\Big)\mathrm{d}x\\ &= \int\limits_{|x|\geqslant a_N} V_N(x)\mathrm{d}x - \int\limits_{|x|\geqslant a_N} V_N(x)\frac{a_N}{|x|}\mathrm{d}x\\ &\geqslant \int\limits_{|x|\geqslant \frac{b_N}{8\pi\mu}} V_N(x)\,\mathrm{d}x - \int_{\mathbb{R}^3} V_N(x)\frac{b_N}{|x|}\mathrm{d}x\\ &\geqslant \int_{\mathbb{R}^3} V_N(x)\,\mathrm{d}x - \int\limits_{|x|\leqslant \frac{b_N}{8\pi\mu}} V_N(x)\mathrm{d}x - b_N\int_{\mathbb{R}^3} V_N(x)\frac{\mathrm{d}x}{|x|}\quad . \end{aligned} \tag{4.41}$$

Due to (4.36), the last two summands in the r.h.s. of the above inequality are estimated as

$$\begin{aligned}\int\limits_{|x|\leqslant \frac{b_N}{8\pi\mu}} V_N(x)\,\mathrm{d}x &\leqslant \frac{4}{3}\pi\Big(\frac{b_N}{8\pi\mu}\Big)^3\|V_N\|_\infty\\ &= \frac{\|V\|_1^3\|V\|_\infty}{384\pi\mu^3}\frac{1}{\eta_N}\Big(\frac{\xi_N}{\eta_N}\Big)^3 \equiv \frac{C_1}{\eta_N}\Big(\frac{\xi_N}{\eta_N}\Big)^3\end{aligned} \tag{4.42}$$

and as

$$\begin{aligned} b_N\int_{\mathbb{R}^3} V_N(x)\frac{\mathrm{d}x}{|x|} &= \int_{\mathbb{R}^3}\frac{1}{\eta_N}\xi_N^3 V(\xi_N x)\frac{\mathrm{d}x}{|x|}\\ &= \frac{b_N}{\eta_N}\int_{\mathbb{R}^3}\frac{V(\mathbf{x})\,\mathrm{d}\mathbf{x}}{|\mathbf{x}|/\xi_N}\\ &\leqslant \frac{\|V\|_1}{\eta_N}\Big(\frac{\xi_N}{\eta_N}\Big)\|V\|_\infty\int\limits_{|\mathbf{x}|\leqslant R}\frac{\mathrm{d}\mathbf{x}}{|\mathbf{x}|}\\ &= \frac{2\pi R^2\|V\|_1\|V\|_\infty}{\eta_N}\Big(\frac{\xi_N}{\eta_N}\Big) \equiv \frac{C_2}{\eta_N}\Big(\frac{\xi_N}{\eta_N}\Big)\,,\end{aligned} \tag{4.43}$$

respectively, so by plugging (4.42) and (4.43) into (4.41), one gets

$$8\pi\mu\eta_N a_N \geqslant b - C_1\Big(\frac{\xi_N}{\eta_N}\Big)^3 - C_2\Big(\frac{\xi_N}{\eta_N}\Big)\,. \tag{4.44}$$

Since $\xi_N/\eta_N \to 0$, (4.37) and (4.44) give (4.35). □

Chapter 5

Strengthened convergence for reduced density matrices in the derivation of the Gross-Pitaevskiĭ equation

5.1 Time-stability of condensation

An issue which is extensively discussed in the Mathematical Physics community working on B.E.C.-related problems is that of TIME-STABILITY OF CONDENSATION. Its physical formulation is: given a condensate Bose gas confined in a trap, investigate whether the system continues evolving in time as a condensate once the trap is removed and the gas is left to expand under the interparticle repulsive interaction only. In general this turns out to depend on the chemical nature of the particles, on the conditions under which the condensate is initially prepared (trapping potential, dilution, etc.), and on the interaction among particles causing their expansion after removing the confinement.

Persistence of B.E.C. in a many-body system evolving in time, according to definitions 3.3.2 and 3.4.1, means that the corresponding time-dependent one-particle reduced density matrix is essentially a projection at any time, thus identifying the time-dependent one-body condensate wave function, say φ_t.[1] Experiments can reveal $|\varphi_t|^2$, thus its behaviour with t can be studied. Hence, one ends up with the mathematical counterpart of this problem, that is, the problem of deriving the law by which the condensate wave function evolves in time.

Heuristic arguments and theoretical investigations yet based on ill-posed derivations or unproved assumptions, dating back the 50's of last century, as well as the fit of most recent experimental data, lead to a specific dependence: if suitably prepared and left free to expand, the condensate preserves its condensed phase and the condensate wave function is the solution of a cubic nonlinear Schrödinger equation which goes under the name of time-dependent Gross-Pitaevskiĭ equation. The mathematical task is then to derive rigorously such a nonlinear one-body equation, starting from the linear many-body Schrödinger equation governing the evolution of the whole system.

[1] This is not a partial derivative, which instead would be denoted by $\partial_t \varphi_t$. The established convention in the related literature is to denote the time and the space dependence "asymmetrically", by $\varphi_t(x)$.

This chapter, following Ref. [85], reports the state of the art of this issue together with some recent improvements.

In Sec. 5.2 we discuss the standard way to derive it *formally*. In comparison to that, Sec. 5.3 reports the most recent *rigorous* results in the derivation of the cubic nonlinear Scrödinger equation of interest in the present context. In the B.E.C. terminology, they deal with a system where asymptotic 100% condensation is preserved in time, hence, they provide the following diagram:

$$\begin{array}{ccccc} \gamma_N & \xrightarrow{\text{partial trace}} & \gamma_N^{(k)} & \xrightarrow{N\to\infty} & |\varphi\rangle\langle\varphi|^{\otimes k} \\ \text{many-body linear dynamics} \downarrow & & \downarrow & & \downarrow \text{GP}-\text{equation} \\ \gamma_{N,t} & \xrightarrow{\text{partial trace}} & \gamma_{N,t}^{(k)} & \xrightarrow{N\to\infty} & |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \end{array} \tag{5.1}$$

The study of a diagram like (5.1) is part of an analysis which goes beyond this B.E.C.-related topic and concerns the derivation of effective nonlinear (classical or quantum) evolutionary one-body equations from the linear many-body dynamics. It involves a general strategy which has been developed in a series of investigations in the last 30 years. We review the scheme of this analysis in Sec. 5.4, together with a historic survey in Sec. 5.5.

The core of this chapter is then presented in Sec. 5.6, 5.7, and 5.8, where, following [85], we show some improvements in the topologies for the convergence of time-dependent marginals, providing a strengthened unified version of the recent results discussed in Sec. 5.3.

5.2 Formal derivation of the time-dependent Gross-Pitaevskiĭ equation

Before proceeding along the main line of this chapter, towards the rigorous derivation of the time-dependent Gross-Pitaevskiĭ equation, this section presents the typical *formal* derivation from the Condensed Matter Physics viewpoint (see, e.g., Chap. IV-V of [71], Chap. 7 of [96], Chap. 5 of [97]). In the current formalism it goes at follows.

Chose for concreteness the three-dimensional setting. The main point is the formal replacement of the true pair potential with a delta-interaction. In the N-body Hamiltonian

$$H_N = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_{x_i} + \sum_{1\leqslant i<j\leqslant N} V_N(x_i - x_j) \tag{5.2}$$

for a gas expanding under the repulsive interparticle interaction only, one substitutes

$$V_N(x) \mapsto \frac{g}{N}\,\delta(x) \tag{5.3}$$

that is, a delta potential scaling in the mean field sense and with a magnitude fixed by the coupling g. The choice of g that turns out to reproduce the correct GP equation is

$$g = 8\pi\mu a = \frac{4\pi\hbar^2 a}{m} \tag{5.4}$$

($\mu = \frac{\hbar^2}{2m}$ as usual), where a is the s-wave scattering length of the two-body process.

Ansatz (5.3) and (5.4) encodes the physical assumption that particles are coupled by a *hard-core* and *short-range* two-body repulsive potential and it is even thought as "possibly the most important result in the whole of the Physics of the dilute ultra-cold alkali gases".[2]

The outcoming Hamiltonian

$$H_N^{(\delta)} = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_{x_i} + \frac{1}{N}\sum_{1\leqslant i<j\leqslant N} g\,\delta(x_i - x_j)\,, \tag{5.5}$$

apart from any well-posedness question, can be treated formally in the Schrödinger equation

$$i\hbar\,\partial_t\Psi_{N,t} = H_N^{(\delta)}\Psi_{N,t}\,. \tag{5.6}$$

When regarding (5.6) as an Euler-Lagrange equation, its solutions $\Psi_{N,t}$ are critical points of the corresponding functional (namely, the many-body action) and, hence, are given by the stationary condition

$$0 = \delta\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\big(-i\hbar\,\partial_t\Psi_{N,t} + H_N^{(\delta)}\,\Psi_{N,t}\big)\,. \tag{5.7}$$

One then searches for critical points of the form

$$\Psi_{N,t} = \varphi_t^{\otimes N}\,, \qquad \varphi_t \in H^1(\mathbb{R}^3)\,, \quad \|\varphi_t\|_{L^2(\mathbb{R}^3)} = 1 \quad \forall\, t\,, \tag{5.8}$$

and looks for the evolutionary equation that $\varphi_\cdot$ must satisfy.

By treating the delta operators in $H_N^{(\delta)}$ as distributions in the integral (5.7),

$$\begin{aligned}\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\big(-i\hbar\,\partial_t\Psi_{N,t} + H_N^{(\delta)}\Psi_{N,t}\big) &= -\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\,i\hbar\,\partial_t\Psi_{N,t} \\ &\quad + \mu\sum_{i=1}^{N}\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\big(-\Delta_{x_i}\big)\Psi_{N,t} + \frac{g}{N}\sum_{1\leqslant i<j\leqslant N}\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\,\delta(x_i-x_j)\Psi_{N,t}\,,\end{aligned} \tag{5.9}$$

one gets

$$\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\,i\hbar\,\partial_t\Psi_{N,t} = N\int_{\mathbb{R}^3}\overline{\varphi_t(x)}\,i\hbar\,\partial_t\varphi_t(x)\,\mathrm{d}x \tag{5.10}$$

$$\sum_{i=1}^{N}\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\big(-\Delta_{x_i}\big)\Psi_{N,t} = N\int_{\mathbb{R}^3}|\nabla\varphi_t(x)|^2\mathrm{d}x \tag{5.11}$$

$$\begin{aligned}\sum_{1\leqslant i<j\leqslant N}\int_{\mathbb{R}^{3N}} \overline{\Psi}_{N,t}\,\delta(x_i-x_j)\Psi_{N,t} &= \frac{N(N-1)}{2}\int_{\mathbb{R}^6}\delta(x-y)|\varphi_t(x)|^2|\varphi_t(y)|^2\mathrm{d}x\,\mathrm{d}y \\ &= \frac{N(N-1)}{2}\int_{\mathbb{R}^3}|\varphi_t(x)|^4\mathrm{d}x\,.\end{aligned} \tag{5.12}$$

For N large enough (so that $N-1\sim N$), (5.7) becomes

$$0 = \delta\left[N\int_{\mathbb{R}^3}\Big(-\overline{\varphi_t(x)}\,i\hbar\,\partial_t\varphi_t(x) + \mu\,|\nabla\varphi_t(x)|^2 + \frac{g}{2}\,|\varphi_t(x)|^4\Big)\,\mathrm{d}x\right]\,, \tag{5.13}$$

[2] Legget [71], Sec. IV.C

then, for the consistency of prescription (5.8), $\varphi_\cdot$ has to be a critical point of the functional (5.13), hence, a solution of the outcoming Euler-Lagrange equation

$$i\hbar\,\partial_t\varphi_t(x) \;=\; -\mu\,\Delta_x\varphi_t(x) + g\,|\varphi_t(x)|^2\,\varphi_t(x) \;=\; \frac{\delta\,\mathcal{E}_{\frac{g}{8\pi\mu}}[\,\varphi_t\,]}{\delta\,\overline{\varphi_t}} \tag{5.14}$$

where

$$\mathcal{E}_\nu[\,\varphi\,] \;:=\; \int_{\mathbb{R}^3} \big(\,\mu\,|\nabla\varphi(x)|^2 + 4\pi\mu\,\nu\,|\varphi_t(x)|^4\,\big)\,\mathrm{d}x\,. \tag{5.15}$$

With the choice $g = 8\pi\mu a$, (5.14) is the desired time-dependent GP equation (compare it with (5.27)), and functional $\mathcal{E}_{\frac{g}{8\pi\mu}}$ is the corresponding GP functional (compare it with (4.28)).

What makes this derivation formal is the treatment of the Hamiltonian $H_N^{(\delta)}$. First, to fix it unambiguously, an explicit self-adjoint extension $\widetilde{H}_N^{(\delta)}$ of the symmetric operator defined in (5.5) has to be chosen. This identifies the domain of the extension and the action of the delta operators on it. This action is *not* the simple "multiplicative" action that has been used in (5.9) by interpreting δ as a distribution. Instead, each extension is related in a suitable sense to a choice of boundary conditions on the hyperplanes $x_i = x_j$ in $\mathbb{R}^{3N}$. [3]

Moreover, what makes this derivation wrong is the occurrence of an Efimov effect: indeed it can be proved that already for $N = 3$ any *local* [4] extension $\widetilde{H}_3^{(\delta)}$ is unbounded below, making the system unstable.[5] Another instability shows up in two dimensions: there one can prove ([35]) that $\widetilde{H}_N^{(\delta)}$ is bounded below, but one can argue that the infimum of the spectrum goes to $-\infty$ faster than N (presumably the behaviour is $\sim N^2$).

It is noteworthy noticing that in the formal derivation the choice of g is made to let (5.14) agree with the GP equation and one models the true interaction with an effective

$$V_N(x) = 8\pi\mu a\,\frac{1}{N}\,\delta(x)\,; \tag{5.16}$$

on the other side, in the rigorous (although indirect) approach one scales the true interaction with the GP scaling and gets

$$V_N(x) = N^2V(Nx) \sim \frac{b}{N}\,\delta(x) \tag{5.17}$$

where $\frac{b}{8\pi\mu}$ is the Born approximation for a. Contradiction between (5.16) and (5.17) is only apparent. In the rigorous argument one deals with the (unique) self-adjoint extension of $\sum_i(-\mu\Delta_{x_i}) + \sum_{i<j} V_N(x_i - x_j)$, while in the formal argument one ignores to fix one among the infinitely many self-adjoint extensions of (5.5) and no rigorous comparison is known between the two operators in the limit. Moreover in the rigorous treatment one does *not* plug $V_N \sim \frac{b}{N}\,\delta$ directly in the Schrödinger equation,[6] but instead keeps track of the action of V_N in the corresponding hierarchy of density matrices (see

[3] A precise meaning to the formal expression (5.5) was first given by Minlos and Faddeev [88] and more recently by Dell'Antonio, Figari and Teta [35].

[4] in the sense that locality is a property of the corresponding boundary conditions on the union of the hyperplanes, see [5]

[5] This serious and entirely new difficulty with respect to the $N = 2$ case, which falls instead within the theory ([5]) of the one-particle Schrödinger equation with zero range potential, was first pointed out by Minlos and Faddeev [87] and then studied by Dell'Antonio, Figari and Teta [35].

[6] unless in the one-dimensional case where δ is easily controllable as an operator

Sec. 5.4) and then makes a highly non-trivial control of the limit $N \to \infty$ in such a hierarchy: this way the correct coupling $8\pi\mu a$ emerges in the limit.

Incidentally, (5.16) usually suggests to interpret the GP equation as a phenomenological *mean field type* equation. Since (5.16) is only formal, this interpretation has not solid grounds: Hamiltonian $\widetilde{H}_N^{(\delta)}$ after the δ-replacement does not describe a stable system. Conversely, one might emphasize that the same GP equation is rigorously recovered in the GP limit, which being a short-range + hard-core limit is *not* a mean field scaling (the latter is characterized, instead, by a long range interaction with vanishing strength). Probably the most appropriate perspective is simply to declare the conditions of validity of the GP equation, that is, the low density regime.

5.3 Rigorous derivation of the time-dependent cubic nonlinear Schrödinger equation (Gross-Pitaevskiĭ equation)

Two independent construction of the commutative diagram (5.1) have been recently achieved, by Erdős, Schlein and Yau [44, 43], and by Adami, Golse and Teta [3]. The first is set in the $d = 3$ dimensional case, although it can be extended to $d = 1, 2$, whereas the second is intrinsically set in $d = 1$. The two proofs are structured with the same scheme, but they differ both in the techniques and in the topologies that control the existence and the uniqueness of the limit $\gamma_{N,t}^{(k)} \to |\varphi_t\rangle\langle\varphi_t|^{\otimes k}$. In this section, both the results are reviewed, with some comments on their setting that will be needed in the sequel. The year we indicate for each theorem stays for when the statement and the proof was made public by the authors for the first time.

Let us start with the three-dimensional result by Erdős et al., which has been established in [44] after the intermediate results [41], [40] (with Elgart) and [42] – see also [100] for a comprehensive review.

Theorem 5.3.1 (**Erdős, Schlein and Yau (2006), [44]**). Let $V : \mathbb{R}^3 \to \mathbb{R}$ be a non-negative, smooth, compactly supported, spherically symmetric function. Regarding it as a two-body quantum mechanical potential for spinless particles of mass m, let a be its s-wave scattering length. Also, let V be small enough when measured in terms of the dimensionless quantity $\frac{\alpha\hbar^2}{2m}$, with

$$\alpha := \sup_x |x|^2 V(x) + \int_{\mathbb{R}^3} \frac{V(x)}{|x|}\,\mathrm{d}x\,. \tag{5.18}$$

Equivalently,

$$\alpha = \|r^2\mathcal{V}\|_{L^\infty(\mathbb{R}_+)} + 4\pi\|r\mathcal{V}\|_{L^1(\mathbb{R}_+)} \tag{5.19}$$

where $r = |x|$ and $\mathcal{V}(r) := V(x)$. Consider a three-dimensional system of N spinless undistinguishable bosons of mass m interacting with a two-body potential V_N defined by

$$V_N(x) := N^2 V(Nx) \tag{5.20}$$

so that the N-body Hamiltonian is

$$H_N = -\frac{\hbar^2}{2m}\sum_{i=1}^N \Delta_{x_i} + \sum_{1\leqslant i<j\leqslant N} N^2 V\big(N(x_i - x_j)\big) \tag{5.21}$$

acting as a self-adjoint operator on $L^2(\mathbb{R}^{3N})$. Assume that the system is prepared in the initial pure state $\Psi_N \in L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$, with $\|\Psi_N\|_{L^2} = 1$, satisfying either the two properties (A_1), (A_2) or the two properties (B_1), (B_2) below.

A_1) **Uniformly bounded k-th moment of the energy per particle**:

$$\exists\, C > 0 \,:\, \forall k \geqslant 1\ \forall N \geqslant 1 \quad \Big(\Psi_N, \Big(\frac{H_N}{N}\Big)^k \Psi_N\Big) \leqslant C^k \tag{5.22}$$

(the bound on the k-th power of H_N/N is uniform in N, but not in k).

A_2) **100% asymptotic B.E.C.**:

$$\begin{array}{l} \exists\, \varphi \in L^2(\mathbb{R}^3)\,,\ \|\varphi\|_{L^2} = 1 \,: \\ \forall k \geqslant 1 \quad \gamma_N^{(k)} \to |\varphi\rangle\langle\varphi|^{\otimes k} \quad \text{as } N \to \infty\,. \end{array} \tag{5.23}$$

B_1) **Uniformly bounded energy per particle**:

$$\exists\, C > 0 \,:\, \forall N \geqslant 1 \quad (\Psi_N, \frac{H_N}{N}\Psi_N) \leqslant C\,. \tag{5.24}$$

B_2) **Asymptotic factorisation**:

$$\begin{array}{l} \exists\, \varphi \in L^2(\mathbb{R}^3)\,,\ \|\varphi\|_{L^2} = 1 \,: \\ \forall\, N \geqslant 1\,, \forall k = 1, \dots, N \\ \exists\, \Xi^{(N-k)} \in L^2(\mathbb{R}^{3(N-k)})\,, \|\Xi^{(N-k)}\|_{L^2} = 1 \,: \\ \|\Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)}\|_{L^2} \to 0 \quad \text{as } N \to \infty\,. \end{array} \tag{5.25}$$

Then $\varphi \in H^1(\mathbb{R}^3)$. Moreover, if $\gamma_{N,t}^{(k)}$ are the k-th marginals of $\Psi_{N,t} := e^{-\frac{i}{\hbar} H_N t}\Psi_N$ (the time-evolution of Ψ_N under the Hamiltonian H_N), then

$$\forall\, k \geqslant 1\ \forall\, t \in \mathbb{R} \quad \gamma_{N,t}^{(k)} \to |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \quad \text{as } N \to \infty \tag{5.26}$$

where $\varphi_t \in H^1(\mathbb{R}^3)$ is the solution of the TIME-DEPENDENT GROSS-PITAEVSKIĬ EQUATION with initial datum φ :

$$\begin{cases} i\hbar\, \partial_t \varphi_t &= -\dfrac{\hbar^2}{2m}\Delta\varphi_t + \dfrac{4\pi\hbar^2}{m} a\, |\varphi_t|^2 \varphi_t \\ \\ \varphi_t\big|_{t=0} &\equiv \varphi \end{cases} \tag{5.27}$$

Limits in (5.23) and (5.26) are understood with respect to the weak-$*$ topology of $\mathcal{L}^1(L^2(\mathbb{R}^{3k}))$, namely, in trace against any compact operator on $L^2(\mathbb{R}^{3k})$.

Thus, ESY theorem essentially states that when the system is left to evolve under H_N, it still shows 100% asymptotic B.E.C. and the condensate wave function φ_t is the evolution of φ under the GP equation.

Let us collect some explanatory comments.

✓ To start with, the real crucial hypotheses on V, which is a major open problem to relax, turn out to be positivity of V itself and smallness of the parameter (5.18), while smoothness, compact support, and spherical symmetry are just technical assumptions.

- ✓ Recall that, according to the standard definition of the s-wave scattering length (see, e.g., appendix A of [80]), V_N has scattering length $a_N = a/N$. Recall, also, that $NV_N(x) \xrightarrow{\mathcal{S}'} b\,\delta(x)$ as $N \to \infty$, where $b := \|V\|_{L^1}$: hence, the two-body V_N is an approximate delta function on the same scale N^{-1} of its scattering length. The dependence $N \mapsto H_N$ is actually the explicit realization of the GP scaling in use (see Sec. 4.3).

- ✓ The (B) hypotheses are interchangeable with the (A) ones, leading to the same thesis: to relax the energy condition from (A_1) to (B_1) one introduces a technically stronger factorisation property (B_2) than (A_2) (although this property is only apparently stronger, as discussed in Sec. 7.2). Since it turns out that $(B_2) \Rightarrow (A_2)$, *both have the meaning of B.E.C.* according to definition 3.4.1. The (B) version of the theorem follows from the (A) version by suitably smoothing an initial datum satisfying (B) with a cut off of the high energy component. The modified initial datum one ends up with, is admissible for the (A) hypotheses and the theorem follows for it. Finally, the cut off is removed.

- ✓ *Two classes of physically interesting initial states satisfying the* (B) *assumptions above are explicitly shown to exist*: totally factorised wave functions $\varphi^{\otimes N}$, for an arbitrary $\varphi \in H^1(\mathbb{R}^3)$, and almost product wave functions with ground state like short-scale correlations (see [43] and appendices B and C of [44]). We will discuss them in Chap. 6 in the context of the correlation structure in a condensate

In comparison with ESY theorem, Adami et al. have proved an analogous theorem in [3], after the intermediate result [2] (with Bardos).

Theorem 5.3.2 (**Adami, Golse and Teta (2005), [3]**)**.** Let $V : \mathbb{R} \to \mathbb{R}$ be a non-negative function in the Schwartz space $\mathcal{S}(\mathbb{R})$ with L^1-norm b. Consider a one-dimensional system of N spinless undistinguishable bosons of mass m interacting with a two-body potential V_N defined by

$$V_N(x) := N^{\beta-1} V(N^\beta x) \tag{5.28}$$

with $0 < \beta < 1$, so that the N-body Hamiltonian is

$$H_N = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \frac{\partial^2}{\partial x^2} + \sum_{1 \leqslant i < j \leqslant N} N^{\beta-1} V\big(N^\beta (x_i - x_j)\big) \tag{5.29}$$

acting as a self-adjoint operator on $L^2(\mathbb{R}^N)$. Assume that the system is prepared in the totally factorised initial pure state $\Psi_N = \varphi^{\otimes N} \in L^2_{\mathrm{sym}}(\mathbb{R}^N)$, with $\varphi \in \mathcal{S}(\mathbb{R})$ and $\|\varphi\|_{L^2} = 1$, such that any k-th moment of the energy per particle is uniformly bounded in N:

$$\exists\, C > 0 \,:\, \forall k \geqslant 1 \;\forall N \geqslant 1 \quad \Big(\Psi_N \,,\, \Big(\frac{H_N}{N} \Big)^k \Psi_N \Big) \leqslant C^k \,. \tag{5.30}$$

If $\gamma^{(k)}_{N,t}$ are the k-th marginals of $\Psi_{N,t} := e^{-\frac{i}{\hbar} H_N t} \Psi_N$ (the time-evolution of Ψ_N under the Hamiltonian H_N), then $\forall\, k \geqslant 1$

$$\gamma^{(k)}_{N,\cdot} \to |\varphi_\cdot\rangle\langle\varphi_\cdot|^{\otimes k} \quad \text{as } N \to \infty \tag{5.31}$$

weakly-$*$ in $L^\infty(\mathbb{R}, \mathcal{L}^2_{\rm reg})$ – defined in (5.39) below – where $\varphi_t \in H^1(\mathbb{R}^3)$ is the solution of the TIME-DEPENDENT CUBIC NONLINEAR SCHRÖDINGER EQUATION (CNSE) with initial datum φ :

$$\left\{ \begin{array}{rcl} i\hbar\,\partial_t\varphi_t & = & -\dfrac{\hbar^2}{2m}\dfrac{\partial^2}{\partial x^2}\varphi_t + b\,|\varphi_t|^2\varphi_t \\ \\ \varphi_t\big|_{t=0} & \equiv & \varphi\,. \end{array} \right. \tag{5.32}$$

So, AGT theorem provides an analogous persistence in time of the asymptotic factorisation of marginals $\gamma^{(k)}_{N,t}$.

Here, by scaling, $NV_N(x) \xrightarrow{\mathcal{S}'} b\,\delta(x)$ as $N \to \infty$: V_N is an approximate delta function at the scale $N^{-\beta}$. Adami et al. can control that factorised initial states $\varphi^{\otimes N}$ satisfying the energy bounds (5.30) exist, provided that $0 < \beta < \frac{1}{2}$: for that it suffices that φ has compactly supported Fourier transform. In the whole range $0 < \beta < 1$ one can deal with factorised initial states even without (5.30), through the approximation strategy set up by Erdős et al. in the three dimensional problem [42] (details in the appendix of [3]).

Both equations (5.27) and (5.32) go under the name of CUBIC NONLINEAR SCHRÖDINGER EQUATION (CNSE). Its solution has the physical meaning of condensate wave function. In dimension d, each term of the CNSE has dimension

$$(\text{mass})\ (\text{length})^{2-\frac{d}{2}}(\text{time})^{-2}.$$

By solution of (5.27) or (5.32) here one means a *strong H^1_0-solution*, that is, a function

$$\varphi \in C(I, H^1_0(\mathbb{R}^d)) \cap C^1(I, H^1_0(\mathbb{R}^d)) \tag{5.33}$$

such that φ satisfies the equation in $H^{-1}(\mathbb{R}^d)$ for all $t \in I$, interval in $\mathbb{R}$ with $I \ni 0$, and prescribed initial value $\varphi_{t=0} \in H^1(\mathbb{R}^d)$. Notoriously (see [27] for details),

- ✓ the solution exists and is unique,
- ✓ it is *global*, i.e., $I = \mathbb{R}$,
- ✓ its L^2-norm is conserved at any time, i.e., $\|\varphi_t(\,\cdot\,)\|_{\mathbb{R}^d} = \|\varphi_{t=0}(\,\cdot\,)\|_{\mathbb{R}^d}$,
- ✓ its H^1-norm is uniformly bounded in time.

Concerning the topologies for *convergence* in both theorems, let us recall from Sec. 2.1 that the Banach structure of the trace class $\mathcal{L}^1(L^2(\mathbb{R}^{3k}))$ and the Hilbert structure of the Hilbert-Schmidt class $\mathcal{L}^2(L^2(\mathbb{R}^{3k}))$ are given, respectively, by

$$\|A\|_{\mathcal{L}^1} = \mathrm{Tr}_{\mathcal{H}_k}|A|\,, \tag{5.34}$$

$$(B, C)_{\mathcal{L}^2} = \mathrm{Tr}_{\mathcal{H}_k}[B^*C]\,, \tag{5.35}$$

$\forall A \in \mathcal{L}^1(\mathcal{H}_k)$, $\forall B, C \in \mathcal{L}^2(\mathcal{H}_k)$. Their weak-$*$ topology, namely the topology based on the preduals, is determined by the dualities

$$\begin{array}{c} \big(\,\mathcal{L}^1(\mathcal{H}_k)\,,\,\|\ \|_{\mathcal{L}^1}\,\big) = \big(\,\mathrm{Com}(\mathcal{H}_k)\,,\,\|\ \|\,\big)^* \\ \\ \mathrm{Com}(\mathcal{H}_k) \ni C \overset{B\in\mathcal{L}^1}{\longmapsto} \mathrm{Tr}_{\mathcal{H}_k}[BC] \end{array} \tag{5.36}$$

and

$$\begin{aligned} \big(\,\mathcal{L}^2(\mathcal{H}_k)\,,\,\|\;\|_{\mathcal{L}^2}\,\big) &= \big(\,\mathcal{L}^2(\mathcal{H}_k)\,,\,\|\;\|_{\mathcal{L}^2}\,\big)^* \\ \mathcal{L}^2(\mathcal{H}_k) \ni C &\overset{B\in\mathcal{L}^2}{\longmapsto} \mathrm{Tr}_{\mathcal{H}_k}[B^*C] \end{aligned} \tag{5.37}$$

respectively, $\mathrm{Com}(\mathcal{H}_k)$ being the space of compacts on $\mathcal{H}_k$.

In particular, convergence (5.26) in ESY theorem reads

$$\lim_{N\to\infty} \mathrm{Tr}_{\mathcal{H}_k}\big[C(\gamma_{N,t}^{(k)} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k})\big] = 0 \tag{5.38}$$

for any positive integer k, any time t and and any compact C. Convergence in AGT theorem involves a Sobolev space of Hilbert Schmidt operators. To denote that an extra regularity is asked to the density matrices under consideration, one introduces the space

$$\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}^k)) := \big\{\gamma \in \mathcal{L}^2(L^2(\mathbb{R}^k)) \;:\; S_1^2\cdots S_k^2\,\gamma \in \mathcal{L}^2(L^2(\mathbb{R}^k))\big\} \tag{5.39}$$

which is a Hilbert space when endowed with the scalar product

$$\begin{aligned} (\rho,\gamma)_{\mathcal{L}^2_{\mathrm{reg}}} &:= (\,S_1^2\cdots S_k^2\,\rho\,,\,S_1^2\cdots S_k^2\,\gamma\,)_{\mathcal{L}^2} \\ &= \mathrm{Tr}_{\mathcal{H}_k}\big[\,(\,S_1^2\cdots S_k^2\,\rho\,)^*(S_1^2\cdots S_k^2\,\gamma\,)\,\big] \\ &= \mathrm{Tr}_{\mathcal{H}_k}\big[\,(\,S_1\cdots S_k\,\rho\,S_1\cdots S_k\,)^*(\,S_1\cdots S_k\,\gamma\,S_1\cdots S_k\,)\,\big] \end{aligned} \tag{5.40}$$

where the S_i's $(i = 1,\dots k)$ are the positive operators

$$S_i := (1-\partial_{x_i}^2)^{1/2}\,. \tag{5.41}$$

Weak-$*$ topology in $\mathcal{L}^2_{\mathrm{reg}}$ is defined analogously to the $\mathcal{L}^2$ case (5.37). This, in turn, defines the weak-$*$ topology in $L^\infty(\mathbb{R},\mathcal{L}^2_{\mathrm{reg}})$ the theorem deals with: in fact, the $\gamma_{N,t}^{(k)}$'s and $|\varphi_t\rangle\langle\varphi_t|^{\otimes k}$ are shown to have $\mathcal{L}^2_{\mathrm{reg}}$-norms that are uniformly bounded both in N and in time. Thus, convergence (5.31) in AGT theorem reads

$$\begin{gathered} \lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t\,\big(\,\rho_t\,,\,\gamma_{N,t}^{(k)} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k}\,\big)_{\mathcal{L}^2_{\mathrm{reg}}} = 0 \\ \forall\,\rho_\cdot \in L^1(\mathbb{R},\mathcal{L}^2_{\mathrm{reg}}) \end{gathered} \tag{5.42}$$

$\forall k \geqslant 1$.

5.4 Scheme of derivation: method of the hierarchies. Effective nonlinear one-particle equations from the linear many-body dynamics.

We review in this section the scheme of the rigorous derivation of the GP equation in connection with the problem of the time-stability of condensation.

Actually, as we have already mentioned, from the mathematical point of view this problem covers an even larger range of models than the aforementioned B.E.C.-related origin. In general it arises, as a part of Kinetic Equations theory, whenever one wants to extract, under suitable conditions and in some scaling limit $N \to \infty$, the effective *nonlinear* one-particle equation emerging from a *linear* many-particle evolutionary equation.

The picture can be both classical and quantum. In the first case, the many-body dynamics one starts with is given by the Liouville equation (or Boltzmann, or Vlasov equation, depending on the context) and the one-body effective equation one derives is usually referred to as the Poisson-Vlasov equation. In the latter, one starts from the many-body Schrödinger equation to end up with a one-body nonlinear Schrödinger equation which takes the name, depending on the kind of nonlinearity, of Hartree, Schrödinger-Poisson, Gross-Pitaevskiĭ, ..., equation. There is also the related issue of deriving the classical results, both at the N-body and at the single-particle level, through an appropriate semiclassical limit of the corresponding quantum equations.

Along this mainstream, classical and quantum perspective historically alternated in the last three decades and specific tools and approaches have been developed, within *an essentially common strategy*: rewrite the N-body equation in terms of a hierarchy of evolutionary equations for marginals and then identify the limiting k-th marginals, at each fixed k as $N \to \infty$, as the solution of the corresponding infinite hierarchy.

This idea has been introduced in 1977 by Braun and Hepp [20] in the classical framework and in 1980 by Spohn [103] for the quantum case (though this is not the only difference among them, as we are discussing later). A historic survey of how it has been developed and improved since then is postponed to the next section. To present it concretely, according to the spirit of the present work, we choose the quantum perspective, with the notation already introduced so far, of an initially condensed Bose gas for which persistence of condensation is investigated. References to similar investigations will be mentioned throughout.

The **scheme of the analysis of the time-stability of B.E.C.** goes as follows. A N-body Bose system trapped in a confining potentials is given in a initial condensed state γ_N, (like in the model of theorem 3.5.2), possibly with some additional conditions (as in the hypotheses of theorems 5.3.1 and 5.3.2) expressing essentially a finite energy per particle. At $t = 0$ the trap is instantaneously removed and the system evolves under a two-body interaction Hamiltonian

$$H_N = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \Delta_{x_i} + \sum_{1 \leqslant i < j \leqslant N} V_N(x_i - x_j)\,. \tag{5.43}$$

As usual, the pairing, apart from its regularity properties, is assumed to be spherically symmetric and repulsive. The initial state is not invariant under such H_N, then it evolves according to the Heisenberg - von Neumann equation

$$i\hbar\,\partial_t \gamma_{N,t} = [\,H_N\,,\,\gamma_{N,t}\,] \tag{5.44}$$

and at each time the corresponding marginals $\gamma_{N,t}^{(k)}$ are defined.

The goal is to control whether condensation occurs at $t > 0$, that is, whether in some approximation $\gamma_{N,t}^{(1)} \approx |\varphi_t\rangle\langle\varphi_t|$ for N large enough. In particular, some scaling has to be specified as $N \to \infty$, in the spirit of the discussion of Chap. 4 (and Sec. 3.2). One is then interested, for any fixed k, in the limiting objects

$$\gamma_{\infty,t}^{(k)} := \lim_{N \to \infty} \gamma_{N,t}^{(k)}\,, \tag{5.45}$$

in which sense they exist, and when they possibly factorise as a tensor product of projections onto some φ_t that plays the role of time-dependent condensate wave function. Furthermore, φ_t is expected to be the solution of that cubic nonlinear Schrödinger equation

(the Gross-Pitaevskiĭ equation) confirmed by physical heuristic arguments and experimental observations to be the evolutionary equation of the condensate wave function, with initial condition given by the condensate wave function φ at time $t = 0$. So one seeks to complete the following diagram for any positive integer k:

$$\begin{array}{ccc} \gamma_N^{(k)} & \xrightarrow{N\to\infty} & |\varphi\rangle\langle\varphi|^{\otimes k} \\ \text{BBGKY}-\text{hierarchy}\Big\downarrow & & \Big\downarrow \text{GP}-\text{equation} \\ \gamma_{N,t}^{(k)} & \xrightarrow{N\to\infty} & |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \end{array} \tag{5.46}$$

Following the seminal idea of Braun, Hepp and Spohn, the key tool to prove (5.46) is represented by *hierarchies of reduced density matrices*. In fact, through the definition of reduced density matrix (definition 2.3.3 and theorem 2.3.4), (5.44) is equivalent to a finite hierarchy of first order linear PDE's for kernels, where each $\partial_t \gamma_{N,t}^{(k)}$ is expressed in terms of $\gamma_{N,t}^{(k)}$ and $\gamma_{N,t}^{(k+1)}$, the last equation of which is (5.44) itself. A rather established although not unanimous use in the literature is to refer to it as the FINITE BBGKY HIERARCHY after the works of Bogolubov [16], Born and Green [17], Kirkwood [65, 66], and Yvon [112] – in fact, it resembles the BBGKY hierarchy of equations satisfied by the k-particle distributions (i.e., probability densities) in the classical Kinetic Theory, as a consequence of the Liouville equation (see [28, 29]). Sometimes it is also cited as the finite Schrödinger hierarchy, to emphasize its quantum character, while its classical analog is also called the Liouville hierarchy. Its explicit form in terms of kernels is

$$\begin{aligned} i\hbar\,\partial_t(X_k, Y_k) = {} & \frac{\hbar^2}{2m}\sum_{i=1}^{k}\big(-\Delta_{x_i}+\Delta_{y_i}\big)\,\gamma_{N,t}^{(k)}(X_k, Y_k) \\ & + \sum_{1\leqslant i<j\leqslant k}\big(V_N(x_i-x_j)-V_N(y_i-y_j)\big)\,\gamma_{N,t}^{(k)}(X_k, Y_k) \\ & + (N-k)\sum_{i=1}^{k}\int_{\mathbb{R}^d}\mathrm{d}z\,\big(V_N(x_i-z)-V_N(y_i-z)\big)\,\gamma_{N,t}^{(k+1)}(X_k, z, Y_k, z) \\ & \\ & k = 1, \dots, N \end{aligned} \tag{5.47}$$

where d is the space dimension, each $x_i \in \mathbb{R}^d$ and $X_k := (x_1, \dots, x_k) \in \mathbb{R}^{kd}$, etc., and the convention is that $\gamma_{N,t}^{(k)} = \mathbb{O}$ if $k > N$. Hence, the N-th and last equation of the hierarchy is (5.44) itself, which in terms of kernels reads

$$\begin{aligned} i\hbar\,\partial_t\gamma_{N,t}(X_N, Y_N) = {} & \frac{\hbar^2}{2m}\sum_{i=1}^{N}\big(-\Delta_{x_i}+\Delta_{y_i}\big)\,\gamma_{N,t}(X_N, Y_N) \\ & + \sum_{1\leqslant i<j\leqslant N}\big(V_N(x_i-x_j)-V_N(y_i-y_j)\big)\,\gamma_{N,t}(X_N, Y_N)\,. \end{aligned} \tag{5.48}$$

Notice that the first term on the r.h.s. of the hierarchy (5.47) describes the kinetic energy of the first k particles, the second term is associated with the interactions among the first k particles, and the last term corresponds to interactions between the first k particles and the other $N - k$ particles.

Let us also quote other useful equivalent forms of (5.47). Its operator form is

$$\begin{aligned} i\hbar\,\partial_t\gamma_{N,t}^{(k)} = \frac{\hbar^2}{2m}\sum_{i=1}^{k}\Big[-\Delta_{x_i}\,,\,\gamma_{N,t}^{(k)}\Big] + \sum_{1\leqslant i<j\leqslant k}\Big[V_N(x_i-x_j)\,,\,\gamma_{N,t}^{(k)}\Big] \\ + (N-k)\sum_{i=1}^{k}\mathrm{Tr}_{[k+1]}\Big[V_N(x_i-x_{k+1})\,,\,\gamma_{N,t}^{(k+1)}\Big] \end{aligned} \tag{5.49}$$

where $\mathrm{Tr}_{[k+1]} : \mathcal{L}^1(L^2(\mathbb{R}^{(k+1)d})) \to \mathcal{L}^1(L^2(\mathbb{R}^{kd}))$ is the partial trace over the $(k+1)$-th variable. The corresponding integral forms are

$$\begin{aligned} \gamma_{N,t}^{(k)} = \gamma_{N,s}^{(k)} - i\sum_{i=1}^{k}\int_s^t \mathrm{d}r\,\Big[-\Delta_{x_i}\,,\,\gamma_{N,r}^{(k)}\Big] - i\sum_{1\leqslant i<j\leqslant k}\int_s^t \mathrm{d}r\,\Big[V_N(x_i-x_j)\,,\,\gamma_{N,r}^{(k)}\Big] \\ - i(N-k)\sum_{i=1}^{k}\int_s^t \mathrm{d}r\,\mathrm{Tr}_{[k+1]}\Big[V_N(x_i-x_{k+1})\,,\,\gamma_{N,r}^{(k+1)}\Big] \\ (s\leqslant t) \end{aligned} \tag{5.50}$$

and

$$\begin{aligned} \gamma_{N,t}^{(k)} = \mathcal{U}_N^{(k)}(t)\,\gamma_{N,0}^{(k)} + \\ - i(N-k)\sum_{i=1}^{k}\int_0^t \mathrm{d}r\,\mathcal{U}_N^{(k)}(t-r)\,\mathrm{Tr}_{[k+1]}\Big[V_N(x_i-x_{k+1})\,,\,\gamma_{N,r}^{(k+1)}\Big] \end{aligned} \tag{5.51}$$

where

$$\mathcal{U}_N^{(k)}(t)\,\gamma \;:=\; e^{-iH_N^{(k)}t/\hbar}\,\gamma\,e^{iH_N^{(k)}t/\hbar} \tag{5.52}$$

$$H_N^{(k)} \;:=\; -\frac{\hbar^2}{2m}\sum_{i=1}^{k}\Delta_{x_i} + \sum_{1\leqslant i<j\leqslant k}V_N(x_i-x_j)\,. \tag{5.53}$$

This way, the knowledge of a solution $\gamma_{N,t}$ of (5.44) with the initial condition γ_N is equivalent to the knowledge of the finite sequence $\{\gamma_{N,t}^{(k)}\}_{k=1}^{N}$ solving hierarchy (5.47) with initial condition $\{\gamma_N^{(k)}\}_{k=1}^{N}$. To control possible limiting marginals $\gamma_{\infty,t}^{(k)}$ in a scaling limit $N\to\infty$, a twofold analysis is performed. The description we give here is modelled on the concrete case of the recent investigation by Erdős, Schlein and Yau which has led to theorem 5.3.1, that is, a class of initial states with asymptotic 100% B.E.C. for which persistence of asymptotic 100% condensation at later time is proved in the GP scaling limit. Many other related attempts share an analogous scheme.

① **(Extraction of weak-$*$ limits)** – At each fixed k, the sequence $\{\gamma_{N,t}^{(k)}\}_{N\geqslant k}$ is proved to be weakly-$*$ compact in some suitable topology for trace class operators. Then, via a Banach-Alaoglu argument, it admits weak-$*$ limits for subsequences. What ensures such a sequence to stay uniformly inside a ball of a suitable Banach space with predual, is an "a priori estimate" following from assuming that the initial state has a N-uniformly bounded energy per particle, together with the same assumption for higher moments of the energy. Each (possibly not unique) limit point $\gamma_{\infty,t}^{(k)}$ is not a priori a density matrices, yet it is proved to be a positive and permutationally symmetric trace class operator, with "good" trace properties

for the subsequent steps (i.e., the weak-$*$ limit belongs to the same ball as the original sequence).

(Convergence to the infinite hierarchy) – Each infinite family $\{\gamma^{(k)}_{\infty,t}\}_{k\geqslant 1}$ of weak-$*$ limit points is proved to be the solution of an *infinite* hierarchy for trace class operators with initial condition $\{\,|\varphi\rangle\langle\varphi|^{\otimes k}\}_{k\geqslant 1}$. This can be obtained *formally* by taking $N\to\infty$ in the finite hierarchy (5.47). Such an infinite hierarchy maintains the structure of coupled P.D.E.'s where the k-th element of the family is given in terms of the $(k+1)$-th one.

② **(Uniqueness of the solution)** – One proves that the infinite hierarchy admits a *unique* solution $\{\gamma^{(k)}_t\}_{k\geqslant 1}$ with initial condition $\{\,|\varphi\rangle\langle\varphi|^{\otimes k}\}_{k\geqslant 1}$ and such that each $\gamma^{(k)}_t$ is a positive and permutationally symmetric trace class operator inside the same ball as $\gamma^{(k)}_{\infty,t}$ (namely, with the same trace properties). By direct inspection, one checks that $\{\,|\varphi_t\rangle\langle\varphi_t|^{\otimes k}\}_{k\geqslant 1}$ solves such infinite hierarchy with those properties, iff φ_t is a solution of the time-dependent Gross-Pitaevskiĭ equation with initial condition φ. Hence, by uniqueness, the weak-$*$ limit point is unique and $\gamma^{(k)}_{N,t}\to|\varphi_t\rangle\langle\varphi_t|^{\otimes k}$ as $N\to\infty$.

In the three-dimensional setting and under the (A) hypotheses of theorem 5.3.1, the infinite hierarchy satisfied by the limiting sequence $\{\gamma^{(k)}_{\infty,t}\}_{k\geqslant 1}$ reads

$$\begin{aligned} i\hbar\,\partial_t\gamma^{(k)}_t(X_k,Y_k) = &-\frac{\hbar^2}{2m}\sum_{\ell=1}^{k}\Big(-\Delta_{x_\ell}+\Delta_{y_\ell}\Big)\,\gamma^{(k)}_t(X_k,Y_k)\\ &+8\pi\mu a\sum_{\ell=1}^{k}\int_{\mathbb{R}^3}\mathrm{d}z\,\big(\delta(x_\ell-z)-\delta(y_\ell-z)\big)\,\gamma^{(k+1)}_t(X_k,z,Y_k,z)\end{aligned} \tag{5.54}$$

for $k=1,2,3,\dots$, or equivalently, in integral operator form,

$$\gamma^{(k)}_t=\mathcal{U}^{(k)}(t)\,\gamma^{(k)}_0-8i\pi\mu a\sum_{\ell=1}^{k}\int_0^t\mathrm{d}s\,\mathcal{U}^{(k)}(t-s)\,\mathrm{Tr}_{[k+1]}\Big[\,\delta(x_\ell-x_{k+1})\,,\,\gamma^{(k+1)}_s\Big]\,, \tag{5.55}$$

where a_1 is the s-wave scattering length of the unscaled ("true") potential V and

$$\mathcal{U}^{(k)}(t)\,\gamma^{(k)}:=e^{\,i\frac{\mu}{\hbar}t\sum_{\ell=1}^{k}\Delta_{x_\ell}}\,\gamma^{(k)}\,e^{\,-i\frac{\mu}{\hbar}t\sum_{\ell=1}^{k}\Delta_{x_\ell}}\,. \tag{5.56}$$

Here the action of the delta function on (kernels of) density matrices is well-defined (it would not be for general density matrices) through an appropriate limiting procedure (details in Sec. 8 of [42] and Sec. 7 of [44]), which is possible since the limiting objects $\gamma^{(k)}_{\infty,t}$ are proved to have sufficiently regular kernels. This gives

$$\begin{aligned}\Big(\mathrm{Tr}_{[k+1]}\big[\,\delta(x_\ell-x_{k+1})\,\gamma^{(k+1)}_s\big]\,\Big)(X_k,Y_k)&=\int_{\mathbb{R}^3}\mathrm{d}z\,\delta(x_\ell-z)\,\gamma^{(k+1)}_t(X_k,z,Y_k,z)\\&=\gamma^{(k+1)}_s(X_k,x_\ell,Y_k,x_\ell)\end{aligned} \tag{5.57}$$

and (5.54) reads

$$
\begin{aligned}
i\hbar\,\partial_t\gamma_t^{(k)}(X_k,Y_k) = &-\frac{\hbar^2}{2m}\sum_{\ell=1}^{k}\Big(-\Delta_{x_\ell}+\Delta_{y_\ell}\Big)\,\gamma_t^{(k)}(X_k,Y_k)\\
&+8\pi\mu a\sum_{\ell=1}^{k}\Big[\,\gamma_t^{(k+1)}(X_k,x_\ell,Y_k,x_\ell)-\gamma_t^{(k+1)}(X_k,y_\ell,Y_k,y_\ell)\,\Big]\,.
\end{aligned}
\tag{5.58}
$$

It is then straightforward to check that $\{\gamma_t^{(k)}=|\varphi_t\rangle\langle\varphi_t|^{\otimes k}\}_{k\geqslant 1}$ is a solution of (5.58) with initial condition $\{|\varphi\rangle\langle\varphi|^{\otimes k}\}_{k\geqslant 1}$ *iff* $\varphi_\cdot$ is a solution of the problem (5.27), that is, of the GP equation with initial condition φ.

Hierarchy (5.54) is usually called the INFINITE BBGKY HIERARCHY, or sometimes also infinite Schrödinger hierarchy to emphasize its quantum character, while its classical counterpart goes usually under the name of infinite Vlasov hierarchy. Due to the physical relevance of the coupling in front of the interaction term, (5.54) is also called INFINITE GP HIERARCHY whenever this coupling is $8\pi\mu a$, to distinguish it from the a generic coupling α. Such α fixes the magnitude of the nonlinearity in the corresponding nonlinear Schrödinger equation for φ_t, which correspondingly has the form

$$
i\hbar\,\partial_t\varphi_t = -\mu\Delta_x\varphi_t+\alpha|\varphi_t|^2\varphi_t\,.
$$

Only for $\alpha=8\pi\mu a$ one gets the physically relevant GP equation.

The formal limit by which one obtains the infinite hierarchy (5.54) from the finite one (5.47) is itself a major conceptual point. Recalling that in the three-dimensional GP scaling $NV_N \xrightarrow{\mathcal{S}'} b\,\delta$ (with $b=\|V\|_{L^1}$), whence $V_N\sim\frac{1}{N}\,\delta$, it is clear heuristically that the second term in the r.h.s. of (5.47) disappears in the limit $N\to\infty$, since it contains $k(k-1)/2$ terms of the form $\sim\frac{1}{N}\,\delta(x_i-x_j)$, while the third term in the r.h.s. of (5.47) consists of k terms with a pre-factor $(1-k/N)$ and hence approaches asymptotically the second term in the r.h.s. of (5.54). (The same can be argued for different kinds of scalings, like mean field or modified GP.) Yet, this does not provide an even heuristic justification of the the coupling $8\pi\mu a$ instead of the emerging b.

Since to close the infinite hierarchy one would need some relation between $\gamma_t^{(2)}$ and $\gamma_t^{(1)}$, any relation between $\gamma_{N,t}^{(2)}$ and $\gamma_{N,t}^{(1)}$ (*before* the limit) provides in principle a distinct way to take the formal limit in the finite hierarchy. As we will be discussing in Chap. 6, any other relation than the simple factorisation $\gamma_{N,t}^{(2)}=\gamma_{N,t}^{(1)}\otimes\gamma_{N,t}^{(1)}$ corresponds to certain non-trivial *correlations* in the many-body state. Thus, the knowledge of correlations present in the many-body state along its time evolution would allow one to perform a distinct formal limit. Since having such a knowledge would, in turn, require to have already solved the problem, one appeals to some ansatz on the basis of physical intuition.

For some years, attempts have been made without introducing short scale correlations in the marginals: this ansatz leads to an infinite hierarchy with an unexpectedly larger coupling constant than the GP one. Only when Erdős, Schlein, and Yau realized that suitable ground-state-like short scale correlations *persisting in time* had to be plugged into $\gamma_{N,t}^{(k)}$, one finally could perform a formal limit $N\to\infty$ leading to the correct GP hierarchy. We give details of this at the end of Sec. 6.4, in the framework of correlations at the level of reduced density matrices (see also the most recent discussion in Sec. 3 of [44], after the first observations in this direction expressed in the introductory sections of [41] and [40]).

Physically, pairwise correlations are established in the two-body scattering process, provided that a particle has enough time to feel the potential field of the other. The GP scaling is indeed *the* scaling where the strength of the interaction (namely the scattering length) is comparable with its spatial range (both scale as N^{-1}) and, hence, it allows to study emergence and persistence in time of short scale correlations. This is not the case for other scalings like the mean field or the modified GP (4.15), where the range of the interaction is respectively much larger than the scattering length or of the same order than the spatial density of the gas. For those scalings one does not need further correlation ansatz to derive formally the infinite hierarchy from the finite one.

5.5 Historic survey

We have already mentioned that the method of the hierarchies, both in classical and in quantum Kinetic Theory, and the twofold task of "convergence + uniqueness" have a 30 years lasting history.

In 1977 Braun and Hepp [20] derived the Vlasov equation from a system of N classical particles coupled by a mean field interaction $\frac{1}{N}\sum_{i<j} V(x_i - x_j)$. They assumed the interacting potential bounded in C^2 and proved the convergence of the classical Liouville hierarchy to the infinite Vlasov hierarchy. They avoided addressing the question of the uniqueness of solutions of the infinite hierarchy, but nevertheless succeeded to prove that the limits of N-particle marginals (in the large N limit) coincide with factorised products of solution of the Vlasov equation. Their argument rely on a clever use of probabilistic methods (such as the strong law of large numbers to establish the propagation of chaos and the central limit theorem to analyze fluctuations about the mean field). Such an argument does not seem to work for the limit $\hbar \to 0$, $N \to \infty$ of a system of quantum particles. This might explain why Narnhofer and Sewell [90] had to prove a uniqueness theorem for the Vlasov equation. Since the corresponding operator is very singular, stringent assumptions (equivalent to the analyticity of the potential) are required in that paper. The regularity assumption was substantially relaxed by Spohn [104].

In 1980 Spohn [103] came up with the idea of using the trace norm to prove the convergence of the solution of the N-body Schrödinger equation to the factorised N-body state driven by the non-linear Schrödinger equation. His proof rests on the assumption of a bounded interaction potential. The scaling in the interaction is just a mean field limit. Being based on the Duhamel formula, his proof avoids the question of uniqueness for the infinite Schrodinger hierarchy, although in a very different way than that in Braun and Hepp. Along this quantum side of the problem, Spohn's result is remarkable not only because of the introduction of the method of the hierarchies for density matrices, but also because it generalizes a previous (1974) result by Hepp [57], obtained in the context of field operators, where the pair potential needed to be differentiable, instead of only bounded as in Spohn.

Incidentally, Ginibre and Velo [51, 52] greatly extended in the meanwhile these results to cover singular potentials (including Coulomb), but in a completely different approach involving quasi-free states in the second quantized framework: this method requires special initial states with no definite particle number.

The next key contribution was due to Bardos, Golse and Mauser [10] who in 2000 split the strategy into the twofold "convergence + uniqueness" task for the mean field treatment of a gas of spinless undistinguishable bosons expanding under a repulsive pair

potential and prepared in a purely factorised initial state. The outcoming one-body equation is the Hartree equation (also known as the Schrödinger-Poisson equation when the interaction V is a repulsive Coulomb potential), that is, a nonlinear Schrödinger equation with a non-local nonlinearity $(V * |\varphi_t|^2)\varphi_t$, instead of the local nonlinearity $|\varphi_t|^2\varphi_t$ of the CNSE. Convergence of the N-body Schrödinger hierarchy to the infinite Schrödinger hierarchy was shown under fairly general assumptions – a bounded-below potential in $L^2 + L^\infty$, thus including repulsive Coulomb systems, which were not covered by Spohn. Uniqueness, instead, needed stringent assumptions (boundedness and vanishing at infinity) and was based on the abstract Cauchy-Kowalewski theorem by Nirenberg [91] and Nisida [92] (a slight modification of such argument was subsequently obtained by the same authors together with Erdős and Yau in [9]).

Technically, all uniqueness methods known until then relied on the boundedness of the potential via the estimate $\mathrm{Tr}|V\gamma| \leqslant \|V\|_\infty \mathrm{Tr}|\gamma|$. To include the Coulomb repulsion a Hardy-like inequality $\mathrm{Tr}|V\gamma| \leqslant C(\mathrm{Tr}[\nabla\gamma\nabla] + \mathrm{Tr}|\gamma|)$ is needed (which is indeed a consequence of the Hardy inequality $\frac{1}{4}|x|^{-2} \leqslant -\Delta$ in three dimensions). It was Erdős and Yau who first noticed it in 2001 the importance of such an estimate and realized that the solution to the infinite BBGKY hierarchy should be unique under the right Sobolev norm: this led to the full derivation (convergence + uniqueness) of the nonlinear Schrödinger (Hartree) equation from a many-body Coulomb system [45]. The scaling was a mean field and the initial datum was a completely factorised $\varphi^{\otimes N}$, but with an unnatural $\varphi \in H^2$, instead of the natural H^1 space for the nonlinear Schrödinger equation. Convergence was controlled in a weak-$*$ sense for trace class valued bounded functions.

In fact, it is noteworthy noticing that in this scenario the choice of the topology to control the limiting factorisation of the time-dependent marginals is forced by the technique of such analysis. Depending on the topology by which one is able to control *uniqueness*, one has to adopt the topology by which controlling the *existence* of the limit. Explicitly, if to prove uniqueness one has to assume that the solution is bounded in some norm, then the limiting $\gamma^{(1)}_{\infty,t}$ has to be bounded in that norm, so one has to introduce a topology for the convergence which allows to control that kind of boundedness in the limit. It may happen that to prove uniqueness one needs to regard the kernel of $\gamma^{(1)}_{\infty,t}$ as a more regular object than simply a L^2-function and some suitable Sobolev spaces enter the analysis. Thus, one ends up with some regularized weak-$*$ topology.

All treatments until this point had been dealing with a mean field scaling, both in the above quantum side and in the corresponding classical counterpart we are not concerned in here. A complete review of the state of the art (until 2003) of the mean field limit for the dynamics of large classical and quantum particle systems is Ref. [54] by Golse. An analogous review for the quantum side only is Ref. [11] by Bardos and Mauser.

It is reasonable to say that the interest in developing the same conceptual scheme in some GP-like scaling (instead of all previous mean field treatments), which started only at this point, was driven by the increasing knowledge on B.E.C. of that time, both at the experimental level and on the mathematical side. In fact, contributions in the B.E.C.-driven directions, that is, towards the rigorous derivation of the CNSE, followed soon after. The crucial intermediate steps to obtain the somewhat conclusive results of theorems 5.3.1 and 5.3.2 have been (*i*) the identification of the appropriate infinite hierarchy as $N \to \infty$ in a suitable scaling limit (see the discussion concluding the previous section) and (*ii*) a joint control of convergence + uniqueness.

In 2003 Adami, Bardos, Golse, and Teta [2] (see also [1]) addressed the problem

for a one-dimensional system of N spinless undistinguishable bosons interacting through a mean field, repulsive, pairwise delta potential $\frac{1}{N}\sum_{i<j}\delta(x_i - x_j)$, initially prepared in a condensed state with N-uniformly bounded energy per particle, and proved the convergence of (solutions of) the finite hierarchy to (solutions of) the infinite one in a weakly-$*$ regularized (Sobolev) trace class sense, where limit points solve the infinite hierarchy in a distributional sense. A trivial solution of the infinite hierarchy is the sequence of the factorised projections onto the solution of the CNSE, but uniqueness for such a model is still missing today,

Concerning the three-dimensional setting, starting from 2004 a comprehension spread over the obstructions and the technical solutions to obtain the true GP hierarchy, that is, the infinite BBGKY hierarchy with the right coupling $(8\pi\mu a)$ ensuring the emergence of the GP equation (say the discussion at the end of the previous section).

In a first work [41] Erdős, Schlein, and Yau have proposed a model of spinless undistinguishable bosons interacting through a sufficiently small potential in the $\| \, \|_{L^1} + \| \, \|_{L^\infty}$ sense, scaling with N in the GP sense. By considering a modified Hamiltonian $\widetilde{H}_N$ which cuts off the pair interactions whenever at least three particles come into a region with diameter much smaller than the typical interparticle distance, and a class of initial states with a N-uniform bound both in the energy per particle $\widetilde{H}_N/N$, and in $(\widetilde{H}_N/N)^2$, they could prove convergence of (solutions of) the finite hierarchy to (solutions of) the infinite GP hierarchy in a weighted weakly-$*$ Hilbert-Schmidt sense, where limit points solve the infinite hierarchy in a distributional sense. The above assumption on the energy selects admissible initial states which are not necessarily condensates. Yet *for the first time two explicit classes of initial data have been recognised to satisfy the uniform bounds in the energy*, which can be both proved to have asymptotic 100% condensation (although this is not explicitly noticed in the original work): purely factorised $\varphi^{\otimes N}$ and wave functions $\prod_{i<j} f_N(x_i - x_j)\varphi(x_1)\cdots\varphi(x_N)$ (with enough regularity for φ), where correlations are given by the zero energy scattering solution of the two-body process f_N. From the mathematical viewpoint, the modified Hamiltonian removes a technical obstacle: squaring the true H_N instead of $\widetilde{H}_N$, the outcoming derivatives of the pair potential could not be controlled by the kinetic energy operator in the rare situation when many particles come close together, since such derivatives have no definite sign.

In a second work [40] Elgart, Erdős, Schlein, and Yau have taken into account the full Hamiltonian (assuming a non-negative, smooth, compactly supported interaction), but scaling it with a modification of the true GP scaling (see (4.15) in Sec. 4.3 and the discussion thereafter) such that the interaction potential is a delta function at a much larger scale than the scattering length. Consequently, the two-body process does not establish correlations at the scale of the scattering length and formal limit leads to an infinite BBGKY hierarchy with the first Born approximation coefficient in the interaction term. For any initial state with a N-uniform bound on every k-th moment of the energy per particle (i.e., $\mathrm{Tr}[(H_N/N)^k\gamma_N] \leqslant C^k$), convergence to such hierarchy is established in the same sense as in [41].

In both these 2004 contributions uniqueness is missing. Only in 2005 the first joint control of convergence and uniqueness in a GP-like scaling has been achieved. Since in the GP-like scaling the interaction is asymptotically $\sim \frac{1}{N}\,\delta$, in two and three dimensions this is such a singular potential that one cannot apply any longer the joint control developed by Bardos et al. for the mean field scaling. In one dimension, instead, this is still possible. Thus, by combining convergence technique of [40] and uniqueness technique

of [9], Adami, Golse and Teta [3] have completed the project of the rigorous derivation of the CNSE in dimension one from the dynamics of pairwise coupled bosons (see theorem 5.3.2) where the interaction scales with a much larger range than the scattering length (in analogy with the three-dimensional setting in [40]) and for any initial state with N-uniform bounds on every energy moment – which includes, in particular, suitably regular factorised state $\varphi^{\otimes N}$.

At the same time, Erdős, Schlein, and Yau [42] have completed the analogous three-dimensional project (with a proof that can be extended to lower dimensions). Interaction is scaled again as $V_N(x) = N^{3\beta-1}V(N^\beta x)$ for a non-negative, smooth, compactly supported and spherically symmetric V (but limited to the regime $0 < \beta < \frac{1}{2}$, unlike the regime $0 < \beta < \frac{3}{5}$ of the previous attempt [40]). The admitted initial states are of the form $\varphi^{\otimes N}$. Convergence holds weakly-$*$ in the trace class to an infinite hierarchy where the coupling constant is the first Born approximation of the scattering length. Uniqueness is achieved via a diagrammatic proof to control the Duhamel-type expansion of the integral form (5.55) of the hierarchy. The whole proof deals with a regularization of the initial data where the high energy component is cut off: this allows an appropriate a priori estimate for the regularized initial states guaranteeing the convergence + uniqueness control, that holds true even after the removal of the cutoff.

Along this line, 2006 theorem 5.3.1 by Erdős, Schlein, and Yau [44] combines the uniqueness technique of [42] with a new control of the convergence in the true three-dimensional GP scaling $V_N(x) = N^2V(Nx)$. In the (A) hypotheses, admitted initial states have asymptotic 100% B.E.C. and N-uniformly bounded energy moments. This leads to the control of convergence in the weakly-$*$ trace class sense and to the boundedness of a suitable Sobolev trace norm for each limiting point $\gamma^{(k)}_{\infty,t}$, which is exactly that norm under which the uniqueness theorem of [42] holds. For finite N, marginals $\gamma^{(k)}_{N,t}$ are recognised to contain a short scale structure emerging in the two-body process: the many-body dynamics *builds up* and *preserves* short-scale correlations (see Sec. 6.5).

5.6 Strengthened convergence in ESY theorem

In this section it will be recalled how the trace class weak-$*$ convergence in ESY theorem actually implies convergence in the trace norm. This is due to the same mechanism used in the proof of B.E.C. for dilute trapped gas by Lieb and Seiringer [74]. We have already reported it in theorem 3.4.2 and the result is the following.

Theorem 5.6.1 (**ESY strengthened convergence**)**.** Limit (5.38) in ESY theorem implies, and hence can be lifted to,

$$\lim_{N\to\infty} \left\| \gamma^{(k)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \right\|_{\mathcal{L}^1} = 0 \tag{5.59}$$

for any positive integer k and any time $t \in \mathbb{R}$.

As a matter of fact, the most recent formulation of ESY theorem already mentions trace norm (see theorem 1 in [43]), while its original version (see theorems 2.1 and 2.2 in [44]), as well as the preliminary results along this mainstream (see theorem 2.1 in [41], theorem 1 in [40], theorem 1 in [42]), all accounts for the weak-$*$ topology only. One has to recognise that the real power of ESY theorem is

① to show that weak-$*$ limits of sequence of marginals do exist via a Banach-Alaoglu weak-$*$ compactness argument; for this one needs that the k-th marginal is uniformly (in N) bounded in a suitable Banach norm,

② to prove that all these limits satisfy an *infinite* BBGKY hierarchy,

③ to prove that the hierarchy admits a unique solution which is explicitly seen to be tensor power of projections onto the solution of the GP equation.

In this perspective, stress in ESY theorem is not given to any stronger topology: the primitive way to think to limit of marginals as $N \to \infty$ (and to deal with them in computations) is to take *weak-$*$ limits*.

5.7 Strengthened convergence in the AGT theorem

In this section it will be shown how the $L^\infty(\mathbb{R}, \mathcal{L}^2_{\rm reg})$ weak-$*$ convergence in AGT theorem actually implies convergence in a suitable time-weighted trace norm.

Theorem 5.7.1 (**AGT strengthened convergence, [85]**). For any positive integer k, limit (5.42) in AGT theorem implies, and hence can be lifted to,

$$\lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t\, f(t) \left\| \gamma^{(k)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \right\|^2_{\mathcal{L}^2} = 0 \qquad \forall\, f \in L^1(\mathbb{R}) \tag{5.60}$$

and also to

$$\lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t\, f(t) \left\| \gamma^{(k)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \right\|_{\mathcal{L}^1} = 0 \qquad \forall\, f \in L^1(\mathbb{R})\,. \tag{5.61}$$

Proof. We give details for $k = 1$; generalization to $k > 1$ is elementary. Also, we consider only the case $f \geqslant 0$ and $\|f\|_{L^1(\mathbb{R})} = 1$: indeed, once (5.60) and (5.61) are proved under this restriction, the general case follows writing any non zero $f \in L^1(\mathbb{R})$ as the linear combination

$$f = \|f_+\|_{L^1(\mathbb{R})} \frac{f_+}{\|f_+\|_{L^1(\mathbb{R})}} - \|f_-\|_{L^1(\mathbb{R})} \frac{f_-}{\|f_-\|_{L^1(\mathbb{R})}} \tag{5.62}$$

where the positive $L^1(\mathbb{R})$-functions $f_\pm$ are the positive and negative part of f respectively, that is,

$$f_\pm(x) = \max\{\pm f(x), 0\}\,. \tag{5.63}$$

Step 1. One chooses a suitable $\rho_\cdot \in L^1(\mathbb{R}, \mathcal{L}^2_{\rm reg})$ in (5.42):

$$\rho_t := f(t) S^{-2} |\varphi_t\rangle\langle\varphi_t| S^{-2} \tag{5.64}$$

$\forall f \in L^1(\mathbb{R})$ such that $f \geqslant 0$ and $\|f\|_{L^1(\mathbb{R})} = 1$, where S is the operator

$$S := \left(1 - \frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)^{1/2}. \tag{5.65}$$

Indeed $S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2}$ is a positive trace class operator with uniformly (in time) bounded $\mathcal{L}^2_{\rm reg}$-norm:

$$\begin{aligned}\| S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2} \|^2_{\mathcal{L}^2_{\rm reg}} &= \| S^2 S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2} \|^2_{\mathcal{L}^2} \\ &= {\rm Tr}\big[\, |\varphi_t\rangle\langle\varphi_t|S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2}\,\big] \\ &= (\varphi_t, S^{-2}\varphi_t)^2_{L^2(\mathbb{R})} \\ &= \Big(\int_{-\infty}^{+\infty} \frac{1}{1+4\pi^2k^2}\,|\widehat{\varphi}_t(k)|^2 {\rm d}k\Big)^2 \\ &\leqslant \|\widehat{\varphi}_t\|^4_{L^2(\mathbb{R})} = \|\varphi_t\|^4_{L^2(\mathbb{R})} = 1\,.\end{aligned} \tag{5.66}$$

(Here the unimportant constants entering the computations are due to the convention

$$\widehat{\varphi}(k) = \int_{-\infty}^{+\infty} e^{-2\pi i k x}\varphi(x){\rm d}x$$

to define the Fourier transform.)

The $L^2(\mathbb{R})$-norm of φ_t is preserved in time and it amounts to 1 as at $t = 0$. Then $\rho_\cdot \in L^1(\mathbb{R}, \mathcal{L}^2_{\rm reg})$ because

$$\int_{-\infty}^{+\infty} {\rm d}t \,\| \,\rho_t\,\|_{\mathcal{L}^2_{\rm reg}} = \int_{-\infty}^{+\infty} {\rm d}t\, f(t) \| \,S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2}\,\|_{\mathcal{L}^2_{\rm reg}} \leqslant 1\,. \tag{5.67}$$

STEP 2. With this choice of $\rho_\cdot$, AGT convergence gives

$$\begin{aligned}&\int_{-\infty}^{+\infty} {\rm d}t\, f(t) \,\| \,\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|\,\|^2_{\mathcal{L}^2} \leqslant \\ &\quad\leqslant -\,2\int_{-\infty}^{+\infty} {\rm d}t\, f(t)\,{\rm Tr}_{\mathcal{H}_1}\big[\,|\varphi_t\rangle\langle\varphi_t|\,(\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|)\,\big] \\ &\quad= -\,2\int_{-\infty}^{+\infty} {\rm d}t\,{\rm Tr}_{\mathcal{H}_1}\big[\,S^2 f(t) S^{-2}|\varphi_t\rangle\langle\varphi_t|S^{-2}S^2(\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|)\,\big] \\ &\quad= -\,2\int_{-\infty}^{+\infty} {\rm d}t\,\big(\,\rho_t\,,\,\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|\,\big)_{\mathcal{L}^2_{\rm reg}}\end{aligned} \tag{5.68}$$

the inequality following by

$$\begin{aligned}&\| \,\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|\,\|^2_{\mathcal{L}^2} = \\ &\qquad= {\rm Tr}_{\mathcal{H}_1}\big[\,(\gamma^{(1)}_{N,t})^2 - \gamma^{(1)}_{N,t}|\varphi_t\rangle\langle\varphi_t| - |\varphi_t\rangle\langle\varphi_t|\gamma^{(1)}_{N,t} + (|\varphi_t\rangle\langle\varphi_t|)^2\,\big] \\ &\qquad\leqslant 2 - 2\,{\rm Tr}_{\mathcal{H}_1}\big[\,\gamma^{(1)}_{N,t}|\varphi_t\rangle\langle\varphi_t|\,\big] \\ &\qquad= -2\,{\rm Tr}_{\mathcal{H}_1}\big[\,|\varphi_t\rangle\langle\varphi_t|\,(\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|)\,\big]\,.\end{aligned} \tag{5.69}$$

Since, by assumption,

$$\begin{aligned}\lim_{N\to\infty}\int_{-\infty}^{+\infty} {\rm d}t\,\big(\,\rho_t\,,\,\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|\,\big)_{\mathcal{L}^2_{\rm reg}} = 0 \\ \forall\,\rho_\cdot \in L^1(\mathbb{R}, \mathcal{L}^2_{\rm reg})\,,\end{aligned} \tag{5.70}$$

then

$$\lim_{N\to\infty}\int_{-\infty}^{+\infty} {\rm d}t\, f(t) \,\| \,\gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|\,\|^2_{\mathcal{L}^2} = 0\,. \tag{5.71}$$

Step 3. Here one arrives to the trace class norm to prove (5.61). First one has

$$\begin{aligned}
&\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \|_{\mathcal{L}^1} = \\
&\quad = \| \, (\, |\varphi_t\rangle\langle\varphi_t| + (\, \mathbb{1} - |\varphi_t\rangle\langle\varphi_t| \,) \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} \\
&\quad \leqslant \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} + \| \, (\, \mathbb{1} - |\varphi_t\rangle\langle\varphi_t| \,) \, \gamma_{N,t}^{(1)} \, \|_{\mathcal{L}^1} \\
&\qquad\qquad + \| \, (\, \mathbb{1} - |\varphi_t\rangle\langle\varphi_t| \,) \, |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^1} \\
&\quad = \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} + 1 - \mathrm{Tr}\big[\, |\varphi_t\rangle\langle\varphi_t| \, \gamma_{N,t}^{(1)} \, \big] \\
&\quad = \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} - \mathrm{Tr}\big[\, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \big] \, ,
\end{aligned} \tag{5.72}$$

and, by (5.68),

$$\lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \mathrm{Tr}\big[\, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \big] = 0 \, . \tag{5.73}$$

At any time t and for all N, all the operators $|\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,)$ have rank one, so their trace class and their Hilbert-Schmidt norms coincide and

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} = \\
&\qquad\qquad = \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^2} \\
&\qquad\qquad \leqslant \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2} \cdot \| \, |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2} \\
&\qquad\qquad = \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2} \, .
\end{aligned} \tag{5.74}$$

Write the r.h.s. as a scalar product in $L^2(\mathbb{R})$

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \mathrm{d}t \, \sqrt{f(t)} \cdot \sqrt{f(t)} \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2} \\
&\qquad \leqslant \Big(\int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \Big)^{1/2} \Big(\int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2}^2 \Big)^{1/2} \\
&\qquad = \Big(\int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^2}^2 \Big)^{1/2}
\end{aligned} \tag{5.75}$$

so that, by (5.71),

$$\lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, |\varphi_t\rangle\langle\varphi_t| \, (\, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \,) \, \|_{\mathcal{L}^1} = 0 \, . \tag{5.76}$$

Then, plugging (5.73) and (5.76) into (5.72), one gets

$$\lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t \, f(t) \, \| \, \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \, \|_{\mathcal{L}^1} = 0 \tag{5.77}$$

which completes the proof. □

5.8 Unified strengthened convergence from ESY/AGT theorems

In this section a unified strong convergence will be shown of any k-th marginal to the k-th tensor power of the projection onto the solution of the CNSE. In other words, it is possible to obtain the same stronger convergence *both* if the starting point is the result of ESY theorem *and* if one starts from AGT theorem. Thus, despite the difference in the statement of such theorems (and in the techniques to derive them), they both admit the corollary that convergence can be lifted to trace-norm convergence at any time.

Since half of this program has been completed in Sec. 5.6, all that remains to do is to prove the following.

Theorem 5.8.1 (**AGT strengthened convergence at any time, [85]**)**.** For any positive integer k, limit (5.60) in the improved version of AGT theorem implies, and hence can be lifted to,

$$\lim_{N\to\infty} \left\| \gamma^{(k)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|^{\otimes k} \right\|_{\mathcal{L}^1} = 0 \tag{5.78}$$

for any time $t \in \mathbb{R}$.

Theorems 5.6.1 and 5.8.1 together provide a unification to the same strong convergence at any time, starting from two distinct weak convergences which have been obtained in ESY and AGT theorem, respectively, through a similar conceptual scheme but different techniques.

Let us focus on the origin of such differences. In ESY theorem, marginals $\gamma^{(k)}_{N,t}$'s are proved to have a bounded trace class norm uniformly in N and in t. By weak-$*$ compactness of the ball they all belong to, one extracts weak-$*$ limit points $\gamma^{(k)}_{\infty,t}$: each sequence $\{\gamma^{(k)}_{\infty,t}\}_{k\geqslant 1}$ of limit points is shown to solve the infinite BBGKY hierarchy with initial data $\gamma^{(k)}_{\infty,0} = |\varphi\rangle\langle\varphi|^{\otimes k}$. This is done *at any time* t in the space $\mathcal{L}^1(L^2(\mathbb{R}^{3k}))$. On the other hand, objects of interest in AGT theorem are not the operators $\gamma^{(k)}_{N,t}$'s at any fixed time, but instead the operator-valued functions $\gamma^{(k)}_{N,t} \in L^\infty(\mathbb{R}, \mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}^{3k})))$ and one proves that $\gamma^{(k)}_{N,\cdot} \to \gamma^{(k)}_{\infty,\cdot}$ weakly-$*$ in such a space. This choice, in turn, is forced by the way one controls that each limit point is the unique solution of the infinite hierarchy: this is not done at any t as in ESY theorem, but instead it is done in terms of functions $(t, X_k, Y_k) \mapsto \gamma_{\infty,t}(X_k, Y_k)$: they are proved to solve the infinite hierarchy (for kernels) in the sense of distributions $\mathcal{D}'(\mathbb{R}^{2k+1})$, that is, $2k$ space coordinates and one time coordinate.

Let us prepare some preliminaries to the proof of theorem 5.8.1. The notation

$$f_N(t) \leqslant \mathrm{const}(N, t) \tag{5.79}$$

will indicate that the sequence of functions $\{f_N\}_{N\geqslant N_0}$ is uniformly in $L^\infty(\mathbb{R})$, i.e., $f_N(t)$ is bounded by a constant independent of N and t, for all $t \in \mathbb{R}$ and for all N exceeding a fixed N_0. Our strategy will be to estimate time derivatives by regularity in space: for $\varphi_t(x)$ this is done by shifting from the l.h.s. to the r.h.s. of the CNSE, while for the kernel $\gamma^{(1)}_{N,t}(x, y)$ this is done by shifting from the l.h.s. to the r.h.s. of the P.D.E. it satisfies, namely the first equation of the BBGKY hierarchy for marginals. The regularity properties we will need are summarized as follows.

Theorem 5.8.2. Under the hypotheses of AGT theorem, the following holds.

① The first equation of the finite BBGKY hierarchy reads, in terms of kernels,

$$\begin{aligned} i\hbar\,\partial_t\gamma^{(1)}_{N,t}(x,y) &= \frac{\hbar^2}{2m}(-\partial_x^2+\partial_y^2)\,\gamma^{(1)}_{N,t}(x,y) + \\ &+ (N-1)\int_{\mathbb{R}}\big[V_N(x-z)-V_N(y-z)\big]\,\gamma^{(2)}_{N,t}(x,z;y,z)\,\mathrm{d}z\,. \end{aligned} \tag{5.80}$$

② For each integer k, $\gamma^{(k)}_{N,t}$ has bounded $\mathcal{L}^2_{\mathrm{reg}}$-norm uniformly in N and in t. When $k=1$, in terms of kernels this reads

$$\int_{\mathbb{R}^2}\big|(1-\partial_x^2)^{1/2}(1-\partial_y^2)^{1/2}\gamma^{(1)}_{N,t}(x,y)\big|^2\,\mathrm{d}x\,\mathrm{d}y < \mathrm{const}(\not{N},\not{t})\,. \tag{5.81}$$

This, in turns, implies that the kernel $\gamma^{(1)}_{N,t}(x,y)$ is bounded in $H^1(\mathbb{R}^2)$ and then also in $L^2(\mathbb{R}^2)$, uniformly in N and in t.

③ Define

$$\sigma_{N,t}(x,y) := (N-1)\int_{\mathbb{R}}\big[V_N(x-z)-V_N(y-z)\big]\,\gamma^{(2)}_{N,t}(x,z;y,z)\,\mathrm{d}z\,. \tag{5.82}$$

Then $\sigma_{N,t}(x,y)$ is the kernel of an operator $\sigma_{N,t}$ that has bounded $\mathcal{L}^2_{\mathrm{reg}}$-norm, and then also $\mathcal{L}^2$-norm, uniformly in N and in t. Equivalently, the kernel $\sigma_{N,t}(x,y)$ is bounded in $H^1(\mathbb{R}^2)$ and then in $L^2(\mathbb{R}^2)$, uniformly in N and in t.

Proof. (5.80) and (5.81) are obtained along the proof of AGT theorem (see [3], equations (1.20) and (1.14), respectively). Also, (5.81) gives immediately uniform $H^1(\mathbb{R}^2)$-boundedness of $\gamma^{(1)}_{N,t}(x,y)$:

$$\begin{aligned} \mathrm{const}(\not{N},\not{t}) &\geqslant \int_{\mathbb{R}^2}\big|(1-\partial_x^2)^{1/2}(1-\partial_y^2)^{1/2}\gamma^{(1)}_{N,t}(x,y)\big|^2\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^2}\overline{\gamma^{(1)}_{N,t}(x,y)}\,(1-\partial_x^2)\,(1-\partial_y^2)\,\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &\geqslant \int_{\mathbb{R}^2}\overline{\gamma^{(1)}_{N,t}(x,y)}\,(1-\partial_x^2-\partial_y^2)\,\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|^2_{H^1(\mathbb{R}^2)}\,. \end{aligned} \tag{5.83}$$

Finally, according to proposition 2.4. of [3], the operator with kernel

$$\widetilde{\sigma}_{N,t}(x,y) := \int_{\mathbb{R}} V_N(x-z)\,\gamma^{(2)}_{N,t}(x,z;y,z)\,\mathrm{d}z \tag{5.84}$$

has uniformly bounded $\mathcal{L}^2_{\mathrm{reg}}$-norm estimated by

$$\|\widetilde{\sigma}_{N,t}\|^2_{\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}))} \leqslant \frac{32\pi}{N^2}\|\hat{V}\|^2_{L^\infty(\mathbb{R})}\|\gamma^{(2)}_{N,t}\|^2_{\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}^2))}\,, \tag{5.85}$$

whence

$$\begin{aligned} \|\sigma_{N,t}\|^2_{\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}))} &\leqslant 64\,\pi(1-1/N)^2\|\hat{V}\|^2_{L^\infty(\mathbb{R})}\|\gamma^{(2)}_{N,t}\|^2_{\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}^2))} \\ &\leqslant \mathrm{const}(\not{N},\not{t}) \end{aligned} \tag{5.86}$$

(recall that each k-th marginal turns out to be $\mathcal{L}^2_{\mathrm{reg}}(L^2(\mathbb{R}^k))$-bounded uniformly in N and in t, the bound depending only on k). □

We are now ready to prove theorem 5.8.1.

Proof (of theorem 5.8.1). We give details only for $k = 1$ case, the generic case being analogous, after a natural modification of theorem 5.8.2 to higher values of k. Also, it suffices to prove $\mathcal{L}^2$-norm convergence at any fixed t: then one applies theorem 3.4.3 to go from Hilbert-Schmidt to trace class norm.

The key point of the proof is the following. By assumption

$$\begin{aligned} 0 &= \lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t\, f(t) \left\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^2}^2 \\ &= \lim_{N\to\infty} \int_{-\infty}^{+\infty} \mathrm{d}t\, f(t)\, G_N(t)\,, \qquad \forall\, f \in L^1(\mathbb{R}) \end{aligned} \tag{5.87}$$

having set

$$G_N(t) := \left\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^2}^2 . \tag{5.88}$$

To prove that $G_N \to 0$ pointwise, let $\tilde{t}$ be any fixed time and let $I \ni \tilde{t}$ be a finite measure interval in $\mathbb{R}$: we will show that

$$\|G_N\|_{H^1(I)} < \text{const}\,, \qquad \text{uniformly in } N\,. \tag{5.89}$$

If (5.89) holds, then $\{G_N\}_N$ is a sequence of uniformly bounded functionals on $H^{-1}(I)$. Since $L^1(I)$ is dense in $H^{-1}(I)$, then (5.87) reads

$$\lim_{N\to\infty} \int_I \mathrm{d}t\, f(t)\, G_N(t) = 0\,, \qquad \forall\, f \in H^{-1}(I)\,. \tag{5.90}$$

Choosing $f(t) = \delta(t - \tilde{t})$ one has $G_N(\tilde{t}) \to 0$ (notice that $\delta \in H^{-\sigma}\ \forall \sigma > \frac{1}{2}$). Since $\tilde{t}$ is arbitrary, then

$$\lim_{N\to\infty} \left\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^2}^2 = 0 \tag{5.91}$$

$\forall\, t \in \mathbb{R}$. Finally, by theorem 3.4.3,

$$\lim_{N\to\infty} \left\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^1} = 0 \tag{5.92}$$

$\forall\, t \in \mathbb{R}$.

So one needs to prove that

$$\|G_N\|_{H^1(I)}^2 = \int_I \left(|G_N(t)|^2 + |G_N'(t)|^2 \right) \mathrm{d}t \ < \text{const} \tag{5.93}$$

uniformly in N. The first summand in the integrand of (5.93) is easily bounded:

$$\begin{aligned} |G_N(t)|^2 &= \left\| \gamma_{N,t}^{(1)} - |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^2}^4 \\ &\leqslant 8 \left(\left\| \gamma_{N,t}^{(1)} \right\|_{\mathcal{L}^2}^4 + \left\| |\varphi_t\rangle\langle\varphi_t| \right\|_{\mathcal{L}^2}^4 \right) \leqslant 16 \end{aligned} \tag{5.94}$$

whence

$$\int_I |G_N(t)|^2\, \mathrm{d}t \leqslant 16|I| \tag{5.95}$$

$|I|$ being the measure of the interval I. Thus, the non-trivial task is to estimate the second summand in the integrand of (5.93).

We compute

$$\begin{aligned} G'_N(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} \big|\,\gamma^{(1)}_{N,t}(x,y) - \varphi_t(x)\overline{\varphi_t(y)}\,\big|^2 \mathrm{d}x\,\mathrm{d}y \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} \Big(\overline{\gamma^{(1)}_{N,t}(x,y)}\,\gamma^{(1)}_{N,t}(x,y) - \overline{\gamma^{(1)}_{N,t}(x,y)}\,\varphi_t(x)\overline{\varphi_t(y)} \\ &\qquad - \gamma^{(1)}_{N,t}(x,y)\overline{\varphi_t(x)}\varphi_t(y) + |\varphi_t(x)|^2|\varphi_t(y)|^2 \Big)\,\mathrm{d}x\,\mathrm{d}y \end{aligned} \tag{5.96}$$

and we treat separately each of the four summands in the r.h.s. of (5.96). The fourth, actually, does not give any contribution, since $\int_{\mathbb{R}^2}|\varphi_t(x)|^2|\varphi_t(y)|^2\mathrm{d}x\,\mathrm{d}y = 1$, so its time derivative is zero. Let us denote the others by

$$\begin{aligned} I_1(N,t) &:= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\,\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ I_2(N,t) &:= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\,\varphi_t(x)\overline{\varphi_t(y)}\,\mathrm{d}x\,\mathrm{d}y \\ I_3(N,t) &:= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} \gamma^{(1)}_{N,t}(x,y)\overline{\varphi_t(x)}\varphi_t(y)\,\mathrm{d}x\,\mathrm{d}y = \overline{I_2(N,t)}\,. \end{aligned} \tag{5.97}$$

The first term gives

$$\begin{aligned} |I_1(N,t)| &= \\ &= \Big|\int_{\mathbb{R}^2} \Big(\overline{\partial_t\gamma^{(1)}_{N,t}(x,y)}\cdot\gamma^{(1)}_{N,t}(x,y) + \overline{\gamma^{(1)}_{N,t}(x,y)}\cdot\partial_t\gamma^{(1)}_{N,t}(x,y)\Big)\mathrm{d}x\,\mathrm{d}y\Big| \\ &\leqslant 2\Big|\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\cdot\partial_t\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &= \frac{\hbar}{m}\Big|\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\cdot\Big[(-\partial_x^2+\partial_y^2)\,\gamma^{(1)}_{N,t}(x,y) + \sigma_{N,t}(x,y)\Big]\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\leqslant \frac{\hbar}{m}\int_{\mathbb{R}^2}\Big(\big|\partial_x\gamma^{(1)}_{N,t}(x,y)\big|^2 + \big|\partial_y\gamma^{(1)}_{N,t}(x,y)\big|^2\Big)\,\mathrm{d}x\,\mathrm{d}y \\ &\qquad + \frac{\hbar}{m}\Big|\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\,\sigma_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \end{aligned} \tag{5.98}$$

where equation (5.80) has been used. By theorem 5.8.2,

$$\begin{aligned} &\int_{\mathbb{R}^2}\Big(\big|\partial_x\gamma^{(1)}_{N,t}(x,y)\big|^2 + \big|\partial_y\gamma^{(1)}_{N,t}(x,y)\big|^2\Big)\,\mathrm{d}x\,\mathrm{d}y \\ &\qquad\qquad < \|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|^2_{H^1(\mathbb{R}^2)} < \mathrm{const}(N,t) \end{aligned} \tag{5.99}$$

and

$$\begin{aligned} &\Big|\int_{\mathbb{R}^2} \overline{\gamma^{(1)}_{N,t}(x,y)}\,\sigma_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\qquad < \|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)}\,\|\sigma_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)} < \mathrm{const}(N,t) \end{aligned} \tag{5.100}$$

so that

$$|I_1(N,t)| < \mathrm{const}(N,t)\,. \tag{5.101}$$

The second and third terms give

$$\begin{aligned} |I_2(N,t)| &= |I_3(N,t)| = \\ &= \Big|\int_{\mathbb{R}^2}\Big(\overline{\partial_t\varphi_t(x)}\cdot\varphi_t(y)\cdot\gamma^{(1)}_{N,t}(x,y) + \overline{\varphi_t(x)}\cdot\partial_t\varphi_t(y)\cdot\gamma^{(1)}_{N,t}(x,y) \\ &\qquad + \overline{\varphi_t(x)}\cdot\varphi_t(y)\cdot\partial_t\gamma^{(1)}_{N,t}(x,y)\Big)\,\mathrm{d}x\,\mathrm{d}y\Big|\,. \end{aligned} \tag{5.102}$$

Set

$$\begin{aligned} J_1(N,t) &:= \int_{\mathbb{R}^2} \overline{\partial_t \varphi_t(x)} \cdot \varphi_t(y) \cdot \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ J_2(N,t) &:= \int_{\mathbb{R}^2} \overline{\varphi_t(x)} \cdot \partial_t \varphi_t(y) \cdot \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ J_3(N,t) &:= \int_{\mathbb{R}^2} \overline{\varphi_t(x)} \cdot \varphi_t(y) \cdot \partial_t \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \end{aligned} \tag{5.103}$$

so that

$$|I_2(N,t)| = |I_3(N,t)| \leqslant |J_1(N,t)| + |J_2(N,t)| + |J_3(N,t)|\,. \tag{5.104}$$

Let us now show the uniform boundedness of each J_i, $i = 1, 2, 3$. By the CNSE (5.32),

$$\begin{aligned} |J_1(N,t)| &= \Big| \int_{\mathbb{R}^2} \overline{\partial_t \varphi_t(x)} \cdot \varphi_t(y) \cdot \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &= \Big| \int_{\mathbb{R}^2} \overline{\Big(-\frac{\hbar^2}{2m}\partial_x^2 \varphi_t(x) + b_1 |\varphi_t(x)|^2 \varphi_t(x) \Big)}\, \varphi_t(y)\, \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &\leqslant \frac{\hbar^2}{2m} \Big| \int_{\mathbb{R}^2} \overline{\partial_x^2 \varphi_t(x)}\, \varphi_t(y)\, \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &\qquad + \frac{\hbar^2 b_1}{2m} \Big| \int_{\mathbb{R}^2} |\varphi_t(x)|^2\, \overline{\varphi_t(x)}\, \varphi_t(y)\, \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \end{aligned} \tag{5.105}$$

By Schwartz inequality and by theorem 5.8.2,

$$\begin{aligned} &\Big| \int_{\mathbb{R}^2} \overline{\partial_x^2 \varphi_t(x)}\, \varphi_t(y)\, \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &\qquad = \Big| \int_{\mathbb{R}^2} \overline{\partial_x \varphi_t(x)}\, \varphi_t(y) \cdot \partial_x \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &\qquad \leqslant \|\partial_x \varphi_t(\,\cdot\,)\|_{L^2(\mathbb{R})}\, \|\varphi_t(\,\cdot\,)\|_{L^2(\mathbb{R})}\, \|\partial_x \gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \\ &\qquad \leqslant \|\varphi_t(\,\cdot\,)\|_{H^1(\mathbb{R})}\, \|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{H^1(\mathbb{R}^2)} \\ &\qquad \leqslant \mathrm{const}(N,t) \end{aligned} \tag{5.106}$$

and

$$\begin{aligned} &\Big| \int_{\mathbb{R}^2} |\varphi_t(x)|^2\, \overline{\varphi_t(x)}\, \varphi_t(y)\, \gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y \Big| \\ &\qquad \leqslant \|\varphi_t(\,\cdot\,)\|^2_{L^\infty(\mathbb{R})} \int_{\mathbb{R}^2} |\varphi_t(x)|\, |\varphi_t(y)|\, |\gamma^{(1)}_{N,t}(x,y)|\,\mathrm{d}x\,\mathrm{d}y \\ &\qquad \leqslant \|\varphi_t(\,\cdot\,)\|^2_{L^\infty(\mathbb{R})}\, \|\varphi_t(\,\cdot\,)\|^2_{L^2(\mathbb{R})}\, \|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \\ &\qquad \leqslant \mathrm{const}(N,t)\,. \end{aligned} \tag{5.107}$$

The treatment of J_2 is completely analogous. So

$$\begin{aligned} |J_1(N,t)| &\leqslant \mathrm{const}(N,t) \\ |J_2(N,t)| &\leqslant \mathrm{const}(N,t)\,. \end{aligned} \tag{5.108}$$

We now estimate J_3:

$$\begin{aligned}\frac{2m}{\hbar}|J_3(N,t)| &= \frac{2m}{\hbar}\Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\cdot\varphi_t(y)\cdot\partial_t\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &= \Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\,\varphi_t(y)\left[(-\partial_x^2+\partial_y^2)\,\gamma^{(1)}_{N,t}(x,y)+\sigma_{N,t}(x,y)\right]\mathrm{d}x\,\mathrm{d}y\Big| \\ &\leqslant \Big|\int_{\mathbb{R}^2}\overline{\partial_x\varphi_t(x)}\,\varphi_t(y)\cdot\partial_x\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\qquad+\Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\,\partial_y\varphi_t(y)\cdot\partial_y\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\qquad+\Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\,\varphi_t(y)\cdot\sigma_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big|\end{aligned} \tag{5.109}$$

where we have used equation (5.80). By Schwartz inequality and theorem 5.8.2

$$\begin{aligned}&\Big|\int_{\mathbb{R}^2}\overline{\partial_x\varphi_t(x)}\,\varphi_t(y)\cdot\partial_x\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\qquad\leqslant \|\partial_x\varphi_t(\,\cdot\,)\|_{L^2(\mathbb{R})}\,\|\varphi_t(\,\cdot\,)\|_{L^2(\mathbb{R})}\,,\|\partial_x\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \\ &\qquad\leqslant \|\varphi_t(\,\cdot\,)\|_{H^1(\mathbb{R})}\,\|\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)\|_{H^1(\mathbb{R}^2)} \\ &\qquad\leqslant \mathrm{const}(N,t)\,,\end{aligned} \tag{5.110}$$

and analogously

$$\Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\,\partial_y\varphi_t(y)\cdot\partial_y\gamma^{(1)}_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big|\leqslant\mathrm{const}(N,t)\,, \tag{5.111}$$

and also

$$\begin{aligned}&\Big|\int_{\mathbb{R}^2}\overline{\varphi_t(x)}\,\varphi_t(y)\cdot\sigma_{N,t}(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big| \\ &\qquad\leqslant \|\varphi_t(\,\cdot\,)\|^2_{L^2(\mathbb{R})}\,\|\sigma_{N,t}(\,\cdot\,,\,\cdot\,)\|_{L^2(\mathbb{R}^2)} \\ &\qquad\leqslant \mathrm{const}(N,t)\,.\end{aligned} \tag{5.112}$$

So

$$|J_3(N,t)|\leqslant\mathrm{const}(N,t)\,. \tag{5.113}$$

Plugging (5.108) and (5.113) into (5.104), one gets

$$|I_2(N,t)|=|I_3(N,t)|\leqslant\mathrm{const}(N,t)\,. \tag{5.114}$$

Finally, by (5.101) and (5.114), one gets

$$|G'_N(t)|\leqslant|I_1(N,t)|+|I_2(N,t)|+|I_3(N,t)|\leqslant\mathrm{const}(N,t)\,. \tag{5.115}$$

Call C this constant. This in turn implies that

$$\int_I|G'_N(t)|^2\,\mathrm{d}t\leqslant C^2|I| \tag{5.116}$$

and (5.95) and (5.116) lead to equation (5.93). The desired $H^1(I)$-boundedness of G_N, uniformly in N, is proved. □

It is noteworthy to remark that $G'_N(t) \leqslant \mathrm{const}(N,t)$ amounts to the uniform boundedness of

$$\int_{\mathbb{R}^2} \partial_t \big| \, \gamma^{(1)}_{N,t}(x,y) - \varphi_t(x)\overline{\varphi_t(y)} \, \big|^2 \, \mathrm{d}x \, \mathrm{d}y \,,$$

but neither $\partial_t \gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)$ nor $\partial_t \varphi_t(\,\cdot\,)$ are separately L^2-bounded, uniformly in N and in t.

Indeed, since φ solves the CLSE (5.32), it is seen that $L^2(\mathbb{R})$-boundedness of $\partial_t \varphi_t(\,\cdot\,)$ requires $H^2(\mathbb{R})$-boundedness of $\varphi_t(\,\cdot\,)$. But it is known (see, e.g., theorem 4.8.1. in [27]) that for every initial datum in H^2 there exist $T_\pm > 0$ and a unique maximal solution

$$\varphi \in C((-T_-, T_+), H^2(\mathbb{R})) \cap C^1((-T_-, T_+), L^2(\mathbb{R}))$$

of (5.32), with H^2-norm blow-up as $t \to T_\pm$, so that one cannot localise the interval I in (5.93) around any time t and independently of the initial datum.

Similarly, since $\gamma^{(1)}_{N,t}$ solves (5.80), it is seen that $L^2(\mathbb{R}^2)$-boundedness of $\partial_t \gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)$ requires uniform $H^2(\mathbb{R}^2)$-boundedness of $\gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)$, while only uniform $H^1(\mathbb{R}^2)$-boundedness is known.

Uniform L^2-boundedness of $\partial_t \gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,)$ and $\partial_t \varphi_t(\,\cdot\,)$ would have helped the proof via the estimate

$$\begin{aligned} |G'_N(t)| &= \Big| \frac{\mathrm{d}}{\mathrm{d}t} \big\| \, \gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t| \, \big\|^2_{\mathcal{L}^2} \Big| \\ &\leqslant 4 \, \big\| \, \partial_t \gamma^{(1)}_{N,t}(\,\cdot\,,\,\cdot\,) \, \big\|_{L^2(\mathbb{R}^2)} + 4 \, \big\| \, \partial_t \varphi_t(\,\cdot\,) \, \big\|_{L^2(\mathbb{R}^2)} \,, \end{aligned} \tag{5.117}$$

the inequality following by

$$\begin{aligned} &\Big| \frac{\mathrm{d}}{\mathrm{d}t} \big\| \, \rho_{N,t} \, \big\|^2_{\mathcal{L}^2} \Big| \leqslant \Big| \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^2} \big| \, \rho_{N,t}(x,y) \, \big|^2 \, \mathrm{d}x \, \mathrm{d}y \Big| \\ &= \Big| \int_{\mathbb{R}^2} \Big(\overline{\partial_t \rho_{N,t}(x,y)} \cdot \rho_{N,t}(x,y) + \overline{\rho_{N,t}(x,y)} \cdot \partial_t \rho_{N,t}(x,y) \Big) \, \mathrm{d}x \, \mathrm{d}y \Big| \\ &\leqslant 2 \, \big\| \, \partial_t \rho_{N,t}(\,\cdot\,,\,\cdot\,) \, \big\|_{L^2(\mathbb{R}^2)} \, \big\| \, \rho_{N,t}(\,\cdot\,,\,\cdot\,) \, \big\|_{L^2(\mathbb{R}^2)} \,, \end{aligned}$$

where $\rho_{N,t} = \gamma^{(1)}_{N,t} - |\varphi_t\rangle\langle\varphi_t|$ (so that $\big\| \, \rho_{N,t}(\,\cdot\,,\,\cdot\,) \, \big\|_{L^2(\mathbb{R}^2)} \leqslant 2$).

Instead of estimate (5.117), our strategy has been to perform the full computation in (5.96), with the insertion of the CNSE or the first BBGKY equation in each of the resulting summands (5.97). This enabled us to substitute each time derivative with a double space derivative and, consequently, to deal with duality products of H^{-1} functions tested on H^1-functions, thus leading to the desired uniform boundedness.

Chapter 6

Interparticle correlations in the condensate

Interaction among particles establishes *correlations* in the many-body state: their theoretical understanding is crucial both in view of their experimental observation and for their influence on other features of the state under consideration. In the presence of B.E.C., the question becomes more concrete, because one expects that the condensed state is roughly speaking a totally factorised state.

In fact, in the trivial case of systems of non-interacting particles, B.E.C. shows up as a totally factorised N-body state. In the more realistic case of even weakly interacting systems, if condensation is present, e.g., in the pure state Ψ_N, and if the condensate wave function is recognised to be some φ, one would still be tempted to make some approximation $\Psi_N \approx \varphi^{\otimes N}$. For instance (see Sec. 1.3), observation of B.E.C. is carried on by illuminating the cloud of atoms of a cooled and trapped alkali gas by resonant laser light and its shadow is imaged onto a CCD camera: the tacit assumption behind this discrete counting is that "almost all" particles have their own individuality in the one-body state φ, as it would be in the N-th tensor power $\varphi^{\otimes N}$.

So here the problem of correlations is the problem of to what extent a condensed state differs from some $\varphi^{\otimes N}$, and in which sense this difference can be expressed and detected. In this chapter, following the work in progress [82], we are discussing the reasonable way to state and to approach to such a problem, together with some related results.

Discussing correlations at *finite* N, within an analysis where rigorous results are *asymptotic* only and where the system appears as totally factorised as $N \to \infty$ (at least when observed at any fixed k-body level), turns out to be a rather subtle issue. In Sec. 6.1 we will sketch it and its main difficulties, emphasizing that the discussion has to be carried on simultaneously at the many-body state level and at the one-body marginal level. This is done by introducing a trial wave function showing condensation at the level of marginals and reproducing the energy of the true state.

In Sec. 6.2 we focus on the specific form of correlations emerging in the ground state of a dilute Bose gas with B.E.C. Their effect is to lower by a macroscopic quantity the expectation value of the energy in the corresponding uncorrelated state. This depends on the presence of an appropriate correlation structure which modifies the purely factorised $\varphi^{\otimes N}$, independently of the latter, as discussed in Sec. 6.3.

The counterpart at the level of marginals is a typical correlation structure which turns out to be crucial in many aspect of the rigorous analysis on B.E.C., as the determination

of the correct evolutionary hierarchy for reduced density matrices (Sec. 6.4).

In fact, emergence and persistence of short scale correlations is recognised in the evolution of the condensate, although one still misses an understanding of the full correlation structure accounting for the energy of a condensate (Sec. 6.5 and 6.5).

6.1 Correlations in many-body condensed states

Let us start with focusing on the main physical and technical obstructions emerging in controlling the difference $(\Psi_N - \varphi^{\otimes N})$ in some sense, where Ψ_N is the many-body state with B.E.C. and φ is the one-body condensate wave function. There are difficulties both at the many body level ("$\mathbb{M}$") and at the reduced one-body level ("$\mathbb{B}$").

($\mathbb{M}_1$) Ψ_N is not accessible explicitly, the many-body problem is too hard.

($\mathbb{M}_2$) One can instead introduce an explicit *trial wave function* $\Psi_N^{\rm trial}$, which has not the form of a tensor product, but by construction is close in norm to $\varphi^{\otimes N}$. If one could prove that Ψ_N and $\Psi_N^{\rm trial}$ too are close in norm, then Ψ_N would be almost factorised as $\Psi_N^{\rm trial}$, which would provide a reasonable structure of the correlations in the true N-body state. Yet, since the distance is taken with the $\|\ \|_{\mathcal{H}_N}$-norm, the choice of the trial function is extremely delicate, as N is very large. A "physically irrelevant" change in $\Psi_N^{\rm trial}$ can make it essentially orthogonal to the Ψ_N, or to the $\varphi^{\otimes N}$ it was close to. For example, if $\varphi, \xi \in \mathcal{H}$ are two orthonormal one-body states, then[1]

$$\left\| \left(\varphi^{\otimes(N-1)} \otimes \xi\right)_{\rm sym} - \varphi^{\otimes N} \right\|_{\mathcal{H}_N} = \sqrt{2} \tag{6.1}$$

independently of N, while one would tend to say that the distinction among such two N-body states is negligible as $N \to \infty$, since they are two bosonic state differing only by one particle out of N. So, evaluating distances in norm is "too detailed" – it carries information on all particles – and does not provide a useful comparison.

($\mathbb{M}_3$) Although example in (6.1) may be fictitious with respect to realistic condensed states, which may indeed be close *in norm* to some $\varphi^{\otimes N}$, it is not trivial at all that $\Psi_N^{\rm trial}$ and Ψ_N are close also *in the energy form*. In fact, $\varphi^{\otimes N}$ typically does not approximates in this sense the "true" Ψ_N (the Hamiltonian being unbounded, vectors close in norms are not necessarily close in the form of H_N) and this causes difficulties in choosing a good trial function: correlations deeply influence the computation of $\langle H_N \rangle$. We are discussing a concrete example of that in Sec. 6.2 and 6.3.

At the one-body level things seem more promising, at least because, unlike ($\mathbb{M}_1$), it is easier to handle marginals and, unlike ($\mathbb{M}_2$), by computing distances of marginals one

[1] To prove (6.1), recall that, by (3.20),

$$\begin{aligned}\left\| \left(\varphi^{\otimes(N-1)} \otimes \xi\right)_{\rm sym} - \varphi^{\otimes N} \right\|^2_{\mathcal{H}_N} = \\ = \left\| \frac{1}{\sqrt{N}} \xi \otimes \varphi \otimes \cdots \otimes \varphi + \frac{1}{\sqrt{N}} \varphi \otimes \xi \otimes \cdots \otimes \varphi + \cdots - \varphi \otimes \varphi \otimes \cdots \otimes \varphi \right\|^2_{\mathcal{H}_N},\end{aligned}$$

that is, the norm of a linear combination of $N+1$ orthonormal terms: the first N terms contribute with $(\frac{1}{\sqrt{N}})^2 \cdot N$, the last one contributes with 1, so that the r.h.s. amounts to 2.

gets rid of physically irrelevant discrepancies between Ψ_N, Ψ_N^{trial}, and $\varphi^{\otimes N}$. This is the case with respect to the example in (6.1): by theorem 3.3.1,

$$\gamma_N^{(1)} = \frac{N-1}{N}|\varphi\rangle\langle\varphi| + \frac{1}{N}|\xi\rangle\langle\xi| \tag{6.2}$$

is the one-particle reduced density matrix corresponding to $(\varphi^{\otimes(N-1)} \otimes \xi)_{\mathrm{sym}}$ and

$$\left\| \gamma_N^{(1)} - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})} = \frac{1}{N}\left\| |\xi\rangle\langle\xi| - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})} \leqslant \frac{2}{N} \xrightarrow{N\to\infty} 0\,. \tag{6.3}$$

Instead, one encounters another kind of difficulties in characterizing correlations present in the many-body state.

($\mathbb{D}_1$) For finite N, in general $\gamma_N^{(1)}$ (or any k-th marginal) does not reconstruct γ_N uniquely, therefore does not controls correlations in the N-body state. See, e.g., the elementary example discussed before definition 3.3.2.

($\mathbb{D}_2$) As $N \to \infty$, convergence of $\gamma_N^{(1)}$ to a projection does not imply in general factorisation of Ψ_N in the limit. This is just the case of (6.3) vs (6.1).

($\mathbb{D}_3$) It may happen that the effects of two-body correlations in the many-body wave function essentially do not contribute to the first (or any k-th) marginal.

Thus, a rigorous analysis of correlations in the condensed many-body state relies on a compromise (and a synergy) in adopting distinct tools.

General strategy. *One tries to build explicit interparticle correlations in a trial wave function Ψ_N^{trial} in such a way that:*

i) the corresponding $\gamma_N^{(1),\mathrm{trial}}$ is essentially $|\varphi\rangle\langle\varphi|$ as it would be for the completely factorised $\varphi^{\otimes N}$,

ii) $\Psi_N^{\mathrm{trial}} \approx \Psi_N$ in norm,

iii) $\Psi_N^{\mathrm{trial}} \approx \Psi_N$ also in the energy form.

For i) to hold it suffices that $\Psi_N^{\mathrm{trial}} \approx \varphi^{\otimes N}$ in norm, since by (2.105)

$$\left\| \gamma_N^{(1),\mathrm{trial}} - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})} \leqslant 2\left\| \Psi_N^{\mathrm{trial}} - \varphi^{\otimes N} \right\|_{\mathcal{H}_N}.$$

By the same estimates (2.72) *and* (2.105)*, i) and ii) prove that Ψ_N has complete condensation:*

$$\begin{aligned}\left\| \gamma_N^{(1)} - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})} &\leqslant \left\| \gamma_N^{(1)} - \gamma_N^{(1),\mathrm{trial}} \right\|_{\mathcal{L}^1(\mathcal{H})} + \left\| \gamma_N^{(1),\mathrm{trial}} - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})} \\ &\leqslant 2\left\| \Psi_N - \Psi_N^{\mathrm{trial}} \right\|_{\mathcal{H}_N} + \left\| \gamma_N^{(1),\mathrm{trial}} - |\varphi\rangle\langle\varphi| \right\|_{\mathcal{L}^1(\mathcal{H})}.\end{aligned}$$

In all that, a limit $N \to \infty$ enters to make asymptotic statements and only a final control of error terms provides complete results for large finite N.

6.2 Ground state correlations: Jastrow and Dyson factors

As an example, we turn back to the case of (asymptotic) 100% B.E.C. in a three-dimensional dilute trapped Bose gas, after the analysis of Lieb, Seiringer and Yngvason (Sec. 3.5). There, condensation is proved in the ground state $\Psi_N^{\text{g.s.}}$ of the Hamiltonian H_N (3.77) and the condensate wave function turns out to be the minimizer φ^{GP} of the GP energy functional (3.72).

The ground state energy asymptotics (3.80), i.e., $(\Psi_N^{\text{g.s.}}, H_N \Psi_N^{\text{g.s.}})/N \xrightarrow{N\to\infty} E^{\text{GP}}$, is obtained in [77] through a separate control from above and from below and the two bounds are proved to converge to the same limit. For the upper bound, in particular, the following trial function Ψ_N^{LSY} is used:

$$\Psi_N^{\text{LSY}}(x_1,\dots,x_N) := \frac{W_N(x_1,\dots,x_N)\prod_{i=1}^N \varphi^{\text{GP}}(x_i)}{\left\| W_N(x_1,\dots,x_N)\prod_{i=1}^N \varphi^{\text{GP}}(x_i) \right\|} . \tag{6.4}$$

Here $x_1,\dots,x_N \in \mathbb{R}^3$, φ^{GP} is the GP minimizer discussed in theorem 3.5.2, and interparticle correlations are encoded in the term

$$W_N(x_1,\dots,x_N) := \prod_{i=2}^N F_N\big(t_i(x_1,\dots,x_i)\big) \tag{6.5}$$

with

$$t_i := \min\left\{ |x_i - x_j| \,,\, 1 \leqslant j \leqslant i-1 \right\} \tag{6.6}$$

$$F_N(r) := \begin{cases} \mathcal{F}_N(r)/\mathcal{F}_N(b_N) & r \leqslant b_N \\ 1 & r \geqslant b_N \end{cases} ; \tag{6.7}$$

b_N is some N-indexed cut-off parameter of order

$$b_N \sim \overline{\rho}_N^{-1/3} \sim N^{-1/3} \tag{6.8}$$

(that is, of the same order as the mean interparticle distance),

$$\mathcal{F}_N(r) := f_N(x) \,, \qquad r := |x| \,, \tag{6.9}$$

and f_N is the spherically symmetric solution of the *zero energy scattering problem*

$$\begin{cases} (-2\mu\Delta + V_N)\, f_N &= 0 \\ \lim\limits_{|x|\to\infty} f_N(x) &= 1 \,. \end{cases} \tag{6.10}$$

It can be proved that

$$0 \leqslant f_N(x) \leqslant 1 \tag{6.11}$$

$$\mathcal{F}_N'(r) \geqslant 0 \tag{6.12}$$

$$f_N(x) \sim 1 - \frac{a_N}{|x|} \quad \text{as } |x| \to \infty \tag{6.13}$$

$$8\pi\mu a_N = \int_{\mathbb{R}^3} V_N(x) f_N(x)\, \mathrm{d}x \tag{6.14}$$

where a_N is the s-wave scattering length of V_N. Consequently, $F_N(r)$ is by construction the constant 1 but at the scale $r \lesssim N^{-1/3}$, where it stays bounded below 1.

In view of the general strategy above, we point out that

i) $\gamma_N^{(1),\mathrm{LSY}} \to |\varphi^{\mathrm{GP}}\rangle\langle\varphi^{\mathrm{GP}}|$,

ii) $\gamma_N^{(1)} \to |\varphi^{\mathrm{GP}}\rangle\langle\varphi^{\mathrm{GP}}|$, but nothing rigorous is known about $\Psi_N^{\mathrm{LSY}} \approx \Psi_N$ in norm,

iii) $\Psi_N^{\mathrm{LSY}} \approx \Psi_N$ in the energy form,

iv) $\Psi_N^{\mathrm{LSY}} \approx \left(\varphi^{\mathrm{GP}}\right)^{\otimes N}$ in norm,

v) Ψ_N^{LSY} is *far* from $\left(\varphi^{\mathrm{GP}}\right)^{\otimes N}$ in energy.

Point *i)* follows from *iv)*, by (2.105):

$$\left\| \gamma_N^{(1),\mathrm{LSY}} - |\varphi^{\mathrm{GP}}\rangle\langle\varphi^{\mathrm{GP}}| \right\|_{\mathcal{L}^1(L^2(\mathbb{R}^3))} \leqslant \left\| \Psi_N^{\mathrm{LSY}} - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2(\mathbb{R}^{3N})}. \tag{6.15}$$

Concerning point *ii)*, the trial function (6.4) has *not* been proved to be close *also in norm* to the true ground state (asymptotically in N), so currently the general strategy above is only partially addressed. Actually, in Lieb and Seiringer's work [74] on B.E.C. for dilute trapped gases, one controls $\gamma_N^{(1)}$ by a detailed examination of Ψ_N, writing it (apart from normalisation) as $\Psi_N(x_1,\dots,x_N) \equiv \varphi^{\mathrm{GP}}(x_1)\cdots\varphi^{\mathrm{GP}}(x_N)\mathcal{W}_N(x_1,\dots,x_N)$, where $\mathcal{W}_N$ is necessarily non-negative and permutationally symmetric and has the role of (but it is distinct from) the analogous W_N defined in (6.5). By combining the "localization of the energy" technique and the generalised Poincaré inequality (see [74, 75] for details), one finally gets B.E.C. as $\gamma_N^{(1)} \to |\varphi^{\mathrm{GP}}\rangle\langle\varphi^{\mathrm{GP}}|$.

Point *iii)* is addressed in [77], as we have just mentioned above. There one proves that $(\Psi_N^{\mathrm{LSY}}, H_N\Psi_N^{\mathrm{LSY}})/N \geqslant (\Psi_N^{\mathrm{g.s.}}, H_N\Psi_N^{\mathrm{g.s.}})/N$ and that the difference between these two quantities asymptotically vanishes, both converging to E^{GP}: that is, $\Psi_N^{\mathrm{g.s.}}$ and Ψ_N^{LSY} are *close in energy*. (Although $\Psi_N^{\mathrm{LSY}} \notin \mathcal{H}_{N,\mathrm{sym}} = L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$, it is nevertheless an acceptable trial wave function, since bosonic and absolute ground state energies coincide.)

Point *iv)* is given by the following.

Theorem 6.2.1. Take Ψ_N^{LSY} as in (6.4). Then

$$\lim_{N\to\infty} \left\| \Psi_N^{\mathrm{LSY}} - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2(\mathbb{R}^{3N})} = 0 \tag{6.16}$$

Proof. We are going to provide an upper bound of $\| \Psi_N^{\mathrm{LSY}} - (\varphi^{\mathrm{GP}})^{\otimes N} \|_{L^2}$ which vanishes as $N \to \infty$. We will be using the following: (a) φ^{GP} is bounded and L^2-normalised, (b) $0 \leqslant F_N(r) \leqslant 1$ and the support of $1 - F_N$ is $[0, b_N]$, which for our purposes can be concretely taken to be $[0, N^{-1/3}]$, (c) a more refined estimate for $1 - F_N$ available in the literature, see (6.30).

Actually we will prove

$$\left\| \Phi_N - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2(\mathbb{R}^{3N})} \xrightarrow{N\to\infty} 0 \tag{6.17}$$

with

$$\Phi_N(x_1,\dots,x_N) := W_N(x_1,\dots,x_N) \prod_{i=1}^{N} \varphi^{\mathrm{GP}}(x_i) \tag{6.18}$$

(so that $\Psi_N^{\mathrm{LSY}} = \Phi_N/\|\Phi_N\|$), because, by (2.106),

$$\left\| \Psi_N^{\mathrm{LSY}} - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2} \leqslant \frac{2 \left\| \Phi_N - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2}}{1 - \left\| \Phi_N - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|_{L^2}} \xrightarrow{N\to\infty} 0\,. \tag{6.19}$$

First, let us re-write (6.5) as

$$W_N(x_1, \dots, x_N) = \prod_{i=2}^{N} F_N(|x_i - x_{j(i)}|) \tag{6.20}$$

where by $x_{j(i)}$ one means the closest coordinate to x_i among $\{x_1, x_2, \dots, x_{i-1}\}$. One has

$$\begin{aligned} &\left\| \Phi_N - \left(\varphi^{\mathrm{GP}}\right)^{\otimes N} \right\|^2_{L^2(\mathbb{R}^{3N})} \\ &\quad = \int_{\mathbb{R}^{3N}} \Big| 1 - \prod_{i=2}^{N} F_N(|x_i - x_{j(i)}|) \Big|^2 \prod_{i=1}^{N} \left|\varphi^{\mathrm{GP}}(x_i)\right|^2 \mathrm{d}x_1 \cdots \mathrm{d}x_N \\ &\quad \leqslant 2 \int_{\mathbb{R}^{3N}} \Big(1 - \prod_{i=2}^{N} F_N(|x_i - x_{j(i)}|) \Big) \prod_{i=1}^{N} \left|\varphi^{\mathrm{GP}}(x_i)\right|^2 \mathrm{d}x_1 \cdots \mathrm{d}x_N \end{aligned} \tag{6.21}$$

because $F_N \leqslant 1$. Notice that one cannot use Lebesgue dominated convergence, because the integration space enlarges with N.

Then set

$$\omega_N(t) := 1 - F_N(t) \tag{6.22}$$

$$\omega^{(i)} := \omega_N(|x_i - x_{j(i)}|) \tag{6.23}$$

Such ω_N is supported in $[0, N^{-1/3}]$ with values in $[0, 1]$. The correlation term in the r.h.s. integrand of (6.21) reads

$$\begin{aligned} 1 - \prod_{i=2}^{N} (1 - \omega^{(i)}) &= \sum_{i=2}^{N} \omega^{(i)} - \sum_{2 \leqslant i < j \leqslant N} \omega^{(i)} \omega^{(j)} \; + \cdots + \; (-)^N \, \omega^{(2)} \omega^{(3)} \cdots \omega^{(N)} \\ &= \sum_{k=1}^{N-1} (-)^{k+1} \sum_{2 \leqslant i_1 < \cdots < i_k \leqslant N} \omega^{(i_1)} \omega^{(i_2)} \cdots \omega^{(i_k)} \end{aligned} \tag{6.24}$$

so that

$$(6.21) \;\leqslant\; 2 \sum_{k=1}^{N-1} (-)^{k+1} \sum_{2 \leqslant i_1 < \cdots < i_k \leqslant N} I^{(N,k)}_{i_1, \dots, i_k} \tag{6.25}$$

with

$$I^{(N,k)}_{i_1, \dots, i_k} := \int_{\mathbb{R}^{3N}} \prod_{s=1}^{k} \omega_N(|x_{i_s} - x_{j(i_s)}|) \prod_{i=1}^{N} \left|\varphi^{\mathrm{GP}}(x_i)\right|^2 \mathrm{d}x_1 \cdots \mathrm{d}x_N\,. \tag{6.26}$$

In each $I^{(N,k)}_{i_1, \dots, i_k}$ the ω_N-factors depend on the variables $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ and other variables $x_{j(i_1)}, x_{j(i_2)}, \dots, x_{j(i_k)}$ that, by construction, are among $\{x_1, x_2, \dots, x_{i_k}\}$, since i_k is the *largest* index among $\{i_1, i_2, \dots i_k\}$. Variables $x_{i_k+1}, \dots, x_N$ are only in the φ^{GP}-factors and are integrated out by $\int_{\mathbb{R}^3} |\varphi^{\mathrm{GP}}(x)|^2 \, \mathrm{d}x = 1$.

Next, partition the remaining domain of integration, that is, $\mathbb{R}^{3i_k}$, into separated regions where coordinates $x_{j(i_1)}, x_{j(i_2)}, \dots, x_{j(i_k)}$ are indexed by constant labels. In each of these regions, take the coordinate change

$$\begin{cases} z_{i_1} &= x_{i_1} - x_{j(i_1)} \\ z_{i_2} &= x_{i_2} - x_{j(i_2)} \\ \vdots & \vdots \\ z_{i_k} &= x_{i_k} - x_{j(i_k)} \end{cases} \tag{6.27}$$

leaving all the other variables unchanged. In the *new* variables, the ω_N-factors take the form $\prod_{s=1}^{k} \omega_N(|z_{i_s}|)$, then there are k factors each with the form $|\varphi^{\mathrm{GP}}(z_{i_s} - x_{i_s})|^2$ and all the others are expressed in the original variable. Bound each $|\varphi^{\mathrm{GP}}(z_{i_s} - x_{i_s})|^2$ by $\|\varphi^{\mathrm{GP}}\|_\infty^2$: this *decouples* the variables $z_{i_1}, \dots, z_{i_k}$ from the others, so that the remaining φ^{GP}-factors can be integrated out. This can be done for all the *disjoint* regions of the integration domain.[2] Thus,

$$\begin{aligned} I^{(N,k)}_{i_1,\dots,i_k} &\leqslant \big(\|\varphi^{\mathrm{GP}}\|_\infty^2 \big)^k \int_{\mathbb{R}^{3k}} \prod_{s=1}^{k} \omega_N(|z_{i_s}|)\, \mathrm{d}z_{i_1} \cdots \mathrm{d}z_{i_k} \\ &= \big(\|\varphi^{\mathrm{GP}}\|_\infty^2 \big)^k \Big(\int_{\mathbb{R}^3} \omega_N(|z|)\, \mathrm{d}z \Big)^k . \end{aligned} \tag{6.28}$$

A more refined estimate is now needed than

$$\int_{\mathbb{R}^3} \omega_N(|z|)\, \mathrm{d}z = 4\pi \int_0^{N^{-1/3}} \omega(r)\, r^2\, \mathrm{d}r \leqslant 4\pi \int_0^{N^{-1/3}} r^2\, \mathrm{d}r = \frac{4\pi}{3}\, \frac{1}{N} . \tag{6.29}$$

To this aim, let us make use of (7.24), which actually applies to a f_N very similar to the one we are dealing with, which is defined in (6.10). There f_N solves the *lowest energy* scattering problem in a ball, and is 1 outside; here it is assumed to solve the *zero-energy* scattering problem on the whole space, with the condition to be 1 at infinity. It can be easily checked that (7.24) in the current case leads to

$$\omega_N(|z|) = 1 - F_N(|z|) = 1 - \frac{\mathcal{F}_N(|z|)}{\mathcal{F}_N(b_N)} \leqslant 1 - f_N(z) \leqslant \frac{a_N}{|z| + a_N} \mathbb{1}_{\mathcal{B}_N}(z) \tag{6.30}$$

[2] As an example, take

$$I^{(N,3)}_{3,5,6} = \int_{(\mathbb{R}^3)^6} \omega_N(|x_3 - x_{j(3)}|)\, \omega_N(|x_5 - x_{j(5)}|)\, \omega_N(|x_6 - x_{j(6)}|) \prod_{i=1}^{6} \big|\varphi^{\mathrm{GP}}(x_i)\big|^2\, \mathrm{d}x_1 \cdots \mathrm{d}x_6$$

and in the region Ω of $(\mathbb{R}^3)^6$ where, e.g., $x_{j(3)} = x_1, x_{j(5)} = x_1, x_{j(6)} = x_5$, one has

$$\begin{aligned} I^{(N,3)}_{3,5,6} &= \int_\Omega \omega_N(|z_3|)\, \omega_N(|z_5|)\, \omega_N(|z_6|)\, \big|\varphi^{\mathrm{GP}}(x_1)\big|^2\, \big|\varphi^{\mathrm{GP}}(x_2)\big|^2\, \big|\varphi^{\mathrm{GP}}(z_3 + x_1)\big|^2 \cdot \\ &\qquad \cdot \big|\varphi^{\mathrm{GP}}(x_4)\big|^2\, \big|\varphi^{\mathrm{GP}}(z_5 + x_1)\big|^2\, \big|\varphi^{\mathrm{GP}}(z_6 + z_5 - x_1)\big|^2\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}z_3\, \mathrm{d}x_4\, \mathrm{d}z_5\, \mathrm{d}z_6 \\ &\leqslant \big(\|\varphi^{\mathrm{GP}}\|_\infty^2 \big)^3 \int_\Omega \omega_N(|z_3|)\, \omega_N(|z_5|)\, \omega_N(|z_6|)\, \big|\varphi^{\mathrm{GP}}(x_1)\big|^2\, \big|\varphi^{\mathrm{GP}}(x_2)\big|^2 \cdot \\ &\qquad \cdot \big|\varphi^{\mathrm{GP}}(x_4)\big|^2\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}z_3\, \mathrm{d}x_4\, \mathrm{d}z_5\, \mathrm{d}z_6 \end{aligned}$$

so that

$$I^{(N,3)}_{3,5,6} \leqslant \big(\|\varphi^{\mathrm{GP}}\|_\infty^2 \big)^3 \Big(\int_{\mathbb{R}^3} \omega_N(|z|)\, \mathrm{d}z \Big)^3 .$$

where $\mathbb{1}_{\mathcal{B}_N}$ is the characteristic function of the ball of radius $b_N = N^{-1/3}$ in $\mathbb{R}^3$, $a_N = N_0\, a/N$ by (3.76) and, with no loss of generality, we take here $a_N = N^{-1}$. Plugging (6.30) into $\int \omega_N$, one gets

$$\begin{aligned}\int_{\mathbb{R}^3} \omega_N(|z|)\,\mathrm{d}z &\leqslant \int_{|z|\leqslant N^{-1/3}} \frac{a_N}{|z|+a_N} = 4\pi a_N \int_0^{N^{-1/3}} \frac{r^2}{r+a_N}\,\mathrm{d}r \\ &= \frac{4\pi}{N^3}\int_0^{N^{2/3}} \frac{s^2}{s+1}\,\mathrm{d}s \\ &\leqslant \frac{4\pi}{N^3}\int_0^{N^{2/3}} s\,\mathrm{d}s \\ &= \frac{2\pi}{N^{5/3}}\end{aligned} \tag{6.31}$$

which is a refined version of (6.29). Plugging (6.31) into (6.28), one gets

$$I^{(N,k)}_{i_1,\dots,i_k} \leqslant \frac{C^k}{N^{5k/3}}\,. \tag{6.32}$$

for some positive constant C.

The whole sum $\sum I^{(N,k)}_{i_1,\dots i_k}$ in (6.25), at fixed k and N, is in turn estimated noticing that it is made of $\binom{N-1}{k}$ summands, corresponding to all possible choices of arranging indices in such a way that $2 \leqslant i_1 < \cdots < i_k \leqslant N$, each leading to a $I^{(N,k)}_{i_1,\dots,i_k}$ which is estimated by (6.32). Then (6.25) and (6.32) give

$$\begin{aligned}(6.21) \;&\leqslant\; 2\sum_{k=1}^{N-1}(-)^{k+1}\sum_{2\leqslant i_1<\cdots<i_k\leqslant N} I^{(N,k)}_{i_1,\dots,i_k} \\ &\leqslant\; 2\sum_{k=1}^{N-1}(-)^{k+1}\binom{N-1}{k}\frac{C^k}{N^{5k/3}} = \sum_{k=1}^{N-1}(-)^{k+1}\,A_N(k)\end{aligned} \tag{6.33}$$

with

$$A_N(k) := \binom{N-1}{k}\frac{2\,C^k}{N^{5k/3}}. \tag{6.34}$$

In the above sum, terms alternate their sign, starting with $A_N(1) > 0$. Moreover,

$$\frac{A_N(k+1)}{A_N(k)} = \frac{C}{(k+1)(N-k)\,N^{5/3}} \leqslant 1\,, \quad \forall\, k = 1,2,\dots,N-1 \tag{6.35}$$

(eventually as $N \to \infty$), whence $A_N(k+1) \leqslant A_N(k)$. This means that

$$\begin{aligned}(6.21) \;&\leqslant\; \sum_{k=1}^{N-1}(-)^{k+1}\,A_N(k) \\ &= A_N(1) - A_N(2) + A_N(3) + \cdots + (-)^N A_N(N-1) \\ &\leqslant A_N(1) =\; 2\,C\,\frac{N-1}{N^{5/3}} \\ &\leqslant\; \text{const}\cdot\frac{1}{N^{2/3}} \xrightarrow{N\to\infty} 0\end{aligned} \tag{6.36}$$

and (6.17) is proved. (One can see that the refined estimate (6.31) is really necessary, because the original estimate (6.29) would have led to a $\mathcal{O}(1)$ term in the r.h.s. of (6.36).) □

Last, concerning point *v*) above, we shall now see in (6.40) that $\Psi_N^{\mathrm{g.s.}}$ and Ψ_N^{LSY} have *lower* expectation value of the energy than $(\varphi^{\mathrm{GP}})^{\otimes N}$: this is due to the lower probability that two particles in the state Ψ_N^{LSY} stay at a distance $\lesssim N^{-1/3}$, with respect to the purely factorised case. Also, correlations present in the ground state decrease its energy with respect to *any* tensor power N-body wave function.

To state all that rigorously, let us tune the scattering length in the GP energy functional (3.72), as previously done in (4.28), by introducing a parameter ν in the generalised functional

$$\mathcal{E}_\nu^{\mathrm{GP}}[\varphi] = \int_{\mathbb{R}^3} \big(\mu|\nabla\varphi|^2 + U|\varphi|^2 + 4\pi\mu\,\nu|\varphi|^4\big)\,\mathrm{d}x \qquad \Big(\mu = \frac{\hbar^2}{2m}\Big)\,, \tag{6.37}$$

on the same variational domain $\mathcal{D}$ (3.73). Also, let us recall that the scattering length a of the pair interaction V is bounded above by its first Born approximation, according to the Spruch-Rosenberg inequality [105]

$$a \leqslant \frac{b}{8\pi\mu} \tag{6.38}$$

where

$$b := \int_{\mathbb{R}^3} V(x)\,\mathrm{d}x = \|V\|_{L^1}\,. \tag{6.39}$$

Theorem 6.2.2. Assume the hypotheses and take the notation of theorem 3.5.2, but with the irrelevant simplification $N_0 = 1$. Assume, also, that V has finite L^1-norm. Then any $\xi^{\otimes N}$ (with $\xi \in \mathcal{D} \subset \mathcal{H}$, $\|\xi\| = 1$) has asymptotically higher expectation value of the energy than the ground state Ψ_N:

$$\lim_{N\to\infty} \Big(\xi^{\otimes N}, \frac{H_N}{N}\,\xi^{\otimes N}\Big) = \mathcal{E}^{\mathrm{GP}}_{b/(8\pi\mu)}[\,\xi\,] \geqslant E^{\mathrm{GP}}\,. \tag{6.40}$$

Assume, in addition, that $b/(8\pi\mu) > a$: then inequality (6.40) becomes strict. Then the best approximation in energy of Ψ_N given by a tensor power N-body state $\xi^{\otimes N}$ is achieved by $\xi = \psi^{\mathrm{GP}}$, the latter being the minimizer of $\mathcal{E}^{\mathrm{GP}}_{b/(8\pi\mu)}$ defined in the same way and with the same features as in theorem 3.5.2.

Proof. One has

$$\begin{aligned}
\big(\,\xi^{\otimes N},\, H_N\,\xi^{\otimes N}\,\big) &= \int_{\mathbb{R}^{3N}} \prod_{r=1}^N \overline{\xi(x_r)}\cdot\Big(\sum_{i=1}^N \big(-\mu\Delta_{x_i} + U(x_i)\big) + \\
&\qquad\qquad + \sum_{1\leqslant i<j\leqslant N} V_N(x_i - x_j)\Big)\cdot\prod_{s=1}^N \xi(x_s)\,\mathrm{d}x_1\cdots\mathrm{d}x_N \\
&= \sum_{i=1}^N \int_{\mathbb{R}^3} \overline{\xi(x_i)}\big(-\mu\Delta_{x_i} + U(x_i)\big)\xi(x_i)\,\mathrm{d}x_i + \\
&\quad + \sum_{1\leqslant i<j\leqslant N} \int_{\mathbb{R}^3\times\mathbb{R}^3} \overline{\xi(x_i)}\;\overline{\xi(x_j)}\,V_N(x_i - x_j)\,\xi(x_i)\,\xi(x_j)\,\mathrm{d}x_i\,\mathrm{d}x_j \\
&= N\int_{\mathbb{R}^3} \big(\,\mu|\nabla\xi(x)|^2 + U(x)|\xi(x)|^2\big)\,\mathrm{d}x + \\
&\quad + \frac{(N-1)N}{2}\int_{\mathbb{R}^3\times\mathbb{R}^3} V_N(x-y)\,|\xi(x)|^2|\xi(y)|^2\mathrm{d}x\,\mathrm{d}y\,.
\end{aligned} \tag{6.41}$$

By (3.78), $NV_N \xrightarrow{N\to\infty} b\,\delta$ in the distributional sense, so

$$\frac{\left(\xi^{\otimes N},\, H_N\,\xi^{\otimes N}\right)}{N} \xrightarrow{N\to\infty} \int_{\mathbb{R}^3} \left(\mu|\nabla\xi(x)|^2 + U(x)|\xi(x)|^2 + \frac{b}{2}|\xi(x)|^4\right) \mathrm{d}x \\ = \mathcal{E}^{\mathrm{GP}}_{b/(8\pi\mu)}[\,\xi\,]\,. \tag{6.42}$$

By (6.38), $\mathcal{E}^{\mathrm{GP}}_{b/(8\pi\mu)}[\,\xi\,] \geqslant \mathcal{E}^{\mathrm{GP}}_{a}[\,\xi\,]$. Since φ^{GP} is the minimizer of $\mathcal{E}^{\mathrm{GP}}_{a} \equiv \mathcal{E}^{\mathrm{GP}}$, then $\mathcal{E}^{\mathrm{GP}}_{a}[\,\xi\,] \geqslant \mathcal{E}^{\mathrm{GP}}[\,\varphi^{\mathrm{GP}}\,] = E^{\mathrm{GP}}$. Then (6.40) is proved. It is also clear that, if $b/(8\pi\mu) > a$, then $\mathcal{E}^{\mathrm{GP}}_{b/(8\pi\mu)}[\,\xi\,] > \mathcal{E}^{\mathrm{GP}}_{a}[\,\xi\,]$. □

It deserves to be mentioned that trial functions of the form (6.4) and analogous, are present throughout the recent literature on rigorous B.E.C. and are all inspired by 1957 Dyson's pioneering work [37]. Before Dyson, attempts had been made by Dingle and Jastrow. A trial wave function of the form $\prod_{i<j} f(x_i - x_j)$ to calculate the N-body ground state energy had already been suggested by Mott for the hard sphere Bose gas and applied in the low density limit by Dingle [36]. Jastrow [61] then made a systematic use of it examining by a variational method the ground state of bosons or of fermions coupled with a strong interaction of short-range. Actually, the term $\prod_{i<j} f(x_i - x_j)$ in the trial wave function typically goes under the name of JASTROW FACTOR and it is the primary way for including electron correlations beyond the mean-field level, especially in quantum mechanical simulations.

While Jastrow's trial is symmetric under permutation of variables, Dyson's one – as well as (6.5) – is not. Also, the latter takes into account the interaction only between nearest neighbours, while in $\prod_{i<j} f(x_i - x_j)$ the product is taken over *all* pairs of particles. However, Jastrow was not able with his wave function to obtain a rigorous upper bound to the energy: his expression for the energy was a series of cluster integrals which he was obliged to cut off without controlling the error. On the other hand, Dyson's work showed first that the "nearest neighbour" type of wave function led to a new approach for the description of many-body systems with strong interactions, in which the typical difficulties associated with cluster-integrals expansions do not arise.

In the presence of the Jastrow-like factor, LSY trial function turns out to be asymptotically equal to a product function in norm in the Gross-Pitaevskiĭ limit. Such a convergence is *not* expected for the ground state function (nor it is expected for ESY trial function that we will be dealing with in Section 6.5, cointaining far more correlation factors). Thus, it is interesting that Ψ^{LSY}_N can "afford" to have less correlations that one might expect and nevertheless give the correct energy to the leading order.

6.3 Ground-state-like correlations lower the energy

The drop of $\langle H_N \rangle$ in the trial function Ψ^{LSY}_N with respect to the factorised $(\varphi^{\mathrm{GP}})^{\otimes N}$ is a typical phenomenon due only to the Dyson-like correlation factor (6.5), independently of the choice of the one-body condensate wave function. That is, the same happens also for many-body states with higher energy expectation value, not necessarily eigenstates of H_N, provided that they have the same correlation structure as Ψ^{LSY}_N.

This is a consequence of the same analysis by Lieb, Seiringer, and Yngvason for the upper bound of the ground state energy of a dilute Bose gas showing B.E.C. in the condensate wave function φ^{GP}. By direct inspection of the proof of theorem III.1 in [77], one can see that the same theorem asserts the following more general result.

Theorem 6.3.1. Let H_N be the Hamiltonian (3.77), i.e., the Hamiltonian of the three-dimensional model of N trapped and interacting bosons considered in theorem 3.5.2. Let $\varphi \in H^1(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ such that $\|\varphi\|_{L^2} = 1$ and let $\Psi_N^{(\varphi)}$ be the many-body wave function

$$\Psi_N^{(\varphi)} \;:=\; \frac{W_N(x_1, \dots, x_N)\, \varphi(x_1) \cdots \varphi(x_N)}{\big\| W_N(x_1, \dots, x_N)\, \varphi(x_1) \cdots \varphi(x_N) \big\|} \tag{6.43}$$

where W_N is the Dyson-like factor defined in (6.5). Let $\mathcal{E}^{\mathrm{GP}}$ be the Gross-Pitaevskiĭ energy functional (3.72). Then

$$\Big(\Psi_N^{(\varphi)}, \frac{H_N}{N} \Psi_N^{(\varphi)} \Big) \;\leqslant\; \mathcal{E}^{\mathrm{GP}}[\varphi] \left(1 + \mathcal{O}(N^{-2/3}) \right) \tag{6.44}$$

hence, $\mathcal{E}^{\mathrm{GP}}[\varphi]$ is an upper bound to the expectation value of the energy per particle on $\Psi_N^{(\varphi)}$.

Lieb et al. made the special choice $\varphi = \varphi^{\mathrm{GP}}$, although it holds in general. As a consequence, combining theorems 6.2.2 and 6.3.1, the expectation value of H_N in $\Psi_N^{(\varphi)}$ is eventually less than that in $\varphi^{\otimes N}$,

$$\Big(\Psi_N^{(\varphi)}, \frac{H_N}{N} \Psi_N^{(\varphi)} \Big) \;\leqslant\; \Big(\varphi^{\otimes N}, \frac{H_N}{N}\, \varphi^{\otimes N} \Big) , \tag{6.45}$$

and this is due to the Dyson-like factor W_N.

Proof of theorem 6.3.1. The scheme is the same as in the original proof of Lieb et al. To agree with the notation there, we let $\hbar = 2m = 1$ and rewrite the above $\Psi_N^{(\varphi)}$ as

$$\Psi_N^{(\varphi)} = \frac{W_N(x_1, \dots, x_N) \prod_{i=1}^N g(x_i)}{\big\| W_N(x_1, \dots, x_N) \prod_{i=1}^N g(x_i) \big\|} \tag{6.46}$$

with

$$g \;:=\; \frac{\varphi}{\| \varphi \|_{L^\infty}} . \tag{6.47}$$

Then one estimates $\big(\Psi_N^{(\varphi)}, \frac{H_N}{N} \Psi_N^{(\varphi)} \big)$ following exactly the same line of [77]. Direct computation produces a number of terms (Eq. (3.11) of [77]) which have to be suitably bounded. For the bound of the term in Eq. (3.19) of [77] a Cauchy-Schwarz inequality is needed, regardless of g (but provided that $g \in L^4(\mathbb{R}^3)$, as here). For the bound of the term in Eq. (3.21) of [77], one needs only $0 \leqslant |g(x)| \leqslant 1$, as it is in the present case. The only difference is that here φ is no longer real-valued as for the minimizer of the GP functional; after an innocent modification one gets the estimate

$$\begin{aligned} \Big(\Psi_N^{(\varphi)}, \frac{H_N}{N} \Psi_N^{(\varphi)} \Big) \leqslant\; & \frac{\tilde{e}_N \int_{\mathbb{R}^3} |g(x)|^2 \, \mathrm{d}x}{\int_{\mathbb{R}^3} |g(x)|^2 \, \mathrm{d}x - N I_N} + N \frac{J_N \int_{\mathbb{R}^3} |g(x)|^4 \, \mathrm{d}x}{\Big[\int_{\mathbb{R}^3} |g(x)|^2 \, \mathrm{d}x - N I_N \Big]^2} \\ & + \frac{\frac{2}{3} N^2 K_N^2}{\Big[\int_{\mathbb{R}^3} |g(x)|^2 - N I_N \Big]^2} \end{aligned} \tag{6.48}$$

which is the analog of Eq. (3.29) in [77], where constants $\tilde{e}_N$, I_N, J_N, and K_N are defined as in the original work, namely

$$I_N := \int_{\mathbb{R}^3} \big(1 - f_N^2(x)\big)\,\mathrm{d}x\,, \tag{6.49}$$

$$J_N := \frac{1}{2}\int_{\mathbb{R}^3} \big(2\,(f_N'(x))^2 + V_N(x) f_N^2(x)\big)\,\mathrm{d}x\,, \tag{6.50}$$

$$K_N := \int_{\mathbb{R}^3} f_N(x) f_N'(x)\,\mathrm{d}x\,, \tag{6.51}$$

$$\tilde{e}_N := \frac{\displaystyle\int_{\mathbb{R}^3} \big[\, -\overline{g(x)}\,\Delta g(x) + U(x)|g(x)|^2\,\big]\,\mathrm{d}x}{\displaystyle\int_{\mathbb{R}^3} |g(x)|^2\,\mathrm{d}x}\,. \tag{6.52}$$

Now recall from (6.8) that the cut-off parameter used by Lieb et al. in the definition of the Dyson-like factor W_N is $b_N \sim N^{-1/3}$: following the same explicit choice as in Eq. (3.39) of [77], let us take b_N such that

$$\frac{4\pi}{3}\overline{\rho}_N b_N^3 = \frac{\displaystyle\int_{\mathbb{R}^3} |g(x)|^4\,\mathrm{d}x}{\displaystyle\int_{\mathbb{R}^3} |g(x)|^2\,\mathrm{d}x} = \frac{\displaystyle\int_{\mathbb{R}^3} |\varphi(x)|^4\,\mathrm{d}x}{\|\,\varphi\,\|_{L^\infty}^2} =: c\,, \tag{6.53}$$

where, in analogy with the original case,

$$\overline{\rho}_N := N\int_{\mathbb{R}^3} |\varphi(x)|^4\,\mathrm{d}x\,. \tag{6.54}$$

Indeed (6.53) and (6.54) give

$$b_N = \Big(\frac{4\pi}{3} N \|\,\varphi\,\|_{L^\infty}^2\Big)^{-1/3} \sim\ N^{-1/3}\,. \tag{6.55}$$

In terms of this b_N and of the scattering length a_N

$$I_N \leqslant 4\pi\Big(\frac{a_N^3}{3} + a_N b_N (b_N - a_N)\Big) \tag{6.56}$$

as in Eq. (3.34) of [77]. This, in turn, leads to

$$\begin{aligned}
\int_{\mathbb{R}^3} |g(x)|^2\,\mathrm{d}x - N I_N &\geqslant \int_{\mathbb{R}^3} |g(x)|^2\Big(1 - \frac{N I_N}{\int_{\mathbb{R}^3}|g(x)|^2\,\mathrm{d}x}\Big)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^3} |g(x)|^2\Big(1 - \frac{4\pi}{3} N a_N^3 \|\,\varphi\,\|_{L^\infty}^2 + \\
&\qquad\qquad - 4\pi N a_N b_N (b_N - a_N)\|\,\varphi\,\|_{L^\infty}^2\Big)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^3} |g(x)|^2\Big(1 - \frac{a_N^3}{b_N^3} - \frac{3\,a_N b_N (b_N - a_N)}{b_N}\Big)\,\mathrm{d}x \\
&\geqslant \int_{\mathbb{R}^3} |g(x)|^2\Big(1 - \frac{a_N}{b_N}\Big)^3\,\mathrm{d}x
\end{aligned} \tag{6.57}$$

as Eq. (3.40) of of [77]. In the r.h.s. of (6.57) we have used

$$\|\,\varphi\,\|_{L^\infty}^2 = \Big(\frac{4\pi}{3} N b_N^3\Big)^{-1} \tag{6.58}$$

which follows from (6.55).

This way, following [77], one has all the estimates to deduce from (6.48) the desired bound, in analogy to Eq. (3.42) of [77]:

$$\begin{aligned}\Big(\Psi_N^{(\varphi)},\frac{H_N}{N}\Psi_N^{(\varphi)}\Big) &\leqslant \frac{\int_{\mathbb{R}^3}\big[\,|\nabla\varphi(x)|^2+U(x)|\varphi(x)|^2\,\big]\,\mathrm{d}x}{\Big(1-\frac{a_N}{b_N}\Big)^3}+\\ &+4\pi a\int_{\mathbb{R}^3}|\varphi(x)|^4\frac{1+\frac{2}{c}\Big(\frac{a_N}{b_N}\Big)-\frac{2}{c}\Big(\frac{a_N}{b_N}\Big)^2+\frac{1}{2c}\Big(\frac{a_N}{b_N}\Big)^3}{\Big(1-\frac{a_N}{b_N}\Big)^8}\,\mathrm{d}x\\ &=\big(1+\mathcal{O}(N^{-2/3})\big)\int_{\mathbb{R}^3}\big[\,|\nabla\varphi(x)|^2+U(x)|\varphi(x)|^2+4\pi a|\varphi(x)|^4\,\big]\,\mathrm{d}x\end{aligned}\tag{6.59}$$

where a is the unscaled scattering length. □

6.4 Correlations at the level of reduced density matrices

We go back to the difficulties ($\mathbb{D}_1$), ($\mathbb{D}_2$), and ($\mathbb{D}_3$) one has in characterizing interparticle correlations of the many-body condensed state through the analysis at the level of k-particle marginals.

Correlations in the N-body wave function show up as modifications of the wave function with respect to a purely factorised form $\prod\varphi(x_i)$ – see, e.g., (6.4). Hence, correlations at the level of marginals are modifications of purely factorised kernels $\prod\varphi(x_i)\overline{\varphi(y_i)}$.

When Ψ_N is thought as a point in an abstract $\mathcal{H}_{N,\mathrm{sym}}$, correlations consist in writing Ψ_N as some non-trivial expansion like (3.22) in a symmetric orthonormal basis – constructed as in theorem 3.1.3 – instead of simply as $\varphi^{\otimes N}$. At the level of marginals one then detects them in the form of linear combinations of distinct tensor power projections $|\varphi\rangle\langle\varphi|^{\otimes k}$, $|\xi\rangle\langle\xi|^{\otimes k}$, …, of mixed projection terms $|\varphi\rangle\langle\varphi|\otimes|\xi\rangle\langle\xi|\otimes\cdots$, and of mixed non-projection terms $|\varphi\rangle\langle\xi|\otimes\cdots$. Such a mechanism, in the particular case of N-body pure states with definite quantum labels, goes as follows.

Theorem 6.4.1. Let $\gamma_N=|\Psi_N\rangle\langle\Psi_N|$ be the N-body ($N\geqslant 2$) pure state with definite quantum labels

$$\Psi_N=|\varphi_{i_1},n_1;\varphi_{i_2},n_2;\ldots;\varphi_{i_r},n_r\rangle_{\mathrm{sym}}\,.\tag{6.60}$$

Then the corresponding two-particle reduced density matrix is

$$\gamma_N^{(2)}=\sum_{k,h=1}^{r}\frac{n_k}{N}\,\frac{n_h-\delta_{k,h}}{N-1}\,|\varphi_{i_k}\rangle\langle\varphi_{i_k}|\otimes|\varphi_{i_h}\rangle\langle\varphi_{i_h}|\tag{6.61}$$

($\delta_{k,h}$ being the Kronecker delta).

Proof. As in the proof of theorem 3.3.1, when constructing $\gamma_N=|\Psi_N\rangle\langle\Psi_N|$ one gets diagonal terms (3.30) and off-diagonal terms (3.31) and, by the same argument, the latter do not contribute when $\mathrm{Tr}_{[N-2]}$ is taken (the partial trace over the last $N-2$

positions). Instead, partial trace of each diagonal term gives

$$\begin{aligned}
&\mathrm{Tr}_{[N-2]}\left[\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\otimes\varphi_{j_2}\otimes\cdots\otimes\varphi_{j_N}\rangle\langle\varphi_{j_1}\otimes\varphi_{j_2}\otimes\cdots\otimes\varphi_{j_N}|\right]\\
&\qquad=\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\rangle\langle\varphi_{j_1}|\otimes|\varphi_{j_2}\rangle\langle\varphi_{j_2}|\prod_{k=3}^{N}\int_\Omega|\varphi_{j_k}(x_k)|^2\,\mathrm{d}x_k \qquad (6.62)\\
&\qquad=\frac{n_1!\,n_2!\cdots n_r!}{N!}\,|\varphi_{j_1}\rangle\langle\varphi_{j_1}|\otimes|\varphi_{j_2}\rangle\langle\varphi_{j_2}|\,.
\end{aligned}$$

Among all terms of the linear combination (3.20), there is a fraction n_k/N having φ_{i_k} in first position; among these ones (the "tail" of which has $N-1$ positions), there is a fraction $n_h/(N-1)$ having φ_{i_h} in second position, if $h\neq k$, and a fraction $(n_k-1)/(N-1)$ having φ_{i_k} also in second position, whence (6.61). □

Thus, $\gamma_N^{(2)}$ accounts not only for tensor power projections $|\varphi_{i_k}\rangle\langle\varphi_{i_k}|^{\otimes 2}$ as for a totally factorised state $\Psi_N=\varphi^{\otimes N}$, but also for mixed projection terms $|\varphi_{i_k}\rangle\langle\varphi_{i_k}|\otimes|\varphi_{i_h}\rangle\langle\varphi_{i_h}|$, $k\neq h$, which are a signature of correlations among particles. Notice that all the $n_{i_1},\ldots,n_{i_r}$ are positive integers, so terms like $|\varphi_{i_k}\rangle\langle\varphi_{i_k}|^{\otimes 2}$ enter (3.29) only when $n_k\geqslant 2$. If Ψ_N has definite quantum labels, only distinct tensor power projections and mixed projection terms appear in its marginals, without mixed non-projection terms. If Ψ_N is in turn a true linear combination, like (6.4), then also non-projection terms may appear in the marginals.

For example, if

$$\begin{aligned}
\Psi_7&=|\varphi,4;\xi,2;\eta,1\rangle_{\mathrm{sym}}\\
&=\sqrt{\frac{30}{105}}\,(\varphi\otimes\varphi\otimes\cdots)+\sqrt{\frac{20}{105}}\,(\varphi\otimes\xi\otimes\cdots)+\sqrt{\frac{10}{105}}\,(\varphi\otimes\eta\otimes\cdots)\\
&+\sqrt{\frac{20}{105}}\,(\xi\otimes\varphi\otimes\cdots)+\sqrt{\frac{5}{105}}\,(\xi\otimes\xi\otimes\cdots)+\sqrt{\frac{5}{105}}\,(\xi\otimes\eta\otimes\cdots)\\
&+\sqrt{\frac{10}{105}}\,(\eta\otimes\varphi\otimes\cdots)+\sqrt{\frac{5}{105}}\,(\eta\otimes\xi\otimes\cdots)
\end{aligned}$$

then one can easily find

$$\begin{aligned}
\gamma_7^{(2)}&=\frac{2}{7}|\varphi\rangle\langle\varphi|^{\otimes 2}+\frac{1}{21}|\xi\rangle\langle\xi|^{\otimes 2}\\
&+\frac{4}{21}|\varphi\rangle\langle\varphi|\otimes|\xi\rangle\langle\xi|+\frac{2}{21}|\varphi\rangle\langle\varphi|\otimes|\eta\rangle\langle\eta|+\frac{4}{21}|\xi\rangle\langle\xi|\otimes|\varphi\rangle\langle\varphi|\\
&+\frac{1}{21}|\xi\rangle\langle\xi|\otimes|\eta\rangle\langle\eta|+\frac{2}{21}|\eta\rangle\langle\eta|\otimes|\varphi\rangle\langle\varphi|+\frac{1}{21}|\eta\rangle\langle\eta|\otimes|\xi\rangle\langle\xi|
\end{aligned}$$

and

$$\gamma_7^{(1)}=\frac{4}{7}|\varphi\rangle\langle\varphi|+\frac{2}{7}|\xi\rangle\langle\xi|+\frac{1}{7}|\eta\rangle\langle\eta|\,.$$

On the other hand, as in the example before definition 3.3.2, if

$$\Psi_8=\frac{1}{\sqrt{6}}\big(\,|\varphi,6;\xi,2\rangle_{\mathrm{sym}}+|\varphi,5;\xi,2;\eta,1\rangle_{\mathrm{sym}}+2\,|\varphi,4;\eta,4\rangle_{\mathrm{sym}}\,\big)$$

then

$$\gamma_8^{(1)}=\frac{27}{48}|\varphi\rangle\langle\varphi|+\frac{4}{48}|\xi\rangle\langle\xi|+\frac{17}{48}|\eta\rangle\langle\eta|+\frac{1}{8\sqrt{6}}|\varphi\rangle\langle\eta|+\frac{1}{8\sqrt{6}}|\eta\rangle\langle\varphi|\,.$$

In the asymptotic analysis ($N \to \infty$) and in presence of complete condensation, correlations disappear in the one-body density matrix: $\gamma_N^{(1)} \to \gamma_\infty^{(1)} = |\varphi\rangle\langle\varphi|$. (As we shall see in Sec. 7.1, the same necessarily happens in any k-body density matrix: $\gamma_N^{(k)} \to \gamma_\infty^{(k)} = |\varphi\rangle\langle\varphi|^{\otimes k}$; by theorem 6.4.1, this is already clear in the special case of pure states with definite quantum labels.) Further, their contribution to k-body marginals before the limit may be essentially negligible for many purposes: for example – see (6.3) – the one-body marginal relative to $\big(\varphi^{\otimes(N-1)} \otimes \xi\big)_{\mathrm{sym}}$ differs in trace norm from its asymptotics $|\varphi\rangle\langle\varphi|$ by an error $\mathcal{O}(N^{-1})$.

Yet, *an accurate knowledge of correlations at the level of reduced density matrices* before *the limit turns out to be crucial in many aspects of the rigorous analysis of B.E.C.*

As an example of that, let us give here details of how an appropriate ansatz on the correlation structure in the marginals determines that formal limit in the finite BBGKY hierarchy which leads to the infinite hierarchy with the correct coupling (namely, the two-body process scattering length, as already mentioned when sketching the scheme of the "methods of the hierarchies" for the derivation of the GP equation, Sec. 5.4).

Let us recall that the first equation of the finite BBGKY hierarchy for the usual model of N undistinguishable bosons expanding under the pairwise repulsion $V_N(x) = N^2 V(Nx)$:

$$\begin{aligned} i\hbar\,\partial_t \gamma_{N,t}^{(1)}(x_1, y_1) = \frac{\hbar^2}{2m}(-\Delta_{x_1} + \Delta_{y_1})\,\gamma_{N,t}^{(1)}(x_1, y_1) + \\ + (N-1)\int_{\mathbb{R}^3} \big[V_N(x_1 - z) - V_N(y_1 - z)\big]\,\gamma_{N,t}^{(2)}(x_1, z, y_1, z)\,\mathrm{d}z\,. \end{aligned} \tag{6.63}$$

The "no-correlation ansatz" $\gamma_{N,t}^{(2)} = \gamma_{N,t}^{(1)} \otimes \gamma_{N,t}^{(1)}$, i.e.,

$$\gamma_{N,t}^{(2)}(x_1, x_2, y_1, y_2) = \gamma_{N,t}^{(1)}(x_1, y_1)\,\gamma_{N,t}^{(1)}(x_2, y_2)\,, \tag{6.64}$$

when plugged into (6.63), gives in the formal limit $N \to \infty$ the following first equation of the infinite BBGKY hierarchy:

$$i\hbar\,\partial_t \gamma_{\infty,t}^{(1)}(x_1, y_1) = \frac{\hbar^2}{2m}(-\Delta_{x_1} + \Delta_{y_1})\,\gamma_{\infty,t}^{(1)}(x_1, y_1) + \big(Q_t(x_1) - Q_t(y_1)\big)\gamma_{\infty,t}^{(1)}(x_1, y_1)\,. \tag{6.65}$$

In (6.65) one has set

$$Q_t(x) \;:=\; \lim_{N\to\infty} N \int_{\mathbb{R}^3} V_N(x - y)\,\gamma_{N,t}^{(1)}(y, y) \;=\; b\,\gamma_{\infty,t}^{(1)}(x, x)\,, \tag{6.66}$$

last equality following if $x \mapsto \gamma_{\infty,t}^{(1)}(x, x)$ is continuous. As usual, a is the s-wave scattering length of V and $\frac{1}{8\pi\mu} b = \frac{1}{8\pi\mu}\int_{\mathbb{R}^3} V(x)\,\mathrm{d}x$ is its first Born approximation. This way, (6.65) turns out to be solved by $\gamma_{\infty,t}^{(1)}(x, y) = \varphi_t(x)\overline{\varphi_t(y)}$ where $\varphi_t(x)$ is the solution of

$$i\hbar\,\partial_t \varphi_t(x) = -\mu\,\Delta_x \varphi_t(x) + b\,|\varphi_t(x)|^2 \varphi_t(x)\,. \tag{6.67}$$

To recover the expected coupling in front of the cubic nonlinearity of (6.67), one instead makes the ansatz

$$\gamma_{N,t}^{(2)}(x_1, x_2, y_1, y_2) = f_N(x_1 - x_2) f_N(y_1 - y_2)\,\gamma_{N,t}^{(1)}(x_1, y_1)\,\gamma_{N,t}^{(1)}(x_2, y_2) \tag{6.68}$$

modifying the pure product form (6.64) with correlations expressed in terms of the zero energy scattering solution f_N defined in (6.10). This, together with (6.14) and the assumption that $x \mapsto \gamma^{(1)}_{\infty,t}(x,x)$ is smooth on scale N^{-1}, in the formal limit $N \to \infty$ gives

$$\begin{aligned}\lim_{N\to\infty} N \int_{\mathbb{R}^3} \big[V_N(x_1 - z) - V_N(y_1 - z)\big]\, \gamma^{(2)}_{N,t}(x_1, z, y_1, z)\, \mathrm{d}z \\ = 8\pi\mu a \left(\gamma^{(1)}_{\infty,t}(x_1,x_1) - \gamma^{(1)}_{\infty,t}(y_1,y_1)\right)\gamma^{(1)}_{\infty,t}(x_1,y_1)\end{aligned} \tag{6.69}$$

for $|x_1 - y_1| \gg N^{-1}$, thus recovering the GP equation with the correct dependence on the scattering length.

6.5 Trial function for the evolution of a condensate

Here the setting changes to that of Sec. 5.3. From the time-dependent analysis of Erdős, Schlein and Yau, a reasonable trial function emerges for the evolution of the originally trapped condensate, expanding under some pairwise repulsion. Let us denote it by $\Psi^{\mathrm{ESY}}_{N,t}$, to distinguish it from Ψ^{LSY}_N considered by Lieb et al. for the static case. In this section we wish to present the motivation for introducing such a $\Psi^{\mathrm{ESY}}_{N,t}$ and in which sense it can/cannot be considered close to the true evolution $\Psi_{N,t}$.

Actually, while Lieb et al. explicitly introduce their trial function, it has to be remarked that throughout the derivation of the ESY theorem 5.3.1 one does not deal with any trial function at all, all the proof being based with a careful treatment of the true marginals extracted from the unknown $\Psi_{N,t}$. Yet, an underlying $\Psi^{\mathrm{ESY}}_{N,t}$ is clearly present.

What motivates the choice of the trial function in this setting is the following key property.

Theorem 6.5.1 (H^2_N-energy estimate, [44]). Assume that H_N is the Hamiltonian of the model considered in theorem 5.3.1. (In particular, α is the small enough parameter defined in (5.18) to measure the strength of the interaction.) Let f_N be the zero-energy scattering solution defined in (6.10). Then, there exists a universal constant $c > 0$ such that, for every fixed indices $i, j \in \{1, \dots, N\}$ with $i \neq j$ and $\forall\, \Psi_N \in L^2_{\mathrm{sym}}(\mathbb{R}^{3N})$,

$$\big(\Psi_N, H^2_N \Psi_N\big) \geqslant (1 - c\,\alpha) N(N-1) \int_{\mathbb{R}^{3N}} f^2_N(x_i - x_j) \Big| \nabla_{x_i} \nabla_{x_j} \frac{\Psi_N}{f_N(x_i - x_j)} \Big|^2 . \tag{6.70}$$

Such an estimate involves a many-body function constructed by "extracting" from Ψ_N the correlation $f_N(x_i - x_j)$ at the short scale $\sim N^{-1}$. It says that, if $\Psi_N / f_N(x_i - x_j)$ is singular when $|x_i - x_j| \to 0$, then $\nabla_{x_i}\nabla_{x_j}(\Psi_N / f_N(x_i - x_j))$ cannot be L^2-integrable. Now, recall that in ESY analysis, one assumes that at the initial time and, consequently, at all later times the true many-body wave function satisfies the energy bound

$$\big(\Psi_{N,t}, H^2_N \Psi_{N,t}\big) \leqslant \mathrm{const} \cdot N^2 . \tag{6.71}$$

Then the short distance behaviour of $\Psi_{N,t}$ is given by $f_N(x_i - x_j)$ when $x_i \sim x_j$ and this holds *at any time*: there is a persistence of such a short scale structure during the time evolution.

There is actually more. Indeed ESY theorem accounts, among others, for purely factorised initial data $\varphi^{\otimes N}$, where any correlation is absent. At later times, the many-body state $e^{-iH_N t/\hbar}\varphi^{\otimes N}$ must contain the above short scale structure. Thus, the many-body dynamics *builds up* and *preserves* short-scale correlations emerging in the two-body process (see [43, 44]).

On this basis one introduces

$$\Psi^{\mathrm{ESY}}_{N,t}(x_1,\dots,x_N) \;:=\; \frac{\mathcal{W}_N(x_1,\dots,x_N)\prod_{i=1}^{N}\varphi_t(x_i)}{\Big\|\,\mathcal{W}_N(x_1,\dots,x_N)\prod_{i=1}^{N}\varphi_t(x_i)\,\Big\|} \tag{6.72}$$

where interparticle correlations are encoded in the Jastrow-like factor

$$\mathcal{W}_N(x_1,\dots,x_N) \;:=\; \prod_{1\leqslant i<j\leqslant N} f_N(x_i-x_j) \tag{6.73}$$

and f_N is the zero-energy scattering solution for the interaction V_N. Here φ_t is the evolution of a given φ under the time-dependent GP equation, where φ is assumed to be the one-body condensate wave function at time $t=0$. Instead, $\mathcal{W}_N$ does not depend on time: such a correlation structure is preserved during time evolution.

Currently, an explicit control that $\Psi^{\mathrm{ESY}}_{N,t}\approx\Psi_{N,t}$ in norm, or in energy form, is lacking. Instead, $\Psi^{\mathrm{ESY}}_{N,t}$ is believed to be close to the true $\Psi_{N,t}$ in a much more indirect sense:

① $\Psi^{\mathrm{ESY}}_{N,t}$ has the same short scale structure that $\Psi_{N,t}$ must have, under the hypothesis of boundedness of the square of the energy per particle;

② the corresponding marginals, $\{\gamma^{(k),\mathrm{ESY}}_{N,t}\}_{k=1}^{N}$, have the right correlations which enable one to close the infinite hierarchy in the formal limit $N\to\infty$ in such a way to recover the expected dependence on the scattering length;

③ estimates hold for the marginals such that each sequence $\{\gamma^{(k),\mathrm{ESY}}_{N,t}\}_{N\geqslant k}$ (at any fixed k) is weakly-$*$ compact in the trace class and any weak-$*$ limit turns out to be the unique solution of the infinite GP hierarchy.

Point ① is the consequence of the H^2_N-energy estimate, as already noticed. Point ② corresponds to the discussion concluding the previous section. Indeed one has

$$\begin{aligned}\gamma^{(2)}_{N,t}(x_1,x_2,y_1,y_2) = f_N(x_1-x_2)\,f_N(y_1-y_2)\,\varphi_t(x_1)\,\varphi_t(x_2)\,\overline{\varphi_t(y_1)}\;\overline{\varphi_t(y_2)}\cdot\\ \cdot\,\mathcal{A}^{(2)}_{N,2}(x_1,x_2,y_1,y_2)\end{aligned} \tag{6.74}$$

with

$$\begin{aligned}\mathcal{A}^{(2)}_{N,2}(x_1,x_2,y_1,y_2) := \int_{\mathbb{R}^{3(N-2)}}\prod_{j=3}^{N} f_N(x_1-z_j)f_N(x_2-z_j)f_N(y_1-z_j)f_N(y_2-z_j)\cdot\\ \cdot\prod_{3\leqslant r<s\leqslant N} f^2_N(z_r-z_s)\,|\varphi_t(z_3)|^2\cdots|\varphi_t(z_N)|^2\,\mathrm{d}z_3\cdots\mathrm{d}z_N\,.\end{aligned} \tag{6.75}$$

Therefore, since $\mathcal{A}^{(2)}_{N,2}\approx 1$, the two-body trial marginal (6.74) has essentially the short scale structure as in (6.68), which in turn leads to the correct GP hierarchy in the formal limit $N\to\infty$.

Let us comment on point ③. In ESY theorem, one needs two fundamental bounds for completing the "convergence+uniqueness" task of the method of the hierarchies. First, the technique to prove the trace class weak-$*$ compactness for each sequence of *true* k-marginals $\{\gamma^{(k)}_{N,t}\}_{N\geqslant k}$, is an Ascoli-Arzelà argument, for which the equicontinuity of each $t\mapsto\gamma^{(k)}_{N,t}$ is needed, uniformly in N. This, in turn, is proved to be guaranteed by

$$\mathrm{Tr}\big[\,(1-\Delta_{x_i})(1-\Delta_{x_j})\,\gamma^{(k)}_{N,t;(i,j)}\,\big]\leqslant\mathrm{const} \tag{6.76}$$

uniformly in t and eventually in N, and $\forall\, i,j\in\{1,\dots,k\}$ with $i<j$, where

$$\gamma^{(k)}_{N,t;(i,j)}(x_1,\dots,x_k,y_1,\dots,y_k)\;:=\;\frac{\gamma^{(2)}_{N,t}(x_1,\dots,x_k,y_1,\dots,y_k)}{f_N(x_i-x_j)f_N(y_i-y_j)}\,. \tag{6.77}$$

The crucial bound (6.76) follows from the H^2_N-energy estimate above. This is a non-trivial technical task: indeed, for finite N, marginals $\gamma^{(k)}_{N,t}$'s do *not* satisfy estimates such as (6.76), at least not uniformly in N, because they contain a short scale structure: only after removal of this structure, as in (6.77), the bound holds. It is also remarkable that the same bound for $\gamma^{(k)}_{N,t}$, instead of $\gamma^{(k)}_{N,t;(i,j)}$, is recovered in the limit $N\to\infty$, where the correlation structure disappears: indeed it is another key achievement along the proof of ESY theorem that

$$\mathrm{Tr}\big[\,(1-\Delta_{x_1})\cdots(1-\Delta_{x_k})\,\gamma^{(k)}_{\infty,t}\,\big]\;\leqslant\;C^k \tag{6.78}$$

$\forall\, k\geqslant 1$, $\forall\, t\in[0,T]$, where C and T are some positive constants. Furthermore, it has to be emphasized here that uniqueness of the solution of the infinite GP hierarchy holds provided that the solution $\{\gamma^{(k)}_t\}_{k\geqslant1}$ satisfies (6.78). Hence, summarizing, *the "a priori bounds"* (6.76) *and* (6.78) *are the essential ingredients for the existence of the limiting marginals and for their uniqueness, respectively.*

The content of point ③ above is thus clear. For example, in the concrete case $k=2$, since essentially one has

$$\gamma^{(2),\mathrm{ESY}}_{N,t}\;\approx\;f_N(x_1-x_2)\,f_N(y_1-y_2)\,\varphi_t(x_1)\,\varphi_t(x_2)\,\overline{\varphi_t(y_1)}\,\overline{\varphi_t(y_2)}\,, \tag{6.79}$$

then

$$\gamma^{(2),\mathrm{ESY}}_{N,t;(1,2)}\;\approx\;\varphi_t(x_1)\,\varphi_t(x_2)\,\overline{\varphi_t(y_1)}\,\overline{\varphi_t(y_2)}\,, \tag{6.80}$$

whence

$$\begin{aligned}\mathrm{Tr}\big[\,(1-\Delta_{x_1})(1-\Delta_{x_2})\,\gamma^{(2),\mathrm{ESY}}_{N,t;(1,2)}\,\big]\;&\approx\\ &\approx\;\big(\,\varphi_t(x_1)\varphi_t(x_2)\,,\,(1-\Delta_{x_1})(1-\Delta_{x_2})\,\varphi_t(x_1)\varphi_t(x_2)\,\big)\\ &=\;\Big(\int_{\mathbb{R}^3}\overline{\varphi_t(x)}\,(1-\Delta_x)\,\varphi_t(x)\,\mathrm{d}x\Big)^2\\ &=\;\|\varphi_t\|^4_{H^1}\end{aligned} \tag{6.81}$$

which is uniformly bounded in time. Then, bound (6.76) is satisfied and $\{\gamma^{(1),\mathrm{ESY}}_{N,t}\}_{N\geqslant1}$ admits trace class weak-$*$ limits, *as for the case of the true marginals.* The same holds at higher k's and the limit is recognised to be the sequence $\{\gamma^{(k),\mathrm{ESY}}_{\infty,t}=|\varphi_t\rangle\langle\varphi_t|^{\otimes k}\}_{k\geqslant1}$. Notice that each $\gamma^{(k),\mathrm{ESY}}_{\infty,t}$ satisfies the bound (6.78) with $C=\|\varphi_t\|^2_{H^1}$.

6.6 Correlations and energy discrepancies in the evolution of the condensate

Despite ESY trial function $\Psi^{\mathrm{ESY}}_{N,t}$ mimes the true many-body state $\Psi_{N,t}$ in the sense that it contains the correct correlation structure at the scale N^{-1}, it is a currently major open problem to understand how good the approximation is in norm and in energy form. Work in progress [82] is dealing with this issue and this section gives a brief account of that.

Concerning the norm distance $\|\Psi_{N,t} - \Psi^{\mathrm{ESY}}_{N,t}\|$, estimates as in theorem 6.2.1 fail because $\Psi^{\mathrm{ESY}}_{N,t}$ contains a far larger number of correlation terms than $\Psi^{\mathrm{LSY}}_{N,t}$ (respectively, $\sim N^2$ versus $\sim N$). A more refined estimate by means of the properties of f_N seems promising.

Concerning the difference $(\Psi_{N,t}, H_N\Psi_{N,t}) - (\Psi^{\mathrm{ESY}}_{N,t}, H_N\Psi^{\mathrm{ESY}}_{N,t})$, it is clear that the first term is conserved in time, because $\Psi_{N,t}$ is the evolution of the initial datum $\Psi_{N,0}$ at $t = 0$ under the dynamics generated by the Hamiltonian H_N of the model considered in theorem 5.3.1. On the other hand the second term fluctuates, since ESY trial is not an eigenstate of H_N. Although an exact asymptotic computation of $(\Psi^{\mathrm{ESY}}_{N,t}, H_N\Psi^{\mathrm{ESY}}_{N,t})$ is still missing (while one can bound this expectation value with some bound proportional to N, through arguments similar to those of lemma B.1 in [44]), it is believed on the basis of the corresponding one-body picture that

$$\Big(\Psi^{\mathrm{ESY}}_{N,t}, \frac{H_N}{N}\Psi^{\mathrm{ESY}}_{N,t}\Big) \sim \mathcal{E}_a[\varphi_t] = \mathcal{E}_a[\varphi] \tag{6.82}$$

where, as usual, φ_t solves the time-dependent GP equation

$$i\hbar\partial_t u_t = -\mu\Delta_x u_t + 8\pi\mu a\,|u_t|^2 u_t = \frac{\delta\mathcal{E}_a[u]}{\delta\overline{u}} \tag{6.83}$$

with initial datum φ and

$$\mathcal{E}_a[u] \;:=\; \int_{\mathbb{R}^3} \big(\,\mu|\nabla u(x)|^2 + 4\pi\mu\, a|u(x)|^4\,\big)\,\mathrm{d}x \tag{6.84}$$

is the corresponding GP energy functional without trapping (compare with (5.14) and (5.15)).

Thus, one understands that $\Psi^{\mathrm{ESY}}_{N,t}$ cannot account correctly for the expectation value of H_N in the state $\Psi_{N,t}$, because of the possibility of choosing distinct initial data $\Psi_{N,0}$ with distinct energies. For instance, starting from an initial datum $\Psi_{N,0} = \varphi^{\otimes N}$ (which is admissible in ESY theorem), the expectation value of the energy in the state $\Psi_{N,t} = e^{-iH_Nt/\hbar}\,\varphi^{\otimes N}$ is

$$\Big(\Psi_{N,t}, \frac{H_N}{N}\Psi_{N,t}\Big) = \Big(\varphi^{\otimes N}, \frac{H_N}{N}\varphi^{\otimes N}\Big) \xrightarrow{N\to\infty} \mathcal{E}_{b/(8\pi\mu)}[\varphi] \tag{6.85}$$

due to the same computation as in the proof of theorem 6.2.2. Hence, for this choice of the initial datum, ESY trial function underestimates $\langle H_N\rangle$.

Such an apparent discrepancy is reasonable because ground-state-like correlations lower the expectation value of the energy, though it does not rely on those firm grounds that the *explicit* computation of $(\Psi^{\mathrm{ESY}}_{N,t}, H_N\Psi^{\mathrm{ESY}}_{N,t})$ would give.

In fact, the net effect of the presence of a Jastrow-like correlation factor $\mathcal{W}_N$ in $\Psi^{\mathrm{ESY}}_{N,t}$, when compared to the initial datum $\varphi^{\otimes N}$, is to lower the expectation value of

the potential energy, because correlations reduce the probability that two particles stay close together and the potential is repulsive, and to increase the expectation value of the kinetic energy, because Laplacians act to f_N-factors which are more and more singular in the origin, as $N \to \infty$. Yet, a detailed balance between these two opposite discrepancies is not clear and it is of interest to have a full control of the energy loss

$$\Delta E_{N,t} \;:=\; \Big(\Psi_{N,t}, \frac{H_N}{N} \Psi_{N,t} \Big) - \Big(\Psi^{\mathrm{ESY}}_{N,t}, \frac{H_N}{N} \Psi^{\mathrm{ESY}}_{N,t} \Big) \tag{6.86}$$

and to understand whether $\Delta E_{N,t} > 0$.

If this turned out to be the case, one should conclude that $\Psi^{\mathrm{ESY}}_{N,t}$ does not capture the leading term in the energy of $\Psi_{N,t}$. Erdős, Schlein, and Yau's tentative explanation (see [43, 44]) is a picture where some energy exists on intermediate length scales of order $N^{-\beta}$, $0 < \beta < 1$. This excess energy apparently does not participate in the evolution of marginals on the macroscopic length scale.

The suspect is that one could recover the correct energy asymptotics by introducing a modified trial function, say $\widetilde{\Psi}^{\mathrm{ESY}}_{N,t}$, with higher energy than $\Psi^{\mathrm{ESY}}_{N,t}$: in $\widetilde{\Psi}^{\mathrm{ESY}}_{N,t}$ one makes the replacement

$$f_N(x_i - x_j) \;\longmapsto\; f_N(x_i - x_j) + \varepsilon_N \, g_N(x_i - x_j) \tag{6.87}$$

with a bounded $g_N(x)$ oscillating around zero and becoming more and more singular for large N, thus increasing the kinetic part of the Hamiltonian, while $\varepsilon_N \xrightarrow{N\to\infty} 0$ in such a way to guarantee that the modification of the potential part is negligible.

Summarizing, in this scenario for the evolution of a Bose gas initially prepared in some state $\varphi^{\otimes N}$ one may consider the following many-body wave functions:

- ✓ the true $\Psi_{N,t} = e^{-iH_N t/\hbar} \varphi^{\otimes N}$ where the expectation value of the energy per particle is asymptotically $\mathcal{E}_{b/(8\pi\mu)}[\varphi]$; $\Psi_{N,t}$ has the characteristic short scale correlations imposed by the H_N^2-energy estimate and possibly certain incoherent excitations living at an intermediate scale;

- ✓ $\varphi_t^{\otimes N}$, in which $\langle H_N \rangle$ is asymptotically the same as in $\Psi_{N,t}$, but does not solve the N-body Schrödinger equation (in fact, correlations are absent);

- ✓ $\Psi^{\mathrm{ESY}}_{N,t}$, a trial function which reproduces the correct correlations at the scale N^{-1} and matches with $\Psi_{N,t}$ at the level of marginals, but has reasonably a lower expectation value of the energy that the true $\Psi_{N,t}$;

- ✓ some modified $\widetilde{\Psi}^{\mathrm{ESY}}_{N,t}$ to be constructed with criteria (see above) guaranteeing its closeness to $\Psi_{N,t}$ in energy form and in norm.

Chapter 7

Equivalent characterizations of B.E.C.

7.1 B.E.C. of infinite systems at any k-particle level

Technical and physical reasons motivate to control condensation not only at the one-body marginal level, but also through *any* k-particle reduced density matrix. In this section, following Ref. [83], we give account of the substantial equivalence of these controls.

From the physical point of view, controlling *one*-particle marginals only does not give full access to interparticle correlations, the knowledge of which turns out to be crucial in determining, e.g., the expectation value of the energy in the state under consideration (see Chap. 6). Moreover, the knowledge of the sequence $\{\gamma_N^{(k)}\}_N$ at level k allows one to control the limit of the expectation values of k-particle observables.

From the mathematical point of view, when condensation is needed as an hypothesis (e.g., assuming condensation at the initial time to investigate the dynamics at further times), it is often necessary to assume it in the apparently stronger form of an asymptotic factorisation of $\gamma_N^{(k)}$ into some $|\varphi\rangle\langle\varphi|^{\otimes k}$, for any fixed positive integer k. The paradigmatic case is the already mentioned ESY theorem 5.3.1.

The behaviour of the *two*-particle (or k-particle) density matrix has then to enter the mathematical formalisation of B.E.C.

First of all, in this section we point out that to define and to control complete condensation in the context of asymptotic B.E.C. *the one-body level is enough.*

Theorem 7.1.1. Let $\{\Psi_N\}_N$ be a sequence of bosonic *pure* states with the property of asymptotic 100% B.E.C. and let φ be the corresponding condensate wave function (see definition 3.4.1). That is,

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi\rangle\langle\varphi| \tag{7.1}$$

in the usual weakly-$*$ or trace-norm sense (see theorem 3.4.2). Then

$$\lim_{N\to\infty} \gamma_N^{(k)} = |\varphi\rangle\langle\varphi|^{\otimes k} \qquad \forall\, k \geqslant 1 \tag{7.2}$$

in the same sense.

This fact was apparently first noticed and proved by Lieb and Seiringer – see the remark after theorem 1 in [74] – within Bogolubov's formalism of second quantization and it applies both to pure and mixed states. This is actually a point where, to be

consistent with our choice of presenting the subject in an elementary way, a longer proof is needed than that with creation and annihilation operators. (In that case, the commutation relations give the upper bound and a convexity argument give a lower bound leading to the limit (7.2).) For completeness, let us prove theorem 7.1.1 in the present formalism. Generalization to *mixed* states is the content of the following theorem 7.1.2.

Proof of theorem 7.1.1. We give details of how (7.1) implies (7.2) for $k = 2$. The general case follows along the same line.

Complete φ to an orthonormal basis $\{\varphi_h\}_{h\geqslant 1}$ of the one-body Hilbert space $\mathcal{H}$, with $\varphi_1 = \varphi$. Let $\mathcal{H}_N$ be the N-body Hilbert space $\mathcal{H}^{\otimes N}$. For each N, let $\{\Phi_j^{(N-1)}\}_j$ be an orthonormal basis of $\mathcal{H}_{N-1}$ and let $\{\Xi_\ell^{(N-2)}\}_\ell$ be an orthonormal basis of $\mathcal{H}_{N-2}$.

By assumption (see theorem 3.4.4),

$$(\varphi, \gamma_N^{(1)}\varphi)_{\mathcal{H}} \xrightarrow{N\to\infty} 1\,, \tag{7.3}$$

and, by characterization (2.48) of partial trace, and since $\gamma_N = |\Psi_N\rangle\langle\Psi_N|$ (pure state),

$$\begin{aligned}\sum_j \left| \left(\varphi \otimes \Phi_j^{(N-1)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 &= \sum_j \left(\varphi \otimes \Phi_j^{(N-1)}, \gamma_N\, \varphi \otimes \Phi_j^{(N-1)}\right)_{\mathcal{H}_N} \\ &= (\varphi, \gamma_N^{(1)}\varphi)_{\mathcal{H}} \xrightarrow{N\to\infty} 1\,,\end{aligned} \tag{7.4}$$

whence

$$\sum_{\substack{j\\ h\neq 1}} \left| \left(\varphi_h \otimes \Phi_j^{(N-1)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 \xrightarrow{N\to\infty} 0\,. \tag{7.5}$$

What one needs to prove (see theorem 3.4.4) is

$$(\varphi^{\otimes 2}, \gamma_N^{(2)}\varphi^{\otimes 2})_{\mathcal{H}_2} \xrightarrow{N\to\infty} 1\,, \tag{7.6}$$

where, by (2.48),

$$\begin{aligned}(\varphi^{\otimes 2}, \gamma_N^{(2)}\varphi^{\otimes 2})_{\mathcal{H}_2} &= \sum_\ell \left(\varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}, \gamma_N\, \varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}\right)_{\mathcal{H}_N} \\ &= \sum_\ell \left| \left(\varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 .\end{aligned} \tag{7.7}$$

Since $\{\varphi_h \otimes \Xi_\ell^{(N-2)}\}_{h,\ell}$ is an orthonormal basis of $\mathcal{H}_{N-1}$, by (7.4)

$$\sum_\ell \left| \left(\varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 = (\varphi, \gamma_N^{(1)}\varphi)_{\mathcal{H}} \; - \sum_{\substack{\ell\\ h\neq 1}} \left| \left(\varphi \otimes \varphi_h \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 \tag{7.8}$$

and, since Ψ_N is symmetric under permutation of variables,

$$\begin{aligned}\sum_{\substack{\ell\\ h\neq 1}} \left| \left(\varphi \otimes \varphi_h \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 &= \sum_{\substack{\ell\\ h\neq 1}} \left| \left(\varphi_h \otimes \varphi \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 \\ &\leqslant \sum_{\substack{\ell,r\\ h\neq 1}} \left| \left(\varphi_h \otimes \varphi_r \otimes \Xi_\ell^{(N-2)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 \\ &= \sum_{\substack{j\\ h\neq 1}} \left| \left(\varphi_h \otimes \Phi_j^{(N-1)}, \Psi_N\right)_{\mathcal{H}_N}\right|^2 \xrightarrow{N\to\infty} 0\,.\end{aligned} \tag{7.9}$$

Then, plugging (7.3) and (7.9) into (7.8) as $N \to \infty$, (7.6) follows. □

Theorem 7.1.2. Let $\{\gamma_N\}_N$ be a sequence of bosonic (pure or *mixed*) states with the property of asymptotic 100% B.E.C. and let φ be the corresponding condensate wave function,

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi\rangle\langle\varphi| \,. \tag{7.10}$$

Then

$$\lim_{N\to\infty} \gamma_N^{(k)} = |\varphi\rangle\langle\varphi|^{\otimes k} \qquad \forall\, k \geqslant 1 \,. \tag{7.11}$$

(Limits in any of the usual senses.)

Proof. (Same notation of last theorem; details for $k = 2$ only.) Decompose each γ_N in its diagonal form

$$\gamma_N = \sum_i \lambda_i |\Psi_N^{(i)}\rangle\langle\Psi_N^{(i)}| \,, \qquad \lambda_i > 0 \,, \qquad \sum_i \lambda_i = 1 \,, \tag{7.12}$$

where $\{\Psi_N^{(i)}\}_i$ is an orthonormal system of $\mathcal{H}_N$. By assumption (see theorem 3.4.4), and by characterization (2.48) of partial trace,

$$\begin{aligned} 1 \xleftarrow{N\to\infty} (\varphi, \gamma_N^{(1)}\varphi)_{\mathcal{H}} &= \sum_j \big(\varphi \otimes \Phi_j^{(N-1)}, \gamma_N\, \varphi \otimes \Phi_j^{(N-1)}\big)_{\mathcal{H}_N} \\ &= \sum_i \lambda_i \sum_j \Big| \big(\varphi \otimes \Phi_j^{(N-1)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 , \end{aligned} \tag{7.13}$$

whence

$$\begin{aligned} &\sum_i \lambda_i \sum_{\substack{j \\ h\neq 1}} \Big| \big(\varphi_h \otimes \Phi_j^{(N-1)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 = \\ &\qquad = \sum_i \lambda_i \Big(1 - \sum_j \Big| \big(\varphi \otimes \Phi_j^{(N-1)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \Big) \xrightarrow{N\to\infty} 0 \,. \end{aligned} \tag{7.14}$$

(since $\sum_i \lambda_i = 1$). On the other hand,

$$\begin{aligned} (\varphi^{\otimes 2}, \gamma_N^{(2)} \varphi^{\otimes 2})_{\mathcal{H}_2} &= \sum_\ell \big(\varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}, \gamma_N\, \varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}\big)_{\mathcal{H}_N} \\ &= \sum_i \lambda_i \sum_\ell \Big| \big(\varphi^{\otimes 2} \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \\ &= \sum_i \lambda_i \Big[\big(\varphi, \mathrm{Tr}_{[N-1]}\big[\, |\Psi_N^{(i)}\rangle\langle\Psi_N^{(i)}| \,\big]\, \varphi\big)_{\mathcal{H}} - \\ &\qquad\qquad - \sum_{\substack{\ell \\ h\neq 1}} \Big| \big(\varphi \otimes \varphi_h \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \Big] \\ &= (\varphi, \gamma_N^{(1)}\varphi)_{\mathcal{H}} \; - \sum_i \lambda_i \sum_{\substack{\ell \\ h\neq 1}} \Big| \big(\varphi \otimes \varphi_h \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 . \end{aligned} \tag{7.15}$$

Since each $\Psi_N^{(i)}$ is symmetric under permutation of variables and $\{\varphi_r \otimes \Xi_\ell^{(N-2)}\}_{r,\ell}$ is an orthonormal basis of $\mathcal{H}_{N-1}$, by (7.14)

$$\begin{aligned}
\sum_i \lambda_i \sum_{\substack{\ell \\ h\neq 1}} \Big| \big(\varphi \otimes \varphi_h \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 = \\
&= \sum_i \lambda_i \sum_{\substack{\ell \\ h\neq 1}} \Big| \big(\varphi_h \otimes \varphi \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \\
&\leqslant \sum_i \lambda_i \sum_{\substack{\ell,r \\ h\neq 1}} \Big| \big(\varphi_h \otimes \varphi_r \otimes \Xi_\ell^{(N-2)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \\
&= \sum_i \lambda_i \sum_{\substack{j \\ h\neq 1}} \Big| \big(\varphi_h \otimes \Phi_j^{(N-1)}, \Psi_N^{(i)}\big)_{\mathcal{H}_N} \Big|^2 \xrightarrow{N\to\infty} 0\,.
\end{aligned} \tag{7.16}$$

Thus, plugging (7.13) and (7.16) into the r.h.s. of (7.15), one has

$$\lim_{N\to\infty} (\varphi^{\otimes 2}, \gamma_N^{(2)} \varphi^{\otimes 2})_{\mathcal{H}_2} = 1 \tag{7.17}$$

which, by theorem 3.4.4, is the same as the trace-norm limit (7.2). □

We can thus conclude that the final content of theorems 3.4.4 and 7.1.2 is: *100% B.E.C. into the one-body state φ, as an asymptotic property of a sequence $\{\gamma_N\}_N$ of N-body bosonic states in the sense of definition 3.4.1, is the factorisation of a whatever k-particle reduced density matrix $\gamma_N^{(k)}$ into the k-th tensor power of the projection onto φ, whatever the topology is for the convergence $\gamma_N^{(k)} \to |\varphi\rangle\langle\varphi|^{\otimes k}$, be it weak-$*$ trace class, Hilbert-Schmidt in norm, or trace class in norm.*

7.2 Asymptotic factorisation of the N-body wave function

There exist other types of complete condensation assumptions in the literature than the asymptotic factorisation (7.2) of the k-th marginals. These are usually needed as apparently stronger technical hypotheses on the initial state when one wants to analyse its time evolution and are in the form of conditions on the N-body state, instead of on the marginals. Following Ref. [83], in this section we deal with the two of them that mostly entered the study of time-stability of condensates and the rigorous derivation of the time-dependent Gross-Pitaevskiĭ equation (Sec. 5.3) and we prove their *equivalence* with the standard definition of asymptotic 100% B.E.C.

This provides a homogenization of the hypotheses appearing in the statements of a number of recent theorems. Yet this does not cancel the merit of having introduced such alternative assumptions of condensation which focus on the specific properties of the initial state that are crucial to prove the desired statements.

Definition 7.2.1 (**partial asymptotic factorisation**)**.** Let $\mathcal{H}$ be a separable Hilbert space, set $\mathcal{H}_N := \mathcal{H}^{\otimes N}$ and let $\mathcal{H}_{N,\mathrm{sym}}$ be its bosonic sector. Let $\{\Psi_N\}_N$ be a sequence of N-body bosonic pure states ($\|\Psi_N\| = 1$ according to our definition), each $\Psi_N \in \mathcal{H}_{N,\mathrm{sym}} \subset \mathcal{H}_N$. Let $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$. The sequence $\{\Psi_N\}_N$ is said to have the

partial asymptotic factorisation (P.A.F.) property in the state φ iff

$$\begin{aligned} &\forall\, N \geqslant 1\,, \quad \forall\, k = 1, \dots, N \\ &\exists\, \Xi^{(N-k)} \in \mathcal{H}_{N-k}\,, \quad \big\| \Xi^{(N-k)} \big\|_{\mathcal{H}_{N-k}} = 1 \;: \\ &\lim_{N\to\infty} \big\| \Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)} \big\|_{\mathcal{H}_N} = 0\,. \end{aligned} \tag{7.18}$$

Notice that the P.A.F. is a statement on the collection $\{\Psi_N\}_N$ compared with a countable infinity of collections $\{\varphi^{\otimes k} \otimes \Xi^{(N-k)}\}_N$ with $k = 1, 2, 3, \dots$: each Ψ_N can be approximated in norm with a $\varphi^{\otimes k} \otimes \Xi^{(N-k)}$ for any k, the larger N, the better the approximation. Unlike Ψ_N, the $\varphi^{\otimes k} \otimes \Xi^{(N-k)}$'s are not necessarily in $\mathcal{H}_{N,\mathrm{sym}}$: fixed any positive integer k and provided that N is large enough, one can always find a vector close to each Ψ_N which is completely factorised in the first k variables (actually it is a k-th tensor power), tensor a tail in the remaining $N - k$. Choosing a different k, the same holds, possibly in a different regime of large N's, that is to say that limit (7.18) is not uniform in k.

From (7.18), it follows that Ψ_N is asymptotically close in norm also to the symmetrized of $\varphi^{\otimes k} \otimes \Xi^{(N-k)}$ – see (3.18): indeed by (2.106)

$$\big\| \Psi_N - \big(\varphi^{\otimes k} \otimes \Xi^{(N-k)}\big)_{\mathrm{sym}} \big\|_{\mathcal{H}_N} \leqslant \frac{2 \big\| \Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)} \big\|_{\mathcal{H}_N}}{1 - \big\| \Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)} \big\|_{\mathcal{H}_N}} \tag{7.19}$$

eventually with N. The converse is false, in general. Actually, in the applications it is crucial that Ψ_N is controlled in norm with the non permutationally symmetric $\varphi^{\otimes k} \otimes \Xi^{(N-k)}$.

Condition (7.18) has been introduced in 2006 by Erdős, Schlein, and Yau [44], first by recognising it to be the convenient assumption which ensures that regularizing the initial wave function Ψ_N by a cut-off on its high-energy component leads to a regularized $\widetilde{\Psi}_N$ which exhibits condensation $\widetilde{\gamma}_N^{(k)} \to |\varphi\rangle\langle\varphi|^{\otimes k}$ and has N-uniformly bounded energy moments (see proposition 8.1 in [44]). Explicit $\{\Psi_N\}_N$ with P.A.F. are shown in appendix B of [44]. Soon after it has been recognised by the same authors that P.A.F. holds for the ground state of the Hamiltonian (3.64) of a trapped Bose gas, but with some more restrictive conditions on the pair interaction, and factorisation holds with respect to the one-body state φ^{GP} (see proposition C.2 in [44]).

Explicitly (see lemma B.1 in [44]), the following bosonic states can be proved to have P.A.F. in the one-body state φ:

$$\Psi_N(x_1, \dots, x_N) \;:=\; \frac{W_N(x_1, \dots, x_N) \prod_{i=1}^N \varphi(x_i)}{\Big\| W_N(x_1, \dots, x_N) \prod_{i=1}^N \varphi(x_i) \Big\|} \tag{7.20}$$

$$W_N(x) \;:=\; \prod_{1\leqslant i<j\leqslant N} f_N(x_i - x_j) \tag{7.21}$$

where now f_N is defined to be identically 1 outside the ball $\mathcal{B}_N$ in $\mathbb{R}^3$ of radius r_N such that $a_N \ll r_N \ll 1$, while inside $\mathcal{B}_N$ it is defined to be the ground state solution of the

Neumann problem

$$\begin{cases} \big(-2\mu\Delta + V_N(x)\big) f_N(x) &= E_N f_N(x) & x \in \mathcal{B}_N \\ f_N(x) &= 1 & x \in \partial\mathcal{B}_N \\ \partial_{\text{radial}}\, f_N(x) &= 0 & x \in \partial\mathcal{B}_N\,. \end{cases} \tag{7.22}$$

Here V_N is the two-body interaction $V_N(x) := N^2 V(Nx)$, where the unscaled V has scattering length a and is assumed to be non-negative, smooth, spherically symmetric and compactly supported. Then $a_N = a/N$ is the scattering length of V_N (with respect to notation of theorem 3.5.2 for trapped Bose gases, here we set unrestrictively $N_0 = 1$). P.A.F. for states (7.20) is verified with respect to the tail

$$\Xi^{(N-k)}(x_{k+1}, \dots, x_N) = \frac{\displaystyle\prod_{k\leqslant i<j\leqslant N} f_N(x_i - x_j) \prod_{i=k+1}^{N} \varphi(x_i)}{\left\| \displaystyle\prod_{k\leqslant i<j\leqslant N} f_N(x_i - x_j) \prod_{i=k+1}^{N} \varphi(x_i) \right\|}\,. \tag{7.23}$$

By lemma A.2 in [41] (see also appendix B in [44]), f_N has the following behaviour: it is spherically symmetric, and there exist positive constants C_1, C_2, independent of N and depending only on V, such that

$$\begin{aligned} 1 - C_1 \frac{a_N}{|x| + a_N} \mathbb{1}_{\mathcal{B}_N}(x) \;\leqslant\; f_N(x) &\;\leqslant\; 1 \\ |\nabla f_N(x)| &\;\leqslant\; C_2 \frac{a_N}{(|x| + a_N)^2} \mathbb{1}_{\mathcal{B}_N}(x) \end{aligned} \tag{7.24}$$

where $\mathbb{1}_{\mathcal{B}_N}$ is the characteristic function of the ball $\mathcal{B}_N$. Thus the W_N-factor (7.21) accounts for correlations at the scale r_N in the wave function (7.20), analogously to the Dyson-like correlations (6.5) used in Lieb and Seiringer's trial function (6.4) for the ground state of a trapped Bose gas.

By comparing (7.20) and (7.23), one sees that $\Xi^{(N-k)} = \Psi_{N-k}$ so that P.A.F. (7.18) for $\{\Psi_N\}_N$ reads simply $\|\Psi_N - \varphi^{\otimes k} \otimes \Psi_{N-k}\| \to 0$ as $N \to \infty$. This motivates to introduce the following.

Definition 7.2.2. Let $\mathcal{H}$ be a separable Hilbert space, set $\mathcal{H}_N := \mathcal{H}^{\otimes N}$, and let $\mathcal{H}_{N,\text{sym}}$ be its bosonic sector. Let $\{\Psi_N\}_N$ be a sequence of N-body bosonic pure states ($\|\Psi_N\| = 1$ according to our definition), each $\Psi_N \in \mathcal{H}_{N,\text{sym}} \subset \mathcal{H}_N$. Let $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$. The sequence $\{\Psi_N\}_N$ is said to have the SELF-TAIL (S.T.) PROPERTY with respect to the one-body state φ iff

$$\lim_{N\to\infty} \big\| \Psi_{N+1} - \varphi \otimes \Psi_N \big\| = 0\,. \tag{7.25}$$

Notice that, as for the P.A.F. property, the S.T. property is a statement on the collection $\{\Psi_N\}_N$ compared with $\{\varphi \otimes \Psi_N\}_N$: adding a particle φ to the N-th element of the sequence $\{\Psi_N\}_N$ to get the $(N+1)$-body state $\varphi \otimes \Psi_N$, one essentially obtains the subsequent Ψ_{N+1} of the sequence, the larger N, the smaller the error in norm. Again, the comparison is with a non permutationally symmetric N-body state and the variables ordering in $\{\varphi \otimes \Psi_N\}_N$ is crucial for the applications: $\{\Psi_N \otimes \varphi\}_N$ would *not* play the same role. Notice, also, that it is just a matter of a finite number of triangle inequalities to prove that

$$\lim_{N\to\infty} \big\| \Psi_{N+1} - \varphi \otimes \Psi_N \big\| = 0 \quad \Rightarrow \quad \lim_{N\to\infty} \big\| \Psi_N - \varphi^{\otimes k} \otimes \Psi_{N-k} \big\| = 0 \quad \forall\, k \tag{7.26}$$

(the converse being obvious), so that one can refer to such a property evoking indifferently (7.25) or the r.h.s. of (7.26).

What we now prove is that *all these are equivalent notion of condensation.*

Theorem 7.2.3 (**equivalent definitions of B.E.C. for pure states**).

$$\begin{array}{c}\text{asymptotic}\\ \text{100\% B.E.C.}\end{array} \quad \Leftrightarrow \quad \begin{array}{c}\text{partial asymptotic}\\ \text{factorisation}\end{array} \quad \Leftrightarrow \quad \begin{array}{c}\text{self-tail}\\ \text{property}\end{array}$$

In detail: let $\mathcal{H}$ be a separable Hilbert space, set $\mathcal{H}_N := \mathcal{H}^{\otimes N}$, and let $\mathcal{H}_{N,\mathrm{sym}}$ be its bosonic sector. Let $\{\Psi_N\}_N$ be a sequence of N-body bosonic pure states ($\|\Psi_N\| = 1$ according to our definition), each $\Psi_N \in \mathcal{H}_{N,\mathrm{sym}} \subset \mathcal{H}_N$. Let $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$. Then any of the conditions of theorem 3.4.4, (7.18), and (7.25) are all equivalent.

Proof of: S.T. property ⇒ P.A.F. property. This is just a finite iteration of triangle inequality to get (7.26). □

Proof of: P.A.F. property ⇒ asymptotic 100% B.E.C.. Assume (7.18) holds: it suffices to use it for $k = 1$. Let $\gamma_N^{(1)}$ be its one-particle reduced density matrix of Ψ_N and let $C \in \mathrm{Com}(\mathcal{H})$, a generic compact on the one-body Hilbert space $\mathcal{H}$. The one-particle reduced density matrix of $\varphi \otimes \Xi^{(N-1)}$ is clearly $|\varphi\rangle\langle\varphi|$. Then ,by characterization (2.49) of partial trace,

$$\begin{aligned}
&\big|\,\mathrm{Tr}\big[\,C(\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|)\,\big]\,\big| = \\
&= \big|\,\mathrm{Tr}\big[\,C\gamma_N^{(1)}\,\big] - \mathrm{Tr}[\,C|\varphi\rangle\langle\varphi|\,\big]\,\big| \\
&= \big|\,\mathrm{Tr}\big[\,(C\otimes\mathbb{1}_{N-1})\,|\Psi_N\rangle\langle\Psi_N|\,\big] - \mathrm{Tr}\big[\,(C\otimes\mathbb{1}_{N-1})\,|\varphi\otimes\Xi^{(N-1)}\rangle\langle\varphi\otimes\Xi^{(N-1)}|\,\big]\,\big| \\
&= \big|\,\big(\Psi_N, (C\otimes\mathbb{1}_{N-1})\Psi_N\big)_{\mathcal{H}_N} - \big(\varphi\otimes\Xi^{(N-1)}, (C\otimes\mathbb{1}_{N-1})\varphi\otimes\Xi^{(N-1)}\big)_{\mathcal{H}_N}\,\big| \\
&\leqslant \|\,C\otimes\mathbb{1}_{N-1}\,\|_{\mathcal{L}(\mathcal{H}_N)} \cdot \big\|\,\Psi_N - \varphi\otimes\Xi^{(N-1)}\,\big\|_{\mathcal{H}_N} \cdot \big(\,2 + \big\|\,\Psi_N - \varphi\otimes\Xi^{(N-1)}\,\big\|_{\mathcal{H}_N}\big) \\
&\leqslant 3\,\|C\|_{\mathcal{L}(\mathcal{H})} \cdot \big\|\,\Psi_N - \varphi\otimes\Xi^{(N-1)}\,\big\|_{\mathcal{H}_N} \xrightarrow{N\to\infty} 0
\end{aligned} \tag{7.27}$$

(first inequality following because in general

$$|\,(x, Ay) - (y, Ay)\,| \leqslant \|A\| \cdot \|x-y\| \cdot (2 + \|x-y\|)$$

for any *bounded* operator A on some Hilbert space and any norm-one vectors x, y), that is, $\gamma_N^{(1)} \to |\varphi\rangle\langle\varphi|$ weakly-$*$ in $\mathcal{L}^1(\mathcal{H})$. This is exactly one of the equivalent ways to define 100% asymptotic B.E.C. □

Alternative proof of: P.A.F. property ⇒ asymptotic 100% B.E.C.. By (2.72) – see theorem 2.4.1,

$$\big\|\,\gamma_N - |\varphi\rangle\langle\varphi|\otimes\big|\,\Xi^{(N-1)}\big\rangle\big\langle\,\Xi^{(N-1)}\big|\,\big\|_{\mathcal{L}^1(\mathcal{H}_N)} \leqslant 2\,\big\|\,\Psi_N - \varphi\otimes\Xi^{(N-1)}\,\big\|_{\mathcal{H}_N}, \tag{7.28}$$

and by (2.104) – see theorem 2.4.4,

$$\big\|\,\gamma_N^{(1)} - |\varphi\rangle\langle\varphi|\,\big\|_{\mathcal{L}^1(\mathcal{H})} \leqslant \big\|\,\gamma_N - |\varphi\rangle\langle\varphi|\otimes\big|\,\Xi^{(N-1)}\big\rangle\big\langle\,\Xi^{(N-1)}\big|\,\big\|_{\mathcal{L}^1(\mathcal{H}_N)} \tag{7.29}$$

so that $\gamma_N^{(1)} \to |\varphi\rangle\langle\varphi|$ in trace norm. □

Proof of: asymptotic 100% B.E.C. ⇒ P.A.F. property. Pick a positive integer k. By assumption (see theorem 3.4.4),

$$\lim_{N\to\infty} \big(\varphi^{\otimes k}, \gamma_N^{(k)} \varphi^{\otimes k}\big)_{\mathcal{H}_k} = 1\,. \tag{7.30}$$

Complete $\varphi^{\otimes k}$ to an orthonormal basis $\{f_i\}_i$ of $\mathcal{H}_k$ in such a way that $f_1 = \varphi^{\otimes k}$. Choose in each $\mathcal{H}_{N-k}$ an orthonormal basis $\{g_j^{(N-k)}\}_j$: hence, each $\Psi_N \in \mathcal{H}_N \simeq \mathcal{H}_k \otimes \mathcal{H}_{N-k}$ can be written as

$$\Psi_N = \sum_{i,j} \alpha_{i,j}^{(N)}\, f_i \otimes g_j^{(N-k)} \tag{7.31}$$

for suitable complex coefficients $\alpha_{i,j}^{(N)}$. Then

$$\begin{aligned}
\gamma_N := |\Psi_N\rangle\langle\Psi_N| = &\sum_j \sum_{j'} \alpha_{1,j}^{(N)}\, \overline{\alpha_{1,j'}^{(N)}}\; |\varphi^{\otimes k}\rangle\langle\varphi^{\otimes k}| \otimes \big|\, g_j^{(N-k)}\,\big\rangle\big\langle\, g_{j'}^{(N-k)}\big| \\
&+ \sum_j \sum_{\substack{i'\neq 1\\ j'}} \alpha_{1,j}^{(N)}\, \overline{\alpha_{i',j'}^{(N)}}\; |\varphi^{\otimes k}\rangle\langle\, f_{i'}\,| \otimes \big|\, g_j^{(N-k)}\,\big\rangle\big\langle\, g_{j'}^{(N-k)}\big| \\
&+ \sum_{\substack{i\neq 1\\ j}} \sum_{j'} \alpha_{i,j}^{(N)}\, \overline{\alpha_{1,j'}^{(N)}}\; |\,f_i\,\rangle\langle\varphi^{\otimes k}| \otimes \big|\, g_j^{(N-k)}\,\big\rangle\big\langle\, g_{j'}^{(N-k)}\big| \\
&+ \sum_{\substack{i\neq 1\\ j}} \sum_{\substack{i'\neq 1\\ j'}} \alpha_{i,j}^{(N)}\, \overline{\alpha_{i',j'}^{(N)}}\; |\,f_i\,\rangle\langle\, f_{i'}\,| \otimes \big|\, g_j^{(N-k)}\,\big\rangle\big\langle\, g_{j'}^{(N-k)}\big|\,.
\end{aligned} \tag{7.32}$$

Taking the partial trace with respect to $\mathcal{H}_{N-k}$, according to (2.52), and using the first of (2.19), gives

$$\begin{aligned}
\gamma_N^{(k)} = |\varphi^{\otimes k}\rangle\langle\varphi^{\otimes k}| \cdot &\sum_j \big|\alpha_{1,j}^{(N)}\big|^2 \\
&+ \sum_{\substack{i'\neq 1\\ j'}} \alpha_{1,j'}^{(N)}\, \overline{\alpha_{i',j'}^{(N)}}\; |\varphi^{\otimes k}\rangle\langle\, f_{i'}\,| \\
&+ \sum_{\substack{i\neq 1\\ j}} \alpha_{i,j}^{(N)}\, \overline{\alpha_{1,j}^{(N)}}\; |\,f_i\,\rangle\langle\varphi^{\otimes k}| \\
&+ \sum_{\substack{i\neq 1\\ i'\neq 1\\ j}} \alpha_{i,j}^{(N)}\, \overline{\alpha_{i',j}^{(N)}}\; |\,f_i\,\rangle\langle\, f_{i'}\,|\,.
\end{aligned} \tag{7.33}$$

As a consequence, by (7.30),

$$\sum_j \big|\alpha_{1,j}^{(N)}\big|^2 = \big(\varphi^{\otimes k}, \gamma_N^{(k)} \varphi^{\otimes k}\big)_{\mathcal{H}_k} \xrightarrow{N\to\infty} 1\,, \tag{7.34}$$

whence

$$\Big\| \Psi_N - \sum_j \alpha_{1,j}^{(N)}\, \varphi^{\otimes k} \otimes g_j^{(N-k)} \Big\|_{\mathcal{H}_N}^2 = \sum_{\substack{i\neq 1\\ j}} \big|\alpha_{i,j}^{(N)}\big|^2 = 1 - \sum_j \big|\alpha_{1,j}^{(N)}\big|^2 \xrightarrow{N\to\infty} 0\,. \tag{7.35}$$

Finally set

$$\Xi^{(N-k)} := \sum_j \alpha_{1,j}^{(N)}\, g_j^{(N-k)} \tag{7.36}$$

so that (7.35) reads $\|\Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)}\| \to 0$ as $N \to \infty$: by the arbitrariness of k this is just the P.A.F. property (7.18). □

Remark, incidentally, that (7.36) accounts for the non uniqueness of $\Xi^{(N-k)}$ in the P.A.F. property.

Proof of: P.A.F. property ⇒ S.T. property. One has

$$\begin{aligned} \big\| \Psi_{N+1} - \varphi \otimes \Psi_N \big\| &\leqslant \big\| \Psi_{N+1} - \varphi^{\otimes(k+1)} \otimes \Xi^{(N-k)} \big\| + \\ &\quad + \big\| \varphi^{\otimes(k+1)} \otimes \Xi^{(N-k)} - \varphi \otimes \Psi_N \big\| \\ &= \big\| \Psi_{N+1} - \varphi^{\otimes(k+1)} \otimes \Xi^{(N-k)} \big\| + \big\| \Psi_N - \varphi^{\otimes k} \otimes \Xi^{(N-k)} \big\| \end{aligned} \tag{7.37}$$

the r.h.s. vanishes as $N \to \infty$ by assumption of P.A.F., hence (7.25) follows. □

The mixed state version of theorem 7.2.3 holds too, although it is not of relevance in the current literature. We state it for completeness. The proof is analogous to the pure states case.

Theorem 7.2.4 (**equivalent definitions of B.E.C. for mixed states**)**.** Let $\mathcal{H}$ be a separable Hilbert space and set $\mathcal{H}_N := \mathcal{H}^{\otimes N}$. Let $\{\gamma_N\}_N$ be a sequence of N-body bosonic mixed states, each γ_N acting on $\mathcal{H}_N$. Let $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$. Then any of the following conditions are equivalent:

- 100% asymptotic B.E.C.

$$\lim_{N\to\infty} \gamma_N^{(1)} = |\varphi\rangle\langle\varphi| \tag{7.38}$$

- partial asymptotic factorisation

$$\begin{aligned} &\forall\, N \geqslant 1\,, \quad \forall\, k = 1, \dots, N \\ &\exists\, \rho_{N-k} \ \text{ bosonic density matrix on } \mathcal{H}_{N-k} \ : \\ &\lim_{N\to\infty} \big\| \gamma_N - |\varphi\rangle\langle\varphi|^{\otimes k} \otimes \rho_{N-k} \big\|_{\mathcal{L}^1(\mathcal{H}_N)} = 0 \end{aligned} \tag{7.39}$$

- self-tail condition

$$\lim_{N\to\infty} \big\| \gamma_{N+1} - |\varphi\rangle\langle\varphi| \otimes \gamma_N \big\|_{\mathcal{L}^1(\mathcal{H}_N)} = 0\,. \tag{7.40}$$

Chapter 8

Concluding remarks

Condensation is an "ancient" and "recent" phenomenon: theoretically predicted 80 years ago, experimentally observed from 1995 on, mathematically studied since the 1950s, until the renewed flurry of interest in the rigorous analysis of the last decade. B.E.C. is a signature of Quantum Mechanics: no explanation is possible in a classical picture. It is a consequence of the quantum statistics of Bosons only. Recent B.E.C. realizations have represented the apex of huge experimental efforts and the confirm of a long theoretical research. On the other side, they have re-fired the need of a fully rigorous comprehension of the phenomenology in terms of first principles.

To conclude this work, let us first collect the key points developed throughout and after outline some perspectives.

The mathematics of B.E.C.

1. Limit $N \to \infty$. Formalisation and symptotic analysis of condensation in many-body dilute Bose gases in the limit $N \to \infty$ relies on the fact that, in practice, the number of particles is so large ($N \sim 100 \div 10^{11}$) to justify the assumption of infinite size and condensation for interacting systems requires to solve a formidably difficult many-body problem in order to access $\gamma_N^{(1)}$, while there are techniques to get its asymptotics (for interacting systems the ground state already poses a challenging problem which is still largely unsolved: B.E.C. itself has, so far, never been proved for many-body Hamiltonians with genuine interactions, except for the special case of hard core bosons on a lattice at half-filling).

2. Scaling. Taking $N \to \infty$ is not just a thermodynamic limit. Instead, a more careful treatment is needed: some scaling has to be precised. The true system $\mathcal{S}_0$ in the state γ_{N_0} and with Hamiltonian H_{N_0} is regarded as an enlarging sequence $\{\mathcal{S}_N\}_N$ of N-body systems, each in the state γ_N and with Hamiltonian H_N. As $N = N_0$, $\mathcal{S}_N$ is exactly $\mathcal{S}_0$, i.e., $H_N|_{N=N_0} = H_{N_0}$ on $\mathcal{H}_N = \mathcal{H}_{N_0}$. For any N, or asymptotically as $N \to \infty$, *certain* physical quantities of $\mathcal{S}_N$ in the state γ_N are the same as the corresponding ones for $\mathcal{S}_0$ in the state γ_{N_0}. This way, each $\mathcal{S}_N$ resembles by construction the original $\mathcal{S}_0$ at least for all those physical quantities that are kept fixed with N. By scaling, one removes the ambiguity one would have in determining the asymptotic of the expectation $f(N, L, a, \dots)$ of an observable, by computing $\lim_{N\to\infty} f(N, L_N, a_N, \dots)$. The Gross-Pitaevskiĭ scaling, in particular, is central in this analysis since it is *the* limit

of ultra-high dilution which still enables one to study the dynamics of the gas, because the ratio between interaction and kinetic energy is kept constant with N.

3. Role of one-body reduced density matrices. $\gamma_N^{(1)}$ is the natural tool to formalise and investigate condensation. One-body marginals enter the description of the many-body system because one usually deals with *local* observables only, i.e., observables of the form $\mathcal{O} \otimes \mathbf{1}_{N-k}$ acting on $\mathcal{H}_N \cong \mathcal{H}_k \otimes \mathcal{H}_{N-k}$. Their expectation values $\mathrm{Tr}[\,\gamma_N(\mathcal{O} \otimes \mathbf{1}_{N-k})\,] = \mathrm{Tr}[\,\gamma_N^{(k)}\mathcal{O}\,]$ involve only the knowledge of $\gamma_N^{(k)}$. In particular, even when dealing with B.E.C. for finite size systems, experimentally one observes a spatial density distribution

$$n(x) = N \int \left|\Psi_N(x, x_2, \dots, x_N)\right|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N = N\, \gamma_N^{(1)}(x, x)$$

providing the average number of particles in some neighborhood of given point (indeed $\int n(x)\,\mathrm{d}x = N$). As $N \to \infty$, the many-body Hilbert space $\mathcal{H}_N$ tends to an infinite tensor power and objects like γ_N act on a far too large space to control. As a consequence, when N is large, two "physically essentially equal" many-body states Ψ_N and Φ_N are essentially *orthogonal*, whereas evaluating their distance at the level of marginals turn out to be indeed meaningful. (Compare, for instance, (6.1) with (6.3): the l.h.s. of $\|\gamma_N^{(1)} - \rho_N^{(1)}\|_{\mathcal{L}^1} \leqslant \|\Psi_N - \Phi_N\|$ vanishes with N, the r.h.s. remaining constant.) Conversely, the sequence $\{\gamma_N^{(k)}\}_N$ is all inside $\mathcal{L}^1(\mathcal{H}_k)$, hence taking $N \to \infty$ makes sense, provided that a suitable topology for the limit is specified.

4. Complete condensation. The ubiquitous form of B.E.C. in all currently known rigorous results is that of asymptotic 100% B.E.C., namely, the asymptotic convergence of the one-body marginal to some rank-one projection. The one-body state it projects onto is the condensate wave function. This is because these results are obtained within the Gross-Pitaevskiĭ scaling, where condensation, if it happens to hold, is necessarily depletionless. (In fact, in the definition of asymptotic B.E.C. one has also to specify how the limit $N \to \infty$ is taken.) Yet, in principle the case with depletion is not excluded: depletion is zero in the limit and (small and) positive before the limit and a (highly non-trivial, actually a major open problem) control of the errors would provide a control of $\gamma_N^{(1)}$ for finite large values of N and, hence, a description of the thermal component around the condensate cloud. The rigorous proof of B.E.C. and of the whole GP theory are today achieved within this scaling. Convergence $\gamma_N^{(1)} \to |\varphi\rangle\langle\varphi|$ can be *equivalently* controlled in the trace class weak-$*$ sense, or as $(\varphi, \gamma_N^{(1)}\varphi) \to 1$, or in the Hilbert-Schmidt norm, or in the trace class norm, and it implies the factorisation $\gamma_N^{(k)} \to |\varphi\rangle\langle\varphi|^{\otimes k}$ for any positive integer k, both in case of pure and of mixed states. Moreover, complete B.E.C. in the limit $N \to \infty$ is *equivalent* to the partial asymptotic factorisation (7.18) and to the self-tail property (7.25) of the many-body pure state (see (7.39) and (7.40) for the mixed state counterpart).

5. Evolution of an originally trapped condensate. The goal is the rigorous derivation of the time-dependent GP equation for the one-body condensate wave function. The strategy is to complete the scheme of derivation of the "method of hierarchies". ESY and AGT theorems are currently the most powerful tool to express rigorously the evolution of the one-body condensate wave function according to the cubic nonlinear Schrödinger equation. AGT asymptotic factorisation of marginals is now known to be equivalent to the same trace norm convergence at any time, as for the ESY case.

6. Correlations in the condensed state. Interparticle interaction establishes correlations in the many-body state. In the presence of condensation, a typical short scale structure is recognised to be present as a modification of the purely factorised $\varphi^{\otimes N}$ and it is also seen to be preserved in time. To control it, one tries to build interparticle correlations in a trial function Ψ_N^{trial}, since the true many-body state is not accessible in practice. Such a Ψ_N^{trial} is chosen to reproduce the limit $\gamma_N^{(1),\mathrm{trial}} \to |\varphi\rangle\langle\varphi|$ and, possibly to be close to the true many-body state both in norm and in energy. This is a subtle task: currently, a comprehension of the full correlation structure is missing because known trial functions fail to capture the energy of the many-body state.

Major open problems

Even in the basics of the mathematics of B.E.C. some major problems remain open (not to mention unsolved questions in other specific B.E.C.-related topics listed in the Introduction, which did not enter this work).

- ✓ Control of errors for finite N in the asymptotic results.
- ✓ Condensate vs thermal cloud, through the control of $\gamma_N^{(1)}$ before the limit.
- ✓ Fragmented condensation (namely, $\gamma_N^{(1)} \overset{N\to\infty}{\longrightarrow} \alpha_1|\varphi_1\rangle\langle\varphi_1| + \alpha_2|\varphi_2\rangle\langle\varphi_2|$).
- ✓ Control of the correlation structure of a condensate evolving in time.
- ✓ Relaxation of a number of simplifying hypotheses towards more singular interaction potentials.
- ✓ True δ-potentials in one dimension.
- ✓ Overlap between the GP and the TF scaling regime.
- ✓ Rigorous statistical treatment of B.E.C. at positive temperature.

As a matter of fact, great efforts are already ongoing in the Mathematical Physics community to address some of these problems.

References

[1] R. Adami, *On the derivation of the Gross-Pitaevskii equation*, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8), 8 (2005), pp. 359–368.

[2] R. Adami, C. Bardos, F. Golse, and A. Teta, *Towards a rigorous derivation of the cubic NLSE in dimension one*, Asymptot. Anal., 40 (2004), pp. 93–108.

[3] R. Adami, F. Golse, and A. Teta, *Rigorous derivation of the cubic NLS in dimension one*, J. Statist. Phys., 127 (2007), pp. 1193–1220.

[4] M. Aizenman, E. H. Lieb, R. Seiringer, J. P. Solovej, and J. Yngvason, *Bose-Einstein quantum phase transition in an optical lattice model*, Phys. Rev. A, 70 (2004), p. 023612.

[5] S. Albeverio, F. Gesztesy, R. Høegh-Krohn, and H. Holden, *Solvable Models in Quantum Mechanics*, Texts and Monographs in Physics, Springer-Verlag, New York, 1988.

[6] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor*, Science, 269 (1995), pp. 198–201.

[7] J. Arazy, *More on convergence in unitary matrix spaces*, Proc. Amer. Math. Soc., 83 (1981), pp. 44–48.

[8] W. Bao, D. Jaksch, and P. A. Markowich, *Numerical solution of the Gross-Pitaevskii equation for Bose-Einstein condensation*, J. Comp. Phys., 187 (2003), pp. 318–342.

[9] C. Bardos, L. Erdős, F. Golse, N. Mauser, and H.-T. Yau, *Derivation of the Schrödinger-Poisson equation from the quantum N-body problem*, C. R. Math. Acad. Sci. Paris, 334 (2002), pp. 515–520.

[10] C. Bardos, F. Golse, and N. J. Mauser, *Weak coupling limit of the N-particle Schrödinger equation*, Methods Appl. Anal., 7 (2000), pp. 275–293.

[11] C. Bardos and N. J. Mauser, *The weak coupling limit for systems of $N \to \infty$ quantum particles: state of the art and applications*, in Actes du CANUM, 2003.

[12] S. T. Beliaev, *Energy spectrum of a non-ideal Bose gas*, Sov. Phys. JETP, 7 (1958), pp. 299–307.

[13] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics*, John Wiley and Sons, New York, 1952, ch. II, p. 74.

[14] N. N. BOGOLUBOV, *Energy levels of the imperfect Bose-Einstein gas*, Bull. Moscow State Univ., 7 (1947), pp. 43–56.

[15] ——, *On the theory of superfluidity*, J. Phys. (USSR), 11 (1947), pp. 23–32.

[16] ——, *Problems of a Dynamical Theory in Statistical Physics*, in Studies in Statistical Mechanics, G. E. de Boer, J. and Uhlenbeck, ed., North-Holland, Amsterdam, 1962. English translation of J. Phys. (USSR), 10 (1946), pp. 256–265.

[17] M. BORN AND H. S. GREEN, *A General Kinetic Theory of Liquids*, Cambridge, at the University Press, 1949.

[18] S. N. BOSE, *Plancks Gesetz und Lichtquantenhypothese*, Z. Phys., 26 (1924), pp. 178–181.

[19] C. C. BRADLEY, C. A. SACKETT, J. J. TOLLETT, AND R. G. HULET, *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions*, Phys. Rev. Lett., 75 (1995), pp. 1687–1690.

[20] W. BRAUN AND K. HEPP, *The Vlasov dynamics and its fluctuations in the* $1/N$ *limit of interacting classical particles*, Comm. Math. Phys., 56 (1977), pp. 101–113.

[21] G. BREIT, *The scattering of slow neutrons by bound protons. I. Methods of calculation*, Phys. Rev., 71 (1947), pp. 215–231.

[22] C. BRISLAWN, *Kernels of trace class operators*, Proc. Amer. Math. Soc., 104 (1988), pp. 1181–1190.

[23] ——, *Traceable integral kernels on countably generated measure spaces*, Pacific J. Math., 150 (1991), pp. 229–240.

[24] J. B. BRU, M. CORREGGI, P. PICKL, AND J. YNGVASON, *The TF limit for rapidly rotating Bose gases in anharmonic traps*, 2007.

[25] K. A. BRUECKNER AND K. SAWADA, *Bose-Einstein gas with repulsive interactions: general theory*, Phys. Rev., 106 (1957), pp. 1117–1127.

[26] Y. CASTIN, *Bose-Einstein condensates in atomic gases: simple theoretical results*, in Lecture Notes of Les Houches Summer School, vol. VII C, EDP Sciences and Springer-Verlag, 2001, pp. 1–136.

[27] T. CAZENAVE, *Semilinear Schrödinger equations*, vol. 10 of Courant Lecture Notes in Mathematics, New York University Courant Institute of Mathematical Sciences, New York, 2003.

[28] C. CERCIGNANI, *The Boltzmann equation and its applications*, vol. 67 of Applied Mathematical Sciences, Springer-Verlag, New York, 1988.

[29] C. CERCIGNANI, R. ILLNER, AND M. PULVIRENTI, *The mathematical theory of dilute gases*, vol. 106 of Applied Mathematical Sciences, Springer-Verlag, New York, 1994.

[30] S. D. CHATTERJEE, *Biographical Memoirs*, in Satyendra Nath Bose, Indian Naional Science Academy, New Delhi, 1984, pp. 59–79.

[31] E. A. Cornell and C. E. Wieman, *Nobel lecture: Bose-Einstein condensation in a dilute gas, the first 70 years and some recent experiments*, Rev. Mod. Phys., 74 (2002), pp. 875–893.

[32] M. Correggi, T. Rindler-Daller, and J. Yngvason, *Rapidly rotating Bose-Einstein condensates in strongly anharmonic traps*, J. Math. Phys., 48 (2007), pp. 042104–042134.

[33] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, *Theory of Bose-Einstein condensation in trapped gases*, Rev. Mod. Phys., 71 (1999), pp. 463–512.

[34] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Bose-Einstein condensation in a gas of sodium atoms*, Phys. Rev. Lett., 75 (1995), pp. 3969–3973.

[35] G. F. Dell'Antonio, R. Figari, and A. Teta, *Hamiltonians for systems of N particles interacting through point interactions*, Ann. Inst. H. Poincaré Phys. Théor., 60 (1994), pp. 253–290.

[36] R. B. Dingle, *The zero-point of a system of particles*, Phil. Mag., 40 (1949), pp. 573–578.

[37] F. J. Dyson, *Ground-state energy of a hard-sphere gas*, Phys. Rev., 106 (1957), pp. 20–26.

[38] A. Einstein, *Quantentheorie des einatomigen idealen Gases*, Sitzber. Kgl. Preuss. Akad. Wiss., (1924), pp. 261–267.

[39] ——, *Quantentheorie des einatomigen idealen Gases, II. Abhandlung*, Sitzber. Kgl. Preuss. Akad. Wiss., (1925), pp. 3–14.

[40] A. Elgart, L. Erdős, B. Schlein, and H.-T. Yau, *Gross-Pitaevskii equation as the mean field limit of weakly coupled bosons*, Arch. Ration. Mech. Anal., 179 (2006), pp. 265–283.

[41] L. Erdős, B. Schlein, and H.-T. Yau, *Derivation of the Gross-Pitaevskii hierarchy for the dynamics of Bose-Einstein condensate*, Comm. Pure Appl. Math., 59 (2006), pp. 1659–1741.

[42] ——, *Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems*, Invent. Mat., 167 (2007), pp. 515–614.

[43] ——, *Rigorous derivation of the Gross-Pitaevskii equation*, Phys. Rev. Lett., 98 (2007), p. 040404.

[44] ——, *Derivation of the Gross-Pitaevskii equation for the dynamics of Bose-Einstein condensate*, math-ph/0606017 (2006).

[45] L. Erdős and H.-T. Yau, *Derivation of the nonlinear Schrödinger equation from a many body Coulomb system*, Adv. Theor. Math. Phys., 5 (2001), pp. 1169–1205.

[46] T. Esslinger, I. Bloch, and T. W. Hänsch, *Bose-Einstein condensation in a quadrupole-ioffe-configuration trap*, Phys. Rev. A, 58 (1998), pp. R2664–R2667.

[47] E. Fermi, Ricerca sci., 7 (1936), pp. 13–52.

[48] R. P. Feynman, *Application of Quantum Mechanics to liquid helium*, in Progress in Low Temperature Physics, vol. 1, Amsterdam: North-Holland, 1955, pp. 17–53.

[49] D. G. Fried, T. C. Killian, L. Willmann, D. Landhuis, S. C. Moss, D. Kleppner, and T. J. Greytak, *Bose-Einstein condensation of atomic hydrogen*, Phys. Rev. Lett., 81 (1998), pp. 3811–3814.

[50] S. A. Gaal, *Linear analysis and representation theory*, Springer-Verlag, New York, 1973, ch. VI.7, pp. 376–387. Die Grundlehren der mathematischen Wissenschaften, Band 198.

[51] J. Ginibre and G. Velo, *The classical field limit of scattering theory for nonrelativistic many-boson systems. I*, Comm. Math. Phys., 66 (1979), pp. 37–76.

[52] ——, *The classical field limit of scattering theory for nonrelativistic many-boson systems. II*, Comm. Math. Phys., 68 (1979), pp. 45–68.

[53] M. Girardeau and R. Arnowitt, *Theory of many-boson systems: Pair theory*, Phys. Rev., 113 (1959), pp. 755–761.

[54] F. Golse, *The mean-field limit for the dynamics of large particle systems*, in Journées "Équations aux Dérivées Partielles", Univ. Nantes, Nantes, 2003, pp. Exp. No. IX, 47.

[55] Hall, H. E. and Vinen, W. F., *The rotation of liquid helium II. I. experiments on the propagation of second sound in uniformly rotating helium II*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 238 (1956), pp. 204–214.

[56] C. E. Hecht, *The possible superfluid behaviour of hydrogen atom gases and liquids*, Physica, 25 (1959), pp. 1159–1161.

[57] K. Hepp, *The classical limit for quantum mechanical correlation functions*, Comm. Math. Phys., 35 (1974), pp. 265–277.

[58] K. Huang and C. N. Yang, *Quantum-mechanical many-body problem with hard-sphere interaction*, Phys. Rev., 105 (1957), pp. 767–775.

[59] N. M. Hugenholtz and D. Pines, *Ground-state energy and excitation spectrum of a system of interacting bosons*, Phys. Rev., 116 (1959), pp. 489–506.

[60] K. Husimi, *Some formal properties of the density matrix*, Proc. Phys. Math. Soc. Jpn., 22 (1940), pp. 264–314.

[61] R. Jastrow, *Many-body problem with strong forces*, Phys. Rev., 98 (1955), pp. 1479–1484.

[62] V. D. Kamble, *Bosons - The Birds That Flock and Sing Together*. http://www.vigyanprasar.gov.in/dream/jan2002/article1.htm.

[63] W. Ketterle, *Nobel lecture: when atoms behave as waves: Bose-Einstein condensation and the atom laser*, Rev. Mod. Phys., 74 (2002), pp. 1131–1151.

[64] W. Ketterle, D. S. Durfee, and D. M. Stamper-Kurn, *Making, probing and understanding Bose-Einstein condensates*, in Bose-Einstein Condensation in Atomic Gases, S. M. Inguscio, ed., vol. 140 of proceedings of the Enrico Fermi summer school in Varenna, 1999.

[65] J. G. Kirkwood, *The statistical mechanical theory of transport processes I. General theory*, The Journal of Chemical Physics, 14 (1946), pp. 180–201. Errata 14, pp. 347.

[66] ——, *The statistical mechanical theory of transport processes II. Transport in gases*, The Journal of Chemical Physics, 15 (1947), pp. 72–76.

[67] L. Landau, *The Theory of Superfluidity of Helium II*, J. Phys. USSR, 5 (1941), p. 71.

[68] L. D. Landau and E. M. Lifshitz, *Statisticheskaya Fizika*, Fizmatgiz, Moscow, 1951.

[69] T. D. Lee, K. Huang, and C. N. Yang, *Eigenvalues and eigenfunctions of a Bose system of hard spheres and its low-temperature properties*, Phys. Rev., 106 (1957), pp. 1135–1145.

[70] T. D. Lee and C. N. Yang, *Many-body problem in Quantum Statistical Mechanics. III. Zero-temperature limit for dilute hard spheres*, Phys. Rev., 117 (1960), pp. 12–21.

[71] A. J. Leggett, *Bose-Einstein condensation in the alkali gases: some fundamental concepts*, Rev. Mod. Phys., 73 (2001), pp. 307–356.

[72] E. H. Lieb, *Simplified approach to the ground-state energy of an imperfect Bose gas*, Phys. Rev., 130 (1963), pp. 2518–2528.

[73] ——, *The Bose fluid*, in Lecture notes in Theoretical Physics, vol. VII C, University of Colorado Press Kluwer Acad. Publ., 1964, p. 175.

[74] E. H. Lieb and R. Seiringer, *Proof of Bose-Einstein condensation for dilute trapped gases*, Phys. Rev. Lett., 88 (2002), p. 170409.

[75] E. H. Lieb, R. Seiringer, J. P. Solovej, and J. Yngvason, *The mathematics of the Bose gas and its condensation*, vol. 34 of Oberwolfach Seminars, Birkhäuser Verlag, Basel, 2005.

[76] ——, *The quantum-mechanical many-body problem: the Bose gas*, in Perspectives in analysis, vol. 27 of Math. Phys. Stud., Springer, Berlin, 2005, pp. 97–183.

[77] E. H. Lieb, R. Seiringer, and J. Yngvason, *Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional*, Phys. Rev. A, 61 (2000), p. 043602.

[78] ——, *A rigorous derivation of the Gross-Pitaevskii energy functional for a two-dimensional Bose gas*, Comm. Math. Phys., 224 (2001), pp. 17–31.

[79] E. H. LIEB AND J. YNGVASON, *Ground state energy of the low density Bose gas*, Phys. Rev. Lett., 80 (1998), pp. 2504–2507.

[80] ——, *The ground state energy of a dilute two-dimensional Bose gas*, J. Statist. Phys., 103 (2001), pp. 509–526.

[81] F. LONDON, *On the Bose-Einstein condensation*, Phys. Rev., 54 (1938), pp. 947–954.

[82] A. MICHELANGELI, *Analysis of interparticle correlations in dilute gases with B.E.C.*, SISSA digital library, 66/2007/MP (2007).

[83] ——, *Reduced density matrices and Bose-Einstein condensation*, SISSA digital library, 39/2007/MP (2007).

[84] ——, *Role of scaling limits in the rigorous analysis of Bose-Einstein condensation*, J. Math. Phys., 48 (2007).

[85] ——, *Strengthened convergence for reduced density matrices in the derivation of the Gross-Pitaevskiui equation*, SISSA digital library, 35/2007/MP (2007).

[86] A. MINGUZZI, S. SUCCI, F. TOSCHI, M. P. TOSI, AND P. VIGNOLO, *Numerical methods for atomic quantum gases with applications to Bose-Einstein condensates and to ultracold fermions*, Phys. Reports, 395 (2004), pp. 223–355.

[87] R. A. MINLOS AND L. D. FADDEEV, *Comment on the problem of three particles with point interactions*, Soviet Physics JETP, 14 (1962), pp. 1315–1316.

[88] ——, *On the point interaction for a three-particle system in Quantum Mechanics*, Soviet Physics Dokl., 6 (1962), pp. 1072–1074.

[89] G. MODUGNO, G. FERRARI, G. ROATI, R. J. BRECHA, A. SIMONI, AND M. INGUSCIO, *Bose-Einstein Condensation of Potassium Atoms by Sympathetic Cooling*, Science, 294 (2001), pp. 1320–1322.

[90] H. NARNHOFER AND G. L. SEWELL, *Vlasov hydrodynamics of a quantum mechanical model*, Comm. Math. Phys., 79 (1981), pp. 9–24.

[91] L. NIRENBERG, *An abstract form of the nonlinear Cauchy-Kowalewski theorem*, J. Differential Geometry, 6 (1972), pp. 561–576.

[92] T. NISIDA, *A note on a theorem of Nirenberg*, J. Differential Geometry, 12 (1978), pp. 629–633.

[93] L. ONSAGER, *Statistical hydrodynamics*, Nuovo Cimento (9), 6 (1949), pp. 279–287.

[94] O. PENROSE, *On the Quantum Mechanics of Helium II*, Philos. Mag., (1951).

[95] O. PENROSE AND L. ONSAGER, *Bose-Einstein Condensation and liquid Helium*, Phys. Rev., 104 (1956), pp. 576–584.

[96] C. Pethick and H. Smith, *Bose-Einstein Condensation in dilute gases*, Cambridge University Press, 2002.

[97] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation*, Clarendon Press, Oxford, 2003.

[98] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, vol. 1, New York Academic Press, 1972.

[99] R. Schatten, *Norm ideals of completely continuous operators*, Ergebnisse der Mathematik und ihrer Grenzgebiete. N. F., Heft 27, Springer-Verlag, Berlin, 1960.

[100] B. Schlein, *Dynamics of Bose-Einstein condensates*, arXiv.org:0704.0813 (2007).

[101] B. Simon, *Trace ideals and their applications*, vol. 35 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1979.

[102] ——, *Convergence in trace ideals*, Proc. Amer. Math. Soc., 83 (1981), pp. 39–43.

[103] H. Spohn, *Kinetic equations from Hamiltonian dynamics: Markovian limits*, Rev. Mod. Phys., 52 (1980), pp. 569–615.

[104] ——, *On the Vlasov hierarchy*, Math. Methods Appl. Sci., 3 (1981), pp. 445–455.

[105] L. Spruch and L. Rosenberg, *Upper bounds on scattering lengths for static potentials*, Phys. Rev., 116 (1959), pp. 1034–1040.

[106] S. Succi, F. Toschi, M. P. Tosi, and P. Vignolo, *Bose-Einstein condensates and the numerical solution of the Gross-Pitaevskii equation*, Computing in Science and Engineering, 07 (2005), pp. 48–57.

[107] F. G. Tricomi, *Integral equations*, Pure and Applied Mathematics. Vol. V, Interscience Publishers, Inc., New York, 1957, ch. 3.12, pp. 124–127.

[108] L. Vestergaard Hau, B. D. Busch, C. Liu, Z. Dutton, M. M. Burns, and J. A. Golovchenko, *Near-resonant spatial images of confined Bose-Einstein condensates in a 4-dee magnetic bottle*, Phys. Rev. A, 58 (1998), pp. R54–R57.

[109] J. Weiner, V. S. Bagnato, S. Zilio, and P. S. Julienne, *Experiments and theory in cold and ultracold collisions*, Rev. Mod. Phys., 71 (1999), pp. 1–85.

[110] T. T. Wu, *Ground state of a Bose system of hard spheres*, Phys. Rev., 115 (1959), pp. 1390–1404.

[111] V. I. Yukalov, *Principal problems in Bose-Einstein condensation of dilute gases*, Laser Phys. Lett., 1 (2004), pp. 435–461.

[112] J. Yvon, *La Théories Statistique des Fluides et l'Equation d'Etat*, in Actualités scientifiques et industrielles, no. 203, Hermann, Paris, 1935, p. 344.

[113] V. A. Zagrebnov and J.-B. Bru, *The Bogoliubov model of weakly imperfect Bose gas*, Phys. Reports, (2001), pp. 291–434.