Hartree–Fock Theory, Lieb's Variational Principle, and their Generalizations

Volker Bach <v.bach@tu-bs.de>,

Institut für Analysis und Algebra
TU Braunschweig
Universitätsplatz 2
38106 Braunschweig
Germany

30.03.2022

Dedicated to Elliott H. Lieb

MSC: 81-02, 81Q05, 81V45, 81V55, 81V74
Keywords: Coulomb Systems · Hartree–Fock · Quasifree States · Lieb's Variational Principle

I Hartree-Fock Theory of Coulomb Systems

One of the biggest triumphes of twentieth century science has been the discovery of quantum mechanics (and quantum field theory) almost one hundred years ago by Heisenberg [58], Born, Heisenberg, and Jordan [18], Schrödinger [89], Dirac [33, 34, 35], and Pauli [84]. It is a remarkable fact that, while quantum mechanics is of key importance for all technologies discovered in the past century and the complexity of theoretic descriptions of quantum systems has increased by several orders of magnitude, the basic conceptual framework of a complex Hilbert space $\mathfrak{H}$ of wave functions $\psi(t) \in \mathfrak{H}$ which represent physical states at time $t \in \mathbb{R}$ and evolve according to the (time-dependent) Schrödinger equation $i\dot{\psi}(t) = H\psi(t)$, with $H = H^*$ being the self-adjoint Hamiltonian operator, is unchanged until today. In absence of external fields, the Hamiltonian H is independent of time t. Then the solution of the time-dependent Schrödinger equation can be traced back to determining the spectral resolution, in particular, all eigenvalues $E \in \mathbb{R}$ and all corresponding eigenvectors $\psi_E \in \mathfrak{H}$ of H. Eigenvalues and corresponding eigenvectors do not cover all possible cases, and the general task to determine the spectral resolution of H has lead to the mathematical theory of spectral analysis of self-adjoint operators, see [86].

For a Coulomb system, i.e., a nonrelativistic atom ($K = 1$) or molecule ($K \geq 2$), consisting of $N \in \mathbb{Z}^+ := \{1, 2, 3, \ldots\}$ dynamical, mutually repelling electrons revolving about $K \in \mathbb{Z}^+$ attractive static nuclei of charges $\underline{Z} := (Z_1, Z_2, \ldots, Z_K) \in [\mathbb{R}_0^+]^K$ at pairwise distinct positions $\underline{R} := (\vec{R}_1, \vec{R}_2, \ldots, \vec{R}_K) \in [\mathbb{R}^3]^K$, the Hilbert space of a single electron is $\mathfrak{h} := L^2(\mathbb{R}^3 \times \{\uparrow, \downarrow\})$, and the Hilbert space

$$\mathfrak{H}^{(N)} \;:=\; \left\{ \Psi \in \mathfrak{h}^{\otimes N} \;\middle|\; \forall\, \pi \in S_N : \Psi(\underline{x}_\pi) \;=\; (-1)^\pi\, \Psi(\underline{x}) \right\} \tag{I.1}$$

of the wave function of the system of N dynamical electrons is the space of square-integrable functions of N coordinates $x_n = (\vec{x}_n, \tau_n) \in \mathbb{R}^3 \times \{\uparrow, \downarrow\}$ which are antisymmetric under permutations $\underline{x}_\pi = (x_{\pi(1)}, \ldots, x_{\pi(N)})$ of these coordinates $(x_1, \ldots, x_N)$. (Here and henceforth we follow the convention from physics and assume for Hilbert space always that $(\varphi, \psi) \mapsto \langle \varphi | \psi \rangle$ is antilinear in φ and linear in ψ.) The Hamiltonian generating the dynamics of these N electrons is

$$H_N(\underline{Z}, \underline{R}) \;:=\; \sum_{n=1}^{N} \left\{ -\Delta_n - \sum_{k=1}^{K} \frac{Z_k}{|\vec{x}_n - \vec{R}_k|} \right\} \;+\; \sum_{1 \leq m < n \leq N} \frac{1}{|\vec{x}_m - \vec{x}_n|}\,. \tag{I.2}$$

Note that the charges $\underline{Z}$ and the positions $\underline{R}$ of the nuclei enter H_N as fixed parameters, and $H_N(\underline{Z}, \underline{R})$ may be considered the Born–Oppenheimer approximation [19] to lowest order. We frequently omit to display the dependence of the Hamiltonian on $\underline{Z}$ and $\underline{R}$ and simply write $H_N \equiv H_N(\underline{Z}, \underline{R})$. The Hamiltonian H_N is

essentially self-adjoint and semibounded on the space $\mathcal{S}^{\wedge N} := \mathcal{S}^{\otimes N} \cap \mathfrak{H}^{(N)}$ of antisymmetric Schwartz test functions of N variables which is a dense subspace of $\mathfrak{H}^{(N)}$. (We henceforth largely ignore domain questions, omit to display $\mathcal{S}^{\wedge N}$, and implicitly assume sufficient regularity of the wave functions under consideration.)

The semiboundedness of H_N ensures the finiteness of the **ground state energy**, i.e., the infimum $\boldsymbol{E_{\rm gs}(N)} \equiv \boldsymbol{E_{\rm gs}(N, \underline{Z}, \underline{R})} := \inf \sigma(H_N)$ of the spectrum of H_N. The ground state energy $E_{\rm gs}(N)$ and, if $E_{\rm gs}(N)$ happens to be an eigenvalue, the corresponding **ground state** (eigenvector) $\boldsymbol{\Psi_{\rm gs}} \in \mathfrak{H}^{(N)}$ are basic quantities for the physical description of the Coulomb system. The actual solution of the corresponding eigenvalue equation $H_N \Psi_{\rm gs} = E_{\rm gs}(N) \Psi_{\rm gs}$, however, is inaccessible to explicit solution or even numerical computation for large molecules due to the large number of variables involved.

At this point the Rayleigh–Ritz principle becomes of key importance because it yields a variational characterization of

$$E_{\rm gs}(N) \;=\; \inf \Big\{ \langle \Psi \mid H_N \Psi \rangle \;\Big|\; \Psi \in \mathfrak{H}^{(N)},\; \|\Psi\| = 1 \Big\} \tag{I.3}$$

as the lowest energy expectation value the Hamiltonian H_N admits. Instead of solving the Schrödinger equation -which is virtually impossible- one computes the energy expectation value of any normalized trial state $\Psi_{\rm trial} \in \mathfrak{H}^{(N)}$. This yields an upper bound $\langle \Psi_{\rm trial} | H_N \Psi_{\rm trial} \rangle \geq E_{\rm gs}(N)$ on the ground state energy. If $\langle \Psi_{\rm trial} | H_N \Psi_{\rm trial} \rangle - E_{\rm gs}(N)$ is small, the trial state $\Psi_{\rm trial}$ is assumed to be a good approximation to (one of) the actual ground state(s) $\Psi_{\rm gs}$. The mathematical justification for this replacement, e.g., in terms of quantitative error bounds, is a difficult and largely open mathematical problem.

The earliest and, perhaps, most natural choice of trial states for Coulomb systems made is known as the **Hartree–Fock approximation**, which had been originally proposed by Hartree [57] but without an antisymmetry contraint on the wave function. This was followed by improvements of Fock [41] and Slater [93, 92], who took the antisymmetry of the trial state correctly into account. It is a variational principle in which the variation in (I.3) is restricted to **Slater determinants**, i.e., to wave functions of the form $\boldsymbol{\Phi(\underline{f})} := f_1 \wedge f_2 \wedge \cdots \wedge f_N$. These are antisymmetrized tensor products

$$f_1 \wedge f_2 \wedge \cdots \wedge f_N \;:=\; \frac{1}{\sqrt{N!}} \sum_{\pi \in \mathcal{S}_N} (-1)^\pi \, f_{\pi(1)} \otimes f_{\pi(2)} \otimes \cdots \otimes f_{\pi(N)} \tag{I.4}$$

of N-tuples $\underline{f} = (f_1, \ldots, f_N) \in \mathfrak{h}^N$ of mutually orthonormal **orbitals**, i.e., vectors $f_i \in \mathfrak{h}$ in the **one-particle Hilbert space**

$$\mathfrak{h} \;=\; L^2\big(\mathbb{R}^3 \times \{\uparrow, \downarrow\}\big)\,, \tag{I.5}$$

obeying $\langle f_i | f_j \rangle_{\mathfrak{h}} = \delta_{i,j}$ The corresponding infimum

$$\boldsymbol{E_{\mathbf{HF}}(N)} \;:=\; \inf\Big\{ \big\langle \Phi(\underline{f}) \,\big|\, H_N\, \Phi(\underline{f}) \big\rangle \;\Big|\; f_1, \dots, f_N \in\, \mathfrak{h}\,,\; \langle f_i | f_j \rangle_{\mathfrak{h}} = \delta_{i,j} \Big\} \quad \text{(I.6)}$$

is called the **Hartree–Fock ground state energy**. A straightforward computation gives

$$\begin{aligned} \mathcal{E}_{\mathrm{HF}}(\underline{f}) \;&:= \big\langle \Phi(\underline{f}) \,\big|\, H_N \Phi(\underline{f}) \big\rangle \qquad \text{(I.7)} \\ &= \sum_{i=1}^{N} \langle f_i \,|\, h f_i \rangle_{\mathfrak{h}} \;+\; \frac{1}{2} \sum_{i,j=1}^{N} \langle f_i \wedge f_j \,|\, V(f_i \wedge f_j) \rangle_{\mathfrak{h}\otimes\mathfrak{h}}\,, \end{aligned}$$

where the **one-particle operator** $\boldsymbol{h} := -\Delta - \sum_{k=1}^{K} Z_k\, |\vec{x} - \vec{R}_k|^{-1}$ is a second-order differential operator acting on (a suitable dense domain in) $\mathfrak{h}$, and the **pair interaction potential** $\boldsymbol{V} := |\vec{x} - \vec{y}|^{-1}$ is a multiplication operator on (a dense domain in) $\mathfrak{h} \otimes \mathfrak{h}$. The energy functional $\mathcal{E}_{\mathrm{HF}}(\underline{f})$ can be written as a sum

$$\mathcal{E}_{\mathrm{HF}}(\underline{f}) \;= T(\underline{f}) \;-\; U(\rho_{\underline{f}}) \;+\; \frac{1}{2} D(\rho_{\underline{f}}) \;-\; \frac{1}{2} X(\gamma_{\underline{f}}) \qquad \text{(I.8)}$$

of the **kinetic energy** $\boldsymbol{T(\underline{f})}$ minus the **nuclear attraction** $\boldsymbol{U(\rho_{\underline{f}})}$,

$$T(\underline{f}) \;:=\; \sum_{\tau=\uparrow,\downarrow} \int |\vec{\nabla} f_i(\vec{x},\tau)|^2 \, d^3x\,, \qquad U(\rho) \;:=\; \sum_{k=1}^{K} \int \frac{Z_k\, \rho(\vec{x})}{|\vec{x} - \vec{R}_k|}\, d^3x\,, \quad \text{(I.9)}$$

plus the **direct term** $\boldsymbol{\frac{1}{2} D(\rho_{\underline{f}})}$, representing the classical electrostatic energy, minus the **exchange term** $\boldsymbol{\frac{1}{2} X(\gamma_{\underline{f}})}$,

$$D(\rho) \;:=\; \iint \frac{\rho(\vec{x})\, \rho(\vec{y})\; d^3x\, d^3y}{|\vec{x} - \vec{y}|}\,, \quad X(\gamma) \;:=\; \iint \frac{|\gamma(\vec{x},\vec{y})|^2 \; d^3x\, d^3y}{|\vec{x} - \vec{y}|}\,, \quad \text{(I.10)}$$

where $\gamma_{\underline{f}}(\vec{x},\vec{y}) := \sum_{\tau=\uparrow,\downarrow} \sum_{i=1}^{N} f_i(\vec{x},\tau)\, \overline{f_i(\vec{y},\tau)}$ and $\rho_{\underline{f}}(\vec{x}) := \gamma_{\underline{f}}(\vec{x},\vec{x})$ are the one-particle density matrix and the one-particle density corresponding to $\Phi(\underline{f})$, respectively. The explicit and relatively simple form of these terms are one main reason for the success of the Hartree–Fock approximation.

For large neutral Coulomb Systems, i.e., for $Z = N \gg 1$, $\underline{Z} = Z\underline{z}$, with $\underline{z} = (z_1, \dots, z_k)$ for fixed $z_k > 0$ summing up to one, and nuclear positions $\underline{R}(Z) = \big(\vec{R}_1(Z), \dots, \vec{R}_k(Z)\big)$ not too close to each other, $\inf_{Z>0} \min_{k<\ell} \{ Z^{1/3} |\vec{R}_k - \vec{R}_\ell| \} > 0$, the Hartree–Fock energy is seen [3] to obey

$$E_{\mathrm{HF}}(Z) \;=\; E_{\mathrm{TF}}(Z, Z\underline{z}, \underline{R}(Z)) \;+\; \frac{Z^2}{4} \sum_{k=1}^{K} z_k^2 \;+\; \mathcal{O}\big(Z^{5/3}\big)\,, \qquad \text{(I.11)}$$

where the main contribution to leading order in Z is the **Thomas–Fermi energy** $\boldsymbol{E_{\mathrm{TF}}\big(Z, Z\underline{z}, \underline{R}(Z)\big)}$ established by Lieb and Simon in [71], which is bounded above and below by universal multiples of $Z^{7/3}$, followed by the **Scott correction** $\frac{\boldsymbol{Z^2}}{\boldsymbol{4}} \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \boldsymbol{z_k^2}$ of order Z^2 derived by Hughes [61] and by Siedentop and Weikard [90, 91] for atoms ($K = 1$), by Ivrii and Sigal [62] for molecules ($K \geq 1$) and by Solovej, Spitzer, and Sørensen in the relativistic and nonrelativistic case for both atoms and molecules [96, 95]. (See also the contribution of Siedentop to this volume.)

For our discussion we observe that if $\underline{f}^{(\mathrm{HF})}$ is a minimizer of $\mathcal{E}_{\mathrm{HF}}$ (or an *approximate minimizer*, i.e., $\mathcal{E}_{\mathrm{HF}}(\underline{f}^{(\mathrm{HF})}) \leq E_{\mathrm{HF}}(N) + \varepsilon$, for $\varepsilon > 0$ sufficiently small) under the orthonormality constraint $\langle f_i | f_j \rangle_{\mathfrak{h}} = \delta_{i,j}$, then there exist universal constants $0 < c < C < \infty$ such that, for any choice of $\underline{z} = (z_1, \ldots, z_k)$ and $\underline{R}(Z) = \big(\vec{R}_1(Z), \ldots, \vec{R}_k(Z)\big)$,

$$c\, Z^{7/3} \;\leq\; T(\underline{f}^{(\mathrm{HF})})\,,\; U(\rho_{\mathrm{HF}})\,,\; D(\rho_{\mathrm{HF}}) \;\leq\; C\, Z^{7/3}\,, \tag{I.12}$$

$$c\, Z^{5/3} \;\leq\; X(\gamma_{\mathrm{HF}}) \;\leq\; C\, Z^{5/3}\,, \tag{I.13}$$

where $\rho_{\mathrm{HF}} := \rho_{\underline{f}^{(\mathrm{HF})}}$ and $\gamma_{\mathrm{HF}} := \gamma_{\underline{f}^{(\mathrm{HF})}}$. That is, the kinetic energy, the nuclear attraction, and the classical electrostatic energy are all or the order $Z^{7/3}$, while the exchange energy is of order $Z^{5/3}$ and hence much smaller in magnitude.

The dominance of the three contributions $T(\underline{f}^{(\mathrm{HF})})$, $U(\rho_{\mathrm{HF}})$, and $D(\rho_{\mathrm{HF}})$ to the energy compared to the contribution of the exchange term $X(\gamma_{\mathrm{HF}})$ can be anticipated from the Cauchy-Schwarz inequality which implies that $X(\gamma_{\underline{f}}) \leq D(\rho_{\underline{f}})$, for any model with repulsive pair interaction $V(x-y) \geq 0$. Note, however, that this takes only *total* ground state energies of the entire system into account; if we compare energy *differences*, then the exchange contribution may become the decisive quantity that determines whether a system binds or not.

Furthermore, if $\underline{f}^{(\mathrm{HF})}$ is a minimizer of $\mathcal{E}_{\mathrm{HF}}$ under the orthonormality constraint $\langle f_i | f_j \rangle_{\mathfrak{h}} = \delta_{i,j}$ then $\mathcal{E}_{\mathrm{HF}}$ is stationary at $\underline{f}^{(\mathrm{HF})} = (f_1^{(\mathrm{HF})}, \ldots, f_N^{(\mathrm{HF})})$ and the Euler–Lagrange equations -known in this context as **Hartree–Fock equations**- become

$$h_{\mathrm{HF}}[\underline{f}^{(\mathrm{HF})}]\, f_i^{(\mathrm{HF})} \;=\; e_i\, f_i^{(\mathrm{HF})}\,, \tag{I.14}$$

for all $i \in \{1, \ldots, N\}$, where the eigenvalues e_i are Lagrange multipliers imposed to fulfill the orthonormality constraint and $\boldsymbol{h_{\mathrm{HF}}[\underline{f}^{(\mathrm{HF})}]}$ is the **Hartree–Fock ef-**

fective Hamiltonian acting on orbitals $g \in \mathfrak{h}$ as

$$\big(h_{\mathrm{HF}}[\underline{f}\,g\big)[\vec{x},\tau] \;:= \tag{I.15}$$
$$\big(h\,g\big)[\vec{x},\tau] \;+\; \bigg(\int \frac{\rho_{\underline{f}}(\vec{y})\,d^3y}{|\vec{x}-\vec{y}|}\bigg) g(\vec{x},\tau) \;-\; \int \frac{\gamma_{\underline{f}}(\vec{x},\vec{y})\,g(\vec{y},\tau)\,d^3y}{|\vec{x}-\vec{y}|}\,.$$

Even though the Hartree–Fock equations form a system of nonlinear partial integro-differential equations in $\underline{f}$, the reduction of N dynamical variables to two makes it accessible to numerical solution. Concrete numerical algorithms to solve the Hartree–Fock equations have been analyzed mathematically by Cancès and Le Bris in [23]. More recently, the numerical solution of the corresponding self-consistent equation of generalized Hartree–Fock theory described in Section VI, the Bogolubov–Hartree–Fock equations, has been studied by Lewin and Paul in [68].

Hartree–Fock theory is closely related to density functional theory and *Kohn–Sham (KS) theory*. These latter two are based on the Hohenberg–Kohn theorem [59] which asserts that the ground state energy of any Coulomb system can be expressed as the infimum of a universal (but unknown) functional of the electron density only. A mathematically precise formulation of the Hohenberg–Kohn theorem was given by Levy [65] and Lieb [72]. We describe the Kohn–Sham theory from the viewpoint of Hartree–Fock theory, although this oversimplifies their physical arguments somewhat. Namely, Kohn and Sham proposed to approximate the exchange term $X(\gamma_{\underline{f}})$ by a functional $\int G[\rho_{\underline{f}}(\vec{x})]\,d^3x$ of the one-particle density $\rho_{\underline{f}}$ only, where G is yet to be determined. A natural candidate for G is $G[\rho] = C_{\mathrm{Dirac}}\,\rho^{4/3}$, which has been proposed by Dirac in [32] and whose quality as an approximation to the exchange term has been analyzed in [4]. This approximation is known as the *local density approximation (LDA)* and the *KS-LDA* theory is widely and successfully used in numerical studies in material science. Its mathematical foundation including a proof of existence of minimizers of the Kohn–Sham energy functional and, hence, of solutions of the corresponding stationarity condition known as Kohn–Sham equations was given by Anantharaman and Cancès in [1]. An important improvement to the local density approximation defined by a function $G[\rho]$ is the *generalized gradient approximation (GGA)*. It accommodates an additional dependence of the exchange term function $G[\rho, \nabla\sqrt{\rho_{\underline{f}}}]$ on the gradient of the (square root of the) density, leading to KS-GGA theory. A very successful proposal for the form of $G[\rho, \nabla\sqrt{\rho_{\underline{f}}}]$ was made by Perdew, Burke, and Ernzerhof in [63] and is known as *PBE*.

The first mathematically rigorous treatment of the Hartree–Fock approximation and the corresponding Hartree–Fock equations was given by Lieb and Simon in [70]. By applying the so-called direct methods of the calculus of variations,

they prove the existence of a minimizer $\underline{f}^{(\mathrm{HF})}$ of $\mathcal{E}_{\mathrm{HF}}$, which then necessarly fulfills the Hartree–Fock equations, under the condition that the number $N-1$ of electrons minus one is strictly less than the total nuclear charge $Z := \sum_{k=1}^{K} Z_k$. This is a natural HVZ-type condition reflecting the fact that, if one electron is spatially separated far away from the nuclei, it is still attracted by a Coulomb force induced by a net charge $Z - N + 1 > 0$. This force binds this outer electron to the molecule and prevents its escape to infinity.

An important technical point in [70] is the conversion of the original orthonormality condition $\langle f_i | f_j \rangle_{\mathfrak{h}} = \delta_{i,j}$ into the equivalent statement $G(\underline{f}, \underline{f}) = \mathbf{1}$ on $\mathbb{C}^N$, where $G(\underline{f}, \underline{g})_{i,j} := \langle f_i | g_j \rangle_{\mathfrak{h}}$ denotes the Gram matrix of $\underline{f} = (f_1, \ldots, f_N), \underline{g} = (g_1, \ldots, g_N) \in \mathfrak{h}^N$. Lieb and Simon then observe that the minimization over $\underline{f}$'s obeying $G(\underline{f}, \underline{f}) = \mathbf{1}$ can be relaxed to the quadratic form inequality $0 \leq G(\underline{f}, \underline{f}) \leq \mathbf{1}$ without changing the minimum. This observation foreshadows Lieb's variational principle [69] formulated shortly after. As opposed to the set of $\underline{f}$'s obeying $G(\underline{f}, \underline{f}) = \mathbf{1}$, the set of $\underline{f}$'s obeying the weaker constraint $0 \leq G(\underline{f}, \underline{f}) \leq \mathbf{1}$ is weakly closed (in the appropriate topology) which is necessary for the application of weak lower semicontinuity.

The HVZ-type condition $E_{\mathrm{HF}}(N) < E_{\mathrm{HF}}(N-1)$ mentioned above is the key condition for the proof of existence of excited states of Coulomb systems in Hartree–Fock theory, too. As the latter leads to nonlinear Euler-Lagrange equations, the concept of excited state as a higher eigenvalue of a linear operator cannot be applied directly, but there is a natural notion for *excited states* in variational analysis, namely, stationary points of the functional under consideration for values strictly above the minimum. The first proof that such excited states exist was given by Lions in [74]. Building up on a contribution by Friesecke [44] and an earlier paper [66], Lewin proved in [67] the existence of infinitely many excited states below $E_{\mathrm{HF}}(N-1)$. The essential step is to prove that below $E_{\mathrm{HF}}(N-1)$, the the Hartree–Fock functional for Coulomb systems fulfills a suitable Palais–Smale condition. Moreover, opposed to [70] and [74], the proof in [66, 67] is entirely given in the space N electrons and does not use any positivity of the pair potential or its Fourier transform.

II Fock Space, Density Matrices, and Second Quantization

Before we turn to Lieb's variational principle we provide a convenient mathematical framework and introduce the second quantization.

Fock Space: We henceforth assume the one-particle Hilbert space $\mathfrak{h}$ to be a complex separable Hilbert space - not necessarily $L^2\big(\mathbb{R}^3 \times \{\uparrow,\downarrow\}\big)$ as specified in (I.5), although this is a good example to keep in mind. For $N \in \mathbb{Z}^+$, the ***N*-particle Hilbert space** is

$$\mathfrak{H}^{(N)} := \bigwedge^N \mathfrak{h} := \overline{\mathrm{span}\Big\{ f_1 \wedge \cdots \wedge f_N \;\Big|\; f_1,\ldots,f_N \in \mathfrak{h} \Big\}}^{\|\cdot\|} \subseteq \mathfrak{h}^{\otimes N}, \quad \text{(II.1)}$$

where $\overline{(\,\cdot\,)}^{\|\cdot\|}$ denotes norm closure. The **fermion Fock space** (over $\mathfrak{h}$) is defined to be the orthogonal sum

$$\mathfrak{F} \equiv \mathfrak{F}[\mathfrak{h}] := \bigoplus_{N=0}^{\infty} \mathfrak{H}^{(N)}, \quad \text{(II.2)}$$

where $\mathfrak{H}^{(0)} := \mathbb{C}\cdot\Omega$ is the one-dimensional **vacuum subspace** spanned by a unit vector Ω called the **vacuum vector**. The vacuum subspace represents the physical state of absence of any particle in the quantum system under consideration. The elements of $\mathfrak{F}$ are sequences $\Psi = (\psi_0, \psi_1, \psi_2, \ldots)$ with $\psi_N \in \mathfrak{H}^{(N)}$. If no confusion is possible, we henceforth consider $\mathfrak{H}^{(N)}$ a subspace of $\mathfrak{F}$ by identifying $\psi_N \in \mathfrak{H}^{(N)}$ with $(0,\ldots,0,\psi_N,0,\ldots) \in \mathfrak{F}$.

Second Quantization: We come to the second quantization of operators. Given $N \geq 2$ and three indices $i,j,k \in \{1,\ldots,N\}$, $i<j$, we define two unitary operators $\Pi_i^{(N)} \in \mathcal{U}(\mathfrak{h}^{\otimes N})$ and $\Pi_{i,j}^{(N)} \in \mathcal{U}(\mathfrak{h}^{\otimes N})$ by

$$\Pi_k^{(N)}[f_1 \otimes \cdots \otimes f_N] := f_k \otimes f_1 \otimes \cdots \otimes f_{k-1} \otimes f_{k+1} \otimes \cdots \otimes f_N, \quad \text{(II.3)}$$

$$\Pi_{i,j}^{(N)}[f_1 \otimes \cdots \otimes f_N] := \quad \text{(II.4)}$$
$$f_i \otimes f_j \otimes f_1 \otimes \cdots \otimes f_{i-1} \otimes f_{i+1} \otimes \cdots \otimes f_{j-1} \otimes f_{j+1} \otimes \cdots \otimes f_N .$$

Next, given a one-particle operator h on $\mathfrak{h}$ and a two-particle operator V on $\mathfrak{h}\wedge\mathfrak{h}$, we define the corresponding N-particle operators h_N, V_N and furthermore H_N on

$\mathfrak{H}^{(N)}$ by $h_0 := V_0 := 0$, $h_1 := h$, $V_1 := 0$, and

$$h_N \;:=\; \sum_{i=1}^{N} \big(\Pi_i^{(N)}\big)^* \big(h \otimes \mathbf{1}^{\otimes(N-1)}\big)\, \Pi_i^{(N)} \tag{II.5}$$

$$V_N \;:=\; 2 \sum_{1\le i<j\le N} \big(\Pi_{i,j}^{(N)}\big)^* \big(V \otimes \mathbf{1}^{\otimes(N-2)}\big)\, \Pi_{i,j}^{(N)}\,, \tag{II.6}$$

$$H_N \;:=\; h_N \;+\; \tfrac{1}{2} V_N\,. \tag{II.7}$$

Note that H_N agrees with the operator in (I.2) provided that $h = -\Delta_x - \sum_{k=1}^{K} Z_k |\vec{x} - \vec{R}_k|^{-1}$ and $V = |\vec{x} - \vec{y}|^{-1}$. Their **second quantizations** are the operators

$$\mathbb{h} \;:=\; \bigoplus_{N=0}^{\infty} h_N\,, \quad \mathbb{V} \;:=\; \bigoplus_{N=0}^{\infty} V_N\,, \quad \mathbb{H} \;:=\; \bigoplus_{N=0}^{\infty} H_N \;=\; \mathbb{h} + \tfrac{1}{2}\mathbb{V}\,, \tag{II.8}$$

acting on finite vectors [defined in (II.19)]. The question whether $\mathbb{H}$ extends to a semibounded quadratic form is subtle, in general. For Coulomb systems, however, Dyson and Lenard [36, 37] and Lieb and Thirring [73] have shown *stability of matter* to hold true, which in our context means that

$$\mathbb{H}_\mu \;:=\; \mathbb{H} \;-\; \mu\,\mathbb{N} \;\ge\; \mu\, Z \;-\; \sum_{1\le k<\ell\le K} Z_k\, Z_\ell\, |R_k - R_\ell|^{-1}\,, \tag{II.9}$$

as a quadratic form on $\mathfrak{F}$, provided that the **chemical potential** $\boldsymbol{\mu} < 0$ is chosen sufficiently small, where

$$\mathbb{N} \;:=\; \bigoplus_{N=0}^{\infty} N \cdot \mathbf{1}_{\mathfrak{H}^{(N)}} \tag{II.10}$$

is the **number operator** on $\mathfrak{F}$. (See also the contribution of Loss to this volume.)

Abstracting from this situation, we assume in the following the operator h to be essentially self-adjoint on a suitable dense domain $\mathfrak{s} \subseteq \mathfrak{h}$ and semibounded, so that $h(m) := h + m \ge 0$, for some sufficiently large constant $m \in \mathbb{R}$. Furthermore, the pair potential is assumed to be an infinitesimal perturbation of h, i.e., V is defined on $\mathfrak{s}$ and, for any $\varepsilon > 0$, there exists a constant $b_\varepsilon < \infty$, such that $\|Vf\|_\mathfrak{h} \le \varepsilon \|hf\|_\mathfrak{h} + b_\varepsilon \|f\|_\mathfrak{h}$ holds true for all $f \in \mathfrak{s}$.

Density Matrices: The energy expectation value $\langle \psi_N \,|\, H_N \psi_N \rangle$ of a state represented by an N-particle wave function $\psi_N \in \mathfrak{H}^{(N)}$ may be written as $\langle \Psi \,|\, \mathbb{H}\Psi \rangle$, where $\Psi = (0, \ldots, 0, \psi_N, 0, \ldots) \in \mathfrak{F}$ has only one non-vanishing component. Allowing for linear combinations, it can be extended to all finite vectors $\Psi \in \mathfrak{F}_{\text{fin}}$ of

sufficient regularity. For these, we can further rewrite $\langle \Psi \,|\, \mathbb{H}\Psi \rangle = \mathrm{Tr}_{\mathfrak{F}}\big(\mathbb{H}|\Psi\rangle\langle\Psi|\big)$, where $|\Psi\rangle\langle\Psi| \in \mathcal{B}(\mathfrak{F})$ denotes the rank-one orthogonal projection onto Ψ.

This suggests to further extend the notion of energy expectation value to all **density matrices**

$$\mathfrak{DM} \;:=\; \Big\{ \rho \in \mathcal{L}^1(\mathfrak{F}) \;\Big|\; \rho \geq 0\,,\;\; \mathrm{Tr}_{\mathfrak{F}}(\rho) \;=\; 1\,,\;\; \rho \text{ is even} \Big\}\,, \tag{II.11}$$

i.e., all even positive trace-class operators ρ on $\mathfrak{F}$ of unit trace. (Here and henceforth we use the convention that $a \geq 0$ includes the self-adjointness of an operator a.) Evenness of ρ means that $\langle \phi | \rho \psi \rangle = 0$, whenever $\phi \in \mathfrak{H}^{(m)}$ and $\psi \in \mathfrak{H}^{(n)}$ with $m - n$ odd. We remark that evenness of density matrices is a natural condition for fermion, but not for boson systems. Since a given density matrix $\rho \in \mathfrak{DM}$ is, in particular, self-adjoint and compact, it can be written in diagonal form as $\rho = \sum_{\nu=1}^{\infty} r_\nu \, |\Psi_\nu\rangle\langle\Psi_\nu|$, where $r_\nu \geq 0$ are its nonnegative eigenvalues, which sum up to one, and $\{\Psi_\nu\}_{\nu=1}^{\infty} \subseteq \mathfrak{F}$ is an orthonormal basis in $\mathfrak{F}$ of eigenvectors of ρ. If a density matrix $\rho \in \mathfrak{DM}$ obeys $\rho\mathbb{N} = \mathbb{N}\rho = N \cdot \rho$, for some $N \in \mathbb{Z}^+$, then ρ is called an $\boldsymbol{N}$**-particle density matrix**. These are collected in

$$\mathfrak{DM}^{(N)} \;:=\; \big\{ \rho \in \mathfrak{DM} \;\big|\; \rho\mathbb{N} = \mathbb{N}\rho = N \cdot \rho \big\}\,, \tag{II.12}$$

Note that $\rho \in \mathfrak{DM}^{(N)}$ if, and only if, all its eigenvectors belong to $\mathfrak{H}^{(N)}$. Moreover, the density matrices form a norm-closed convex subset $\mathfrak{DM} \subseteq \mathcal{L}^1(\mathfrak{F})$ in which rank-one orthogonal projections, such as $|\Psi\rangle\langle\Psi| \in \mathfrak{DM}$ above, are extremal points called **pure**. In particular, the orthogonal projection onto $\psi_N \in \mathfrak{H}^{(N)}$ considered above is a pure N-particle density matrix.

Density matrices (and, in general, states) are not only natural objects mathematically, but from a physics point of view they are also important conceptually: For any reasonable theoretical framework for the description of a physical system, the scenario that this system is a subsystem of a larger system ("the rest of the universe") ought to be built in. In the latter situation, however, density matrices resulting from projecting onto the subsystem are the natural physical states, not wave functions.

Equipped with the definitions of density matrices and N-particle density matrices above, the Rayleigh–Ritz principle (I.3) assumes the following abstract form:

$$E_{\mathrm{gs}}(N) \;=\; \inf \big\{ \mathrm{Tr}_{\mathfrak{F}}(\rho\,\mathbb{H}) \;\big|\; \rho \;\in\; \mathfrak{DM}^{(N)}\,,\; \langle \mathbb{H} \rangle_\rho < \infty \big\}\,, \tag{II.13}$$

where it is implicity assumed that H_N is bounded from below and we denote

$$\langle A \rangle_\rho \;:=\; \mathrm{Tr}_{\mathfrak{F}}(\rho^{1/2}\, A\, \rho^{1/2})\,. \tag{II.14}$$

Similarly, for a sufficiently small chemical potential $\mu < 0$ such that $\mathbb{H}_\mu$ is bounded below, we define the **total ground state energy**

$$E_{\rm gs} := \inf_{N\in\mathbb{Z}^+} E_{\rm gs}(N) = \inf\left\{ {\rm Tr}_{\mathfrak{F}}(\rho\,\mathbb{H}_\mu) \,\middle|\, \rho \in \mathfrak{DM},\ \langle \mathbb{H}\rangle_\rho < \infty \right\}. \quad \text{(II.15)}$$

Creation and Annihilation Operators: Next, we introduce creation operators. Fixing $f \in \mathfrak{h}$, we define $c^*(f) : \mathfrak{H}^{(0)} \to \mathfrak{H}^{(1)}$ by $c^*(f)\Omega := f$ and $c^*(f) : \mathfrak{H}^{(N)} \to \mathfrak{H}^{(N+1)}$, for $N \in \mathbb{Z}^+$, by

$$c^*(f)[f_1 \wedge \cdots \wedge f_N] := f \wedge f_1 \wedge \cdots \wedge f_N \quad \text{(II.16)}$$

and extension by linearity. One then easily checks that $c^*(f)$ extends by continuity to a bounded operator on $\mathfrak{F}$ of norm $\|c^*(f)\|_{\mathcal{B}(\mathfrak{F})} = \|f\|_{\mathfrak{h}}$, called **creation operator** $\boldsymbol{c^*(f)} \in \mathcal{B}(\mathfrak{F})$. We observe that

$$\mathfrak{H}^{(N)} = \overline{{\rm span}\left\{ c^*(f_1)\cdots c^*(f_N)\Omega \,\middle|\, f_1,\ldots,f_N \in \mathfrak{h} \right\}}^{\|\cdot\|} \quad \text{and} \quad \text{(II.17)}$$

$$\mathfrak{F} = \overline{{\rm span}\left\{ c^*(f_1)\cdots c^*(f_N)\Omega \,\middle|\, N\in\mathbb{Z}_0^+,\ f_1,\ldots,f_N \in \mathfrak{h} \right\}}^{\|\cdot\|}, \quad \text{(II.18)}$$

for $N \in \mathbb{Z}_0^+ := \{0,1,2,3,\ldots\}$. Note that if $\mathfrak{s} \subseteq \mathfrak{h}$ is a dense subspace then the **space**

$$\mathfrak{F}_{\rm fin}[\mathfrak{s}] := {\rm span}\left\{ c^*(f_1)\cdots c^*(f_N)\Omega \,\middle|\, N\in\mathbb{Z}_0^+,\ f_1,\ldots,f_N \in \mathfrak{s} \right\} \subseteq \mathfrak{F}[\mathfrak{h}] \quad \text{(II.19)}$$

of finite vectors in $\mathfrak{F}$ containing finite linear combinations of finite wedge-products of orbitals in $\mathfrak{s}$ is a convenient dense domain for second quantizations of one- and two-particle operators on $\mathfrak{F}$.

The adjoint $\boldsymbol{c(f)} := [c^*(f)]^* \in \mathcal{B}(\mathfrak{F})$ of $c^*(f)$ is called **annihilation operator**. We remark that, while $\mathfrak{h} \ni f \mapsto c^*(f) \in \mathcal{B}(\mathfrak{F})$ is linear, the map $\mathfrak{h} \ni f \mapsto c(f) \in \mathcal{B}(\mathfrak{F})$ is antilinear. Moreover, one easily checks that the family $\{c^*(f), c(f)\}_{f\in\mathfrak{h}} \subseteq \mathcal{B}(\mathfrak{F})$ of creation and annihilation operators define a **Fock representation** of the **canonical anticommutation relations (CAR)**, i.e., it fulfills

$$\left\{ c(f), c^*(g) \right\} = \langle f|g\rangle_{\mathfrak{h}} \cdot \mathbf{1}_{\mathfrak{F}} \quad \text{and} \quad c(f)\Omega = 0\,, \quad \text{(II.20)}$$

for all $f, g \in \mathfrak{h}$, where $\{a,b\} := ab + ba$ is the anticommutator of two operators a and b.

The second-quantized operators $\mathbb{h}$, $\mathbb{V}$, $\mathbb{H}$, and $\mathbb{N}$ have a convenient representation in terms of creation and annihilation operators. Assuming that $\mathfrak{h}$ is infinite-dimensional and that $\{f_j\}_{j=1}^{\infty} \subseteq \mathfrak{s} \subseteq \mathfrak{h}$ is an orthonormal basis of $\mathfrak{h}$ of sufficiently regular functions lying in a dense subspace $\mathfrak{s}$, we have that

$$\mathbb{N} \;=\; \sum_{j=1}^{\infty} c^*(f_j)\, c(f_j)\,, \quad \mathbb{h} \;=\; \sum_{j,k=1}^{\infty} \langle f_j |\, h f_k \rangle\, c^*(f_j)\, c(f_k)\,, \tag{II.21}$$

$$\mathbb{V} \;=\; \sum_{i,j,k,\ell=1}^{\infty} \langle f_i \otimes f_j |\, V(f_k \otimes f_\ell) \rangle\, c^*(f_i)\, c^*(f_j)\, c(f_\ell)\, c(f_k)\,, \tag{II.22}$$

Reduced Density Matrices: We come to defining reduced density matrices - the central object of this paper. Given a density matrix $\rho \in \mathfrak{DM}$ with finite expectation value $\langle \mathbb{N} \rangle_\rho < \infty$ of $\mathbb{N}$, we define its **reduced one-particle density matrix (1-RDM) $\boldsymbol{\gamma_\rho^{(1)}}$** as a linear operator on $\mathfrak{h}$. , Likewise, if $\langle \mathbb{N}^2 \rangle_\rho < \infty$ we define its **reduced two-particle density matrix (2-RDM) $\boldsymbol{\gamma_\rho^{(2)}}$** as a linear operator on $\mathfrak{h} \otimes \mathfrak{h}$ by their matrix elements

$$\big\langle g \,\big|\, \gamma_\rho^{(1)} f \big\rangle_{\mathfrak{h}} \;:= \mathrm{Tr}_{\mathfrak{F}}\big[\rho\, c^*(f)\, c(g)\big]\,, \tag{II.23}$$

$$\big\langle g_1 \otimes g_2 \,\big|\, \gamma_\rho^{(2)} (f_1 \otimes f_2) \big\rangle_{\mathfrak{h}\otimes\mathfrak{h}} \;:= \mathrm{Tr}_{\mathfrak{F}}\big[\rho\, c^*(f_1)\, c^*(f_2)\, c(g_2)\, c(g_2)\big]\,, \tag{II.24}$$

for all $f, g, f_1, f_2, g_1, g_2 \in \mathfrak{h}$. These definitions are meaningful because it turns out that they define trace-class operators, as the following lemma asserts.

Lemma 1. *Let $\rho \in \mathfrak{DM}$ be a density matrix of finite particle number variance $\langle \mathbb{N}^2 \rangle_\rho := \mathrm{Tr}_{\mathfrak{F}}[\rho\, \mathbb{N}^2] < \infty$, and define its 1-RDM $\gamma_\rho^{(1)}$ and its 2-RDM $\gamma_\rho^{(2)}$ by* (II.23) *and* (II.24)*, respectively. Then $\gamma_\rho^{(1)}$ and $\gamma_\rho^{(2)}$ possess the following properties:*

(i) The operators $\gamma_\rho^{(1)} \in \mathcal{L}^1(\mathfrak{h})$ and $\gamma_\rho^{(2)} \in \mathcal{L}^1(\mathfrak{h} \otimes \mathfrak{h})$ are positive trace-class operators of trace

$$\mathrm{Tr}_{\mathfrak{h}}[\gamma_\rho^{(1)}] \;=\; \big\langle \mathbb{N} \big\rangle_\rho \quad \textit{and} \quad \mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}[\gamma_\rho^{(2)}] \;=\; \big\langle \mathbb{N}^2 - \mathbb{N} \big\rangle_\rho\,. \tag{II.25}$$

(ii) As quadratic forms,

$$0 \;\leq\; \gamma_\rho^{(1)} \;\leq\; \mathbf{1}_{\mathfrak{h}} \quad \textit{and} \quad 0 \;\leq\; \gamma_\rho^{(2)} \;\leq\; (N-1)\, \mathbf{1}_{\mathfrak{h}\otimes\mathfrak{h}}\,. \tag{II.26}$$

(iii) Suppose that h and V are semibounded and $\langle \mathbb{h} \rangle_\rho, \langle \mathbb{V} \rangle_\rho < \infty$. Then

$$\big\langle \mathbb{H} \big\rangle_\rho \;=\; \mathcal{E}_Q\big(\gamma_\rho^{(1)}, \gamma_\rho^{(2)}\big) \;:=\; \mathrm{Tr}_{\mathfrak{h}}\big[h\, \gamma_\rho^{(1)}\big] \;+\; \tfrac{1}{2} \mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\, \gamma_\rho^{(2)}\big]\,. \tag{II.27}$$

We finally remark that, in case of an N-particle density matrix $\rho \in \mathfrak{DM}^{(N)}$, for some $N \geq 2$, the corresponding 1-RDM $\gamma_\rho^{(1)}$ can be obtained from its 2-RDM $\gamma_\rho^{(2)}$ by taking a partial trace,

$$\big\langle g \,\big|\, \gamma_\rho^{(1)} f \big\rangle_{\mathfrak{h}} = \frac{1}{N-1} \sum_{j=1}^{\infty} \big\langle g \otimes f_j \,\big|\, \gamma_\rho^{(2)} (f \otimes f_j) \big\rangle_{\mathfrak{h}\otimes\mathfrak{h}} \,, \tag{II.28}$$

where $\{f_j\}_{j=1}^{\infty} \subseteq \mathfrak{h}$ is an orthonormal basis.

III Lieb's Variational Principle

The formulation of the Hartree–Fock approximation in Section I turns out to be too rigid and inconvenient, mathematically. A more flexible formulation is provided by Lieb's variational principle which uses the reduced one-particle density matrices introduced in (II.23).

We first link the 1-RDM $\gamma_\rho^{(1)}$ corresponding to a density matrix $\rho \in \mathfrak{DM}$ to the one-particle density matrix $\gamma_{\underline{f}}$ corresponding to an orthonormal family $\underline{f} = (f_1, \dots, f_N) \in \mathfrak{h}^N$ of N orbitals that enter the Hartree–Fock energy functional $\mathcal{E}_{\mathrm{HF}}(\underline{f})$ in (I.7). In fact, for pure states we have the following important relation between these operators:

Lemma 2. *For $N \geq 2$, let $\underline{f} = (f_1, \dots, f_N) \in \mathfrak{h}^N$ with $G(\underline{f}, \underline{f}) = \mathbf{1}$ and assume that $\rho(\underline{f}) = |\Phi(\underline{f})\rangle\langle\Phi(\underline{f})| \in \mathfrak{DM}^{(N)}$ is the orthogonal projection onto the Slater determinant $\Phi(\underline{f}) = f_1 \wedge \cdots \wedge f_N$. Then the following statements hold true.*

(i) The reduced one-particle density matrix of $\rho(\underline{f})$ is the rank-N orthogonal projection

$$\gamma^{(1)}_{\rho(\underline{f})} \;=\; \sum_{i=1}^{N} |f_i\rangle\langle f_i| \tag{III.1}$$

onto the subspace of $\mathfrak{h}$ spanned by $\{f_1, \dots, f_N\}$.

(ii) The reduced two-particle density matrix of $\rho(\underline{f})$ is twice the orthogonal projection of rank $\frac{1}{2}N(N-1)$,

$$\gamma^{(2)}_{\rho(\underline{f})} \;=\; \sum_{i,j=1}^{N} |f_i \wedge f_j\rangle\langle f_i \wedge f_j| \;=\; \big(\mathbf{1} - \mathrm{Ex}\big)\big(\gamma^{(1)}_{\rho(\underline{f})} \otimes \gamma^{(1)}_{\rho(\underline{f})}\big)\,, \tag{III.2}$$

*onto the subspace of $\mathfrak{h} \otimes \mathfrak{h}$ spanned by $\{f_i \wedge f_j | 1 \leq i < j \leq N\}$. Here, $\mathbf{Ex} \in \mathcal{U}(\mathfrak{h} \otimes \mathfrak{h})$ denotes the **exchange operator** $f \otimes g \mapsto g \otimes f$*

Inserting (III.2) into (II.27), we immediately obtain

$$\begin{aligned}\langle \Phi(\underline{f}) \,|\, H_N \Phi(\underline{f})\rangle \;&= \mathcal{E}_Q\big(\gamma^{(1)}_{\rho(\underline{f})},\; (\mathbf{1} - \mathrm{Ex})(\gamma^{(1)}_{\rho(\underline{f})} \otimes \gamma^{(1)}_{\rho(\underline{f})})\big) \\ &= \mathrm{Tr}_{\mathfrak{h}}\big[h\, \gamma^{(1)}_{\rho(\underline{f})}\big] \;+\; \frac{1}{2}\,\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\,(\mathbf{1} - \mathrm{Ex})(\gamma^{(1)}_{\rho(\underline{f})} \otimes \gamma^{(1)}_{\rho(\underline{f})})\big]\,,\end{aligned} \tag{III.3}$$

if the energy expectation is evaluated on a Slater determinant $\Phi(\underline{f}) = f_1 \wedge \cdots \wedge f_N$. The right side of (III.3) is entirely expessed in terms of the rank-N orthogonal projection $\gamma^{(1)}_{\rho(\underline{f})}$, and no other property than that enters the functional. That is,

$$E_{\mathrm{HF}}(N) \;=\; \inf\Big\{\mathcal{E}_{\mathrm{HF}}(\gamma)\;\Big|\;\gamma = \gamma^* \in \mathcal{L}^1(\mathfrak{h})\,,\; \mathrm{Tr}(\gamma) = N\,,\; \gamma = \gamma^2\Big\}\,, \tag{III.4}$$

where

$$\mathcal{E}_{\rm HF}(\gamma) \;=\; {\rm Tr}_{\mathfrak{h}}[h\,\gamma] \;+\; \frac{1}{2}\,{\rm Tr}_{\mathfrak{h}\otimes\mathfrak{h}}[V\,(\mathbf{1}-{\rm Ex})(\gamma\otimes\gamma)]\,. \tag{III.5}$$

Lieb's variational principle [69] asserts that the projection property $\gamma = \gamma^2$ in (III.4) can be relaxed to $0 \leq \gamma \leq \mathbf{1}_{\mathfrak{h}}$ without changing the infimum of the functional.

Theorem 3 (Lieb's Variational Principle)**.** *For $N \geq 2$,*

$$E_{\rm HF}(N) \;=\; \inf\Big\{\mathcal{E}_{\rm HF}(\gamma)\;\Big|\;\gamma\;\in\mathcal{L}^1(\mathfrak{h})\,,\;{\rm Tr}(\gamma)=N\,,\;0\leq\gamma\leq\mathbf{1}_{\mathfrak{h}}\Big\}\,. \tag{III.6}$$

Proof. We define the auxiliary energy $E_{\rm aux}(N)$ to be the infimium on the right side in (III.6) and observe that, clearly, $E_{\rm aux}(N) \leq E_{\rm HF}(N)$. We make three simplifying assumptions which are not essential for the validity of Theorem 3 and can be avoided by suitable limiting arguments. The first is the strict positivity of $\langle\psi|V\psi\rangle > 0$, for nonvanishing ψ, as opposed to merely assuming $\langle\psi|V\psi\rangle \geq 0$. The second simplifying assumption we make is that the infimum $E_{\rm aux}(N)$ is actually a minimum. That is, $E_{\rm aux}(N) = \mathcal{E}_{\rm HF}(\gamma_0)$ is attained by a minimizer γ_0 which fulfills $0 \leq \gamma_0 \leq \mathbf{1}$ and ${\rm Tr}(\gamma_0) = N$. Since γ_0 is compact, there exist an orthonormal basis $\{f_i\}_{i=1}^{\infty} \subseteq \mathfrak{h}$ of eigenvectors of γ_0 with corresponding (not necessarily distinct) eigenvalues $\lambda_i \in [0,1]$ that sum up to N. The third assumption we make is that γ_0 is of finite rank $J < \infty$, so that $\lambda_1, \ldots, \lambda_J > 0$,

$$\gamma_0 \;=\; \sum_{j=1}^{J}\lambda_j|f_j\rangle\langle f_j| \quad\text{and}\quad \mathcal{E}_{\rm HF}(\gamma_0) \;=\; \sum_{i=1}^{J}\lambda_i\,h_i \;+\; \frac{1}{2}\sum_{i,j=1}^{J}\lambda_i\,\lambda_j\,V_{i,j}\,, \tag{III.7}$$

where $h_i := \langle f_i|hf_i\rangle$ and $V_{i,j} := \langle f_i \wedge f_j|V(f_i\wedge f_j)\rangle > 0$.

Before we turn to Lieb's original proof in [69] we sketch the proof that $E_{\rm aux}(N) \geq E_{\rm HF}(N)$ given in [3], which, however, takes a different perspective. First note that it suffices to show that $\gamma_0 = \gamma_0^2$ is a projection. To this end we assume that γ_0 is not a projection and derive a contradiction from this assumption. If γ_0 is not a projection then there are (at least) two indices $p, q \in \mathbb{Z}^+$, $p < q$ such that $\lambda_p, \lambda_q \in (0,1)$ because the sum $\sum_{j=1}^{J}\lambda_j = N$ is an integer. We set $r := \min\{\lambda_p, 1-\lambda_p, \lambda_q, 1-\lambda_q\} > 0$ and $I := [-r, r]$ and observe that $\lambda_p + \delta, \lambda_q + \delta \in [0,1]$, for any $\delta \in I$. We define

$$\gamma_\delta \;:=\; (\lambda_p+\delta)|f_p\rangle\langle f_p| \;+\; (\lambda_q-\delta)|f_q\rangle\langle f_q| \;+\; \sum_{j\in\mathbb{Z}^+\setminus\{p,q\}}\lambda_j|f_j\rangle\langle f_j|\,. \tag{III.8}$$

Then $0 \leq \gamma_\delta \leq \mathbf{1}$ and ${\rm Tr}(\gamma_\delta) = N$, so γ_δ is admissible for any $\delta \in I$. Moreover $\gamma_0 = \frac{1}{2}\gamma_\delta + \frac{1}{2}\gamma_{-\delta}$. A simple computation using that $V_{p,q} > 0$ shows the strict

concavity of $I \ni \delta \mapsto \mathcal{E}_{\rm HF}(\gamma_\delta)$. Hence, $\min\{\mathcal{E}_{\rm HF}(\gamma_r), \mathcal{E}_{\rm HF}(\gamma_{-r})\} < \mathcal{E}_{\rm HF}(\gamma_0)$, which contradicts the assumption that γ_0 is a minimizer of $\mathcal{E}_{\rm HF}$. It follows that γ_0 is a projection, indeed. Note that this proof is constructive in the sense that, fixing the orthonormal orbitals $f_1, \ldots, f_J$, it defines an algorithm to find the γ_* of minimal energy $\mathcal{E}_{\rm HF}(\gamma_*)$ among all rank-J operators of the form $\gamma(\tau_1, \ldots, \tau_J) = \sum_{j=1}^J \tau_j |f_j\rangle\langle f_j|$ with $0 \leq \tau_j \leq 1$ and $\sum_{j=1}^J \tau_j = N$.

We now turn to Lieb's proof of Theorem 3 in [69], starting from (III.7). Its heart is a lemma that, under the assumption that $\lambda_1, \ldots, \lambda_J > 0$ and $\lambda_1 + \ldots + \lambda_J = N$, asserts the existence of N orthonormal vectors $G^{(1)}, \ldots, G^{(N)} \in \mathbb{C}^J$ which fulfill $\sum_{n=1}^N |G_j^{(n)}|^2 = \lambda_j$, for all $j \in \{1, \ldots, J\}$. We omit its interesting proof. Given these vectors $G^{(1)}, \ldots, G^{(N)}$, Lieb defines

$$g_n^{(\theta)} := \sum_{j=1}^J e^{2\pi i \theta_j} \, G_j^{(n)} \, f_j \in \mathfrak{h} , \tag{III.9}$$

for all $n \in \{1, \ldots, N\}$ and any choice $\theta := (\theta_j)_{j=1}^J \in [0,1)^J$ of phases $\theta_1, \ldots, \theta_J$. Using the orthonormality of $\{f_1, \ldots, f_J\} \subseteq \mathfrak{h}$ it is easy to check that the set $\{g_1^{(\theta)}, \ldots, g_N^{(\theta)}\} \subseteq \mathfrak{h}$ is orthonormal, too. That is, $G(\underline{g}^{(\theta)}, \underline{g}^{(\theta)}) = \mathbf{1}$, where $\underline{g}^{(\theta)} = (g_1^{(\theta)} \ldots, g_N^{(\theta)}) \in \mathfrak{h}^N$, and the corresponding Slater determinant is $\Phi(\underline{g}^{(\theta)}) = g_1^{(\theta)} \wedge \cdots \wedge g_N^{(\theta)} \in \mathfrak{H}^{(N)}$. The energy expectation value of this Slater determinant is

$$\begin{aligned} &\big\langle \Phi(\underline{g}^{(\theta)}) \, \big| \, H_N \Phi(\underline{g}^{(\theta)}) \big\rangle \qquad \text{(III.10)} \\ &= \sum_{n=1}^N \langle g_n^{(\theta)} \, | \, h g_n^{(\theta)} \rangle_{\mathfrak{h}} + \frac{1}{2} \sum_{m,n=1}^N \langle g_m^{(\theta)} \wedge g_n^{(\theta)} \, | \, V(g_m^{(\theta)} \wedge g_n^{(\theta)}) \rangle_{\mathfrak{h} \otimes \mathfrak{h}} \, . \end{aligned}$$

This energy expectation value is now averaged over all possible choices of θ by integrating over $[0,1)^J$. That is, for any integrable function $F \in L^1([0,1)^J)$ we write $\mathbb{E}_\theta[F] := \int_0^1 \cdots \int_0^1 F(\theta) \, d\theta_1 \cdots d\theta_J$. Using that $\mathbb{E}_\theta[e^{2\pi i(\Theta_j - \Theta_k)}] = \delta_{j,k}$ and (III.9), Lieb obtains

$$\begin{aligned} \mathbb{E}_\theta \Big[\sum_{n=1}^N \langle g_n^{(\Theta)} \, | \, h g_n^{(\Theta)} \rangle_{\mathfrak{h}} \Big] &= \sum_{n=1}^N \sum_{j,k=1}^J \mathbb{E}_\theta \big[e^{2\pi i(\Theta_j - \Theta_k)} \big] \, G_j^{(n)} \, \overline{G_k^{(n)}} \, \langle f_j \, | \, h f_k \rangle_{\mathfrak{h}} \\ &= \sum_{j=1}^J \bigg(\sum_{n=1}^N |G_j^{(n)}|^2 \bigg) \langle f_j \, | \, h f_j \rangle_{\mathfrak{h}} = \sum_{j=1}^J \lambda_j \, h_j \, . \end{aligned} \tag{III.11}$$

Similarly, if $i \neq j$ and $k \neq \ell$ then $\mathbb{E}_\theta[e^{2\pi i(\Theta_i + \Theta_j - \Theta_k - \Theta_\ell)}] = (\delta_{i,k} \, \delta_{j,\ell} + \delta_{i,\ell} \, \delta_{j,k})$,

and this implies that

$$\mathbb{E}_\theta\Big[\sum_{m,n=1}^{N}\langle g_m^{(\Theta)}\wedge g_n^{(\Theta)}\,|\,V(g_m^{(\Theta)}\wedge g_n^{(\Theta)})\rangle_{\mathfrak{h}\otimes\mathfrak{h}}\Big] \tag{III.12}$$

$$\begin{aligned}
&= \sum_{m,n=1}^{N}\sum_{i,j,k,\ell=1}^{J}\mathbb{E}_\theta\big[e^{2\pi i(\Theta_i+\Theta_j-\Theta_k-\Theta_\ell)}\big]\\
&\qquad G_i^{(m)}\,G_j^{(n)}\,\overline{G_k^{(m)}\,G_\ell^{(n)}}\,\langle f_i\wedge f_j\,|\,V(f_k\wedge f_\ell)\rangle_{\mathfrak{h}\otimes\mathfrak{h}}\\
&= \sum_{m,n=1}^{N}\sum_{i,j=1}^{J}\Big(|G_i^{(m)}|^2\,|G_j^{(n)}|^2\;-\;G_j^{(m)}\,G_i^{(n)}\,\overline{G_i^{(m)}\,G_j^{(n)}}\Big)\,V_{i,j}\\
&= \sum_{i,j=1}^{J}\lambda_i\,\lambda_j\,V_{i,j}\;-\;\sum_{i,j=1}^{J}\Big|\sum_{n=1}^{N}G_i^{(n)}\,\overline{G_j^{(n)}}\Big|^2\,V_{i,j}\;\leq\;\sum_{i,j=1}^{J}\lambda_i\,\lambda_j\,V_{i,j}\,.
\end{aligned}$$

Here, the positivity $V\geq 0$ is crucial, see also (IV.2). Adding up (III.11) and half of (III.12), Lieb arrives at

$$\mathbb{E}_\theta\big[\langle\Phi(\underline{g}^{(\Theta)})\,|\,H_N\Phi(\underline{g}^{(\Theta)})\rangle\big]\;\leq\;\mathcal{E}_{\mathrm{aux}}(\gamma_0)\,. \tag{III.13}$$

Since $\mathbb{E}_\theta$ is an average, Eq. (III.13) implies that there is at least one choice of $\theta\in[0,1)^J$, for which $\langle\Phi(\underline{g}^{(\theta)})\big|\,H_N\Phi(\underline{g}^{(\theta)})\rangle\leq\mathcal{E}_{\mathrm{aux}}(\gamma_0)$. Thus, we finally have $E_{\mathrm{HF}}(N)\leq E_{\mathrm{aux}}(N)$. □

Lieb's variational principle is a formulation of the Hartree–Fock approximation in the natural variable γ. It justifies the introduction of the notion of a **one-particle density matrix** as any self-adjoint trace-class operator

$$\boldsymbol{\gamma}\in\mathcal{L}^1(\mathfrak{h})\quad\text{that obeys}\quad 0\leq\gamma\leq\mathbf{1}_{\mathfrak{h}}\,, \tag{III.14}$$

leaving aside the question whether it is the reduced one-particle density matrix $\gamma=\gamma_\rho^{(1)}$ corresponding to some density matrix $\rho\in\mathfrak{DM}$. We come back to this point in the next section. The one-particle density matrices form a norm-closed, and hence weakly closed, convex subset of $\mathcal{L}^1(\mathfrak{h})$ which makes them suitable for variational analysis.

Lieb's variational principle asserts, briefly speaking, that among one-particle density matrices obeying (III.14) and of trace N the ones with lowest energy are the rank-N projections. Under the assumption of the existence of a minimizer γ_{HF}, this conclusion also follows from the stationarity of the Hartree–Fock functional at $\mathcal{E}_{\mathrm{HF}}$ at γ_{HF}. In fact, the Hartree–Fock equations (I.14) turn into the following **self-consistent equation**:

$$\gamma_{\mathrm{HF}}\;=\;\mathbf{1}_N\big(h_{\mathrm{HF}}[\gamma_{\mathrm{HF}}]\big)\,, \tag{III.15}$$

where $\mathbf{1}_N(A)$ denotes the projection onto the lowest N eigenvalues, counting multiplicities, for a self-adjoint operator A. In other words, $\mathbf{1}_N(A)$ is the projection onto a subspace of dimension N such that $\mathrm{Tr}_{\mathfrak{h}}[A\,\mathbf{1}_N(A)]$ is minimal. (If a minimizer exists, this subspace is actually unique, as follows from the unfilled-shell theorem of Lieb, Loss, Solovej, and the author [11].) Furthermore, $h_{\mathrm{HF}}[\gamma]$ is the corresponding form of the Hartree–Fock effective Hamiltonian, acting on orbitals $g \in \mathfrak{h}$ as

$$\big(h_{\mathrm{HF}}[\gamma]\,g\big)[\vec{x},\tau] \;:= \tag{III.16}$$
$$\big(h\,g\big)[\vec{x},\tau] \;+\; \bigg(\int \frac{\rho_\gamma(\vec{y})\,d^3y}{|\vec{x}-\vec{y}|}\bigg) g(\vec{x},\tau) \;-\; \int \frac{\gamma(\vec{x}|\vec{y})\,g(\vec{y},\tau)\,d^3y}{|\vec{x}-\vec{y}|}\,,$$

with $\rho_\gamma(\vec{x}) := \gamma(\vec{x},\vec{x})$ being the one-particle density corresponding to γ and a partial trace $\gamma(\vec{x},\vec{y}) = \sum_{\tau=\uparrow,\downarrow} \gamma(\vec{x},\tau\,,\,\vec{y},\tau)$ as well as a sufficiently regular choice for the integral kernel for γ are understood.

Comparing Lieb's variational principle to the original Hartree–Fock approximation, it is interesting to observe that the condition $0 \leq G(\underline{f},\underline{f}) \leq \mathbf{1}_{\mathbb{C}^N}$ considered by Lieb and Simon in [70] is actually equivalent to $0 \leq \gamma_{\underline{f}} \leq \mathbf{1}_{\mathfrak{h}}$, if we set $\gamma_{\underline{f}} := \sum_{n=1}^N |f_n\rangle\langle f_n|$. Note, however, that $\gamma_{\underline{f}}$ is of rank N, at most, and hence that $\mathrm{Tr}_{\mathfrak{h}}[\gamma_{\underline{f}}] < N$ unless $\gamma_{\underline{f}}$ is a rank-N projection. It follows that the relaxation of the condition $G(\underline{f},\underline{f}) = \mathbf{1}_{\mathbb{C}^N}$ on the Gram matrix to the bound $0 \leq G(\underline{f},\underline{f}) \leq \mathbf{1}_{\mathbb{C}^N}$ is different from the relaxation of $\gamma = \gamma^2$ to $0 \leq \gamma \leq \mathbf{1}_{\mathfrak{h}}$.

IV Bogolubov Transformations and Representability

We begin our discussion of the concept of *representability* by comparing the two proofs of Theorem 3 given in the previous section. Lieb's original proof seems to be considerably more involved than the one in [3]. One must not overlook, however, that Lieb proves a stronger statement than Eq. (III.6). Namely, the averaging procedure introduced after (III.10) above yields an N-particle density matrix

$$\begin{aligned}\rho_{\mathrm{av}} \;&:= \mathbb{E}_\theta\big(\,|\Phi(\underline{g}^{(\Theta)})\rangle\langle\Phi(\underline{g}^{(\Theta)})|\,\big) \qquad \text{(IV.1)}\\ &= \int_0^1 \cdots \int_0^1 |\Phi(\underline{g}^{(\theta)})\rangle\langle\Phi(\underline{g}^{(\theta)})|\, d\theta_1 \cdots d\theta_J \;\in\; \mathfrak{DM}^{(N)}\,,\end{aligned}$$

whose reduced one-particle density matrix equals the minimizing one-particle density matrix $\gamma_0 = \gamma^{(1)}_{\rho_{\mathrm{av}}}$. Concerning the energy estimate, the key point in Lieb's construction is that

$$\gamma^{(2)}_{\rho_{\mathrm{av}}} \;\leq\; (\mathbf{1} - \mathrm{Ex})\left(\gamma^{(1)}_{\rho_{\mathrm{av}}} \otimes \gamma^{(1)}_{\rho_{\mathrm{av}}}\right), \qquad \text{(IV.2)}$$

which leads to Estimate (III.12), thanks to the positivity $V \geq 0$ of the pair interaction potential V. To describe the significance of this observation we introduce some more definitions and notation. We follow the paper [12] by Lieb, Solovej, and the author, Solovej's lecture notes [94], and the papers [9, 6, 10, 8] by Breteaux, Hach, Menge, Knörr, and the author.

Generalized Reduced Density Matrices The Hamiltonian $\mathbb{H}$ in (II.8) is a linear operator on $\mathfrak{F} = \bigoplus_{N=0}^\infty \mathfrak{H}^{(N)}$ which leaves the N-particle Hilbert spaces $\mathfrak{H}^{(N)}$ invariant. Thus the variation in the Rayleigh-Ritz principles (II.15) for the total ground state energy E_{gs} and (II.13) for the ground state energy $E_{\mathrm{gs}}(N)$ for N particles may both be restricted to density matrices $\rho = \bigoplus_{N=0}^\infty \rho_N \in \mathfrak{DM}$ that are particle-number conserving and even to N-particle density matrices $\rho \in \mathfrak{DM}^{(N)}$ without changing the infimum.

In general, however, density matrices $\rho \in \mathfrak{DM}$ need not leave the N-particle Hilbert spaces $\mathfrak{H}^{(N)}$ invariant, they are only assumed to be even. This can be conveniently formulated with the aid of the self-dual algebra built from creation and annihilation operators which was introduced by Araki [2]. We choose an antiunitary involution $\mathsf{j} : \mathfrak{h} \to \mathfrak{h}$ and define the **self-dual field operator**

$$\boldsymbol{A^*(f_1 \oplus \mathsf{j} f_2)} \;:=\; c^*(f_1) + c(f_2) \;\in\; \mathcal{B}(\mathfrak{F})\,, \qquad \text{(IV.3)}$$

of a **generalized orbital** $\boldsymbol{F} = f_1 \oplus \mathrm{j} f_2 \in \mathfrak{h} \oplus \mathfrak{h}$. Neither $A^*(F)$ and $A^*(G)$ nor $A(F) := [A^*(F)]^*$ and $A^*(G)$ anticommute, but rather

$$A(F) \;=\; A^*(\mathsf{J}F) \quad \text{and} \quad \big\{A(F), A^*(G)\big\} \;=\; \big\langle F \big| G \big\rangle_{\mathfrak{h}\oplus\mathfrak{h}}\,, \tag{IV.4}$$

where $\mathsf{J} : \mathfrak{h} \oplus \mathfrak{h} \to \mathfrak{h} \oplus \mathfrak{h}$ is the antiunitary involution defined by $\mathsf{J}(f_1 \oplus \mathrm{j} f_2) := f_2 \oplus \mathrm{j} f_1$. All creation and annililation operators can be expessed as self-dual field operators $A^*(F)$ for suitable choices of F. We remark that the antiunitary involution $\mathrm{j} : \mathfrak{h} \to \mathfrak{h}$ ensures the linearity of $\mathfrak{h} \oplus \mathfrak{h} \ni F \mapsto A^*(F)$, even though $\mathfrak{h} \ni f \mapsto c(f)$ is antilinear. Its choice is arbitrary and may be adapted to the model under consideration. The Riesz isomorphism $\mathfrak{h} \to \mathfrak{h}^*$, $|f\rangle \mapsto \langle f|$ yields one possible choice. Identifying $\mathfrak{h}^*$ with $\mathfrak{h}$, it is the only choice up to unitary transformation of the domain $\mathfrak{h}$ of definition and of the range $\mathfrak{h}$ of j. The example $\mathfrak{h} = L^2(\mathbb{R})$ with the maps $(\mathrm{j}_1 f)(x) := \overline{f(x)}$ and $(\mathrm{j}_2 \hat{f})(\xi) := \overline{\hat{f}(\xi)}$ gives a good illustration of the freedom in the choice of j.

Now, suppose that $k \in \mathbb{Z}^+$ is a positive integer and $F_1, \ldots, F_{2k} \in \mathfrak{h} \oplus \mathfrak{h}$ are generalized orbitals. The evenness of $\rho \in \mathfrak{DM}$ is equivalent to the vanishing $\mathrm{Tr}_{\mathfrak{F}}[\rho\, A^*(F_1) \cdots A^*(F_{2k-1})] = 0$ of all expectation values of monomials of *odd* degree in the self-dual field operators. If $\rho \in \mathfrak{DM}$ does not preserve the particle number then expectation values $\mathrm{Tr}_{\mathfrak{F}}[\rho\, A^*(F_1) \cdots A^*(F_{2k})]$ of monomials of *even* degree in the self-dual field operators are, in general, non-vanishing - even if the generalized orbitals are all of the form $F_j = f_j \oplus 0$, for all $j = 1, \ldots, 2k$. While the existence of each of these matrix elements is guaranteed by the boundedness of $A^*(F)$, for any $F \in \mathfrak{h} \oplus \mathfrak{h}$, their summability requires an extra assumption. To formulate this we define the subspace

$$\mathcal{L}^1_{\mathbb{N}^k}(\mathfrak{F}) \;:=\; \big\{\rho \in \mathcal{L}^1(\mathfrak{F}) \;\big|\; (\mathbb{N}^{k/2}\rho\mathbb{N}^{k/2}) \in \mathcal{L}^1(\mathfrak{F})\big\} \;\subseteq\; \mathcal{L}^1(\mathfrak{F})\,, \tag{IV.5}$$

which is a Banach space with respect to the norm $\|\rho\|_{\mathbb{N}^k} := \mathrm{Tr}_{\mathfrak{F}}\big|(\mathbb{N}+\mathbf{1})^{k/2}\rho(\mathbb{N}+\mathbf{1})^{k/2}\big|$. We introduce the subset $\mathfrak{DM}_{\mathbb{N}^k} := \mathfrak{DM} \cap \mathcal{L}^1_{\mathbb{N}^k}(\mathfrak{F})$ of all density matrices $\rho \in \mathfrak{DM}$ for which the expectation $\langle \mathbb{N}^k \rangle_\rho < \infty$ of the k^{th} power of the particle number operator is finite.

Given $k \in \mathbb{Z}^+$ and a density matrix $\rho \in \mathfrak{DM}_{\mathbb{N}^k}$, we define its **reduced generalized k-particle density matrix (k-gRDM) $\boldsymbol{\Gamma^{(k)}_\rho} \in \mathcal{B}\big((\mathfrak{h}\oplus\mathfrak{h})^{\otimes k}\big)$** by

$$\begin{aligned}\Big\langle G_1 \otimes \cdots \otimes G_k \;\Big|\; \Gamma^{(k)}_\rho (F_1 \otimes \cdots \otimes F_k)\Big\rangle \\ := \mathrm{Tr}_{\mathfrak{F}}\big[\rho\, A^*(F_1) \cdots A^*(F_k)\, A(G_k) \cdots A(G_1)\,\big]\,,\end{aligned} \tag{IV.6}$$

where $F_1, \ldots, F_k, G_1, \ldots, G_k \in \mathfrak{h}\oplus\mathfrak{h}$. We obtain the **reduced k-particle density matrix (k-RDM) $\boldsymbol{\gamma^{(k)}_\rho} \in \mathcal{B}\big(\mathfrak{h}^{\otimes k}\big)$** by restricting the matrix elements on vectors of

the form $g_i \oplus 0$ and $f_j \oplus 0$, that is,

$$\begin{aligned} &\Big\langle g_1 \otimes \cdots \otimes g_k \,\Big|\, \gamma_\rho^{(k)}(f_1 \otimes \cdots \otimes f_k) \Big\rangle \\ &\quad := \left\langle \begin{pmatrix} g_1 \\ 0 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} g_k \\ 0 \end{pmatrix} \,\middle|\, \Gamma_\rho^{(k)} \left[\begin{pmatrix} f_1 \\ 0 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} f_k \\ 0 \end{pmatrix} \right] \right\rangle . \end{aligned} \tag{IV.7}$$

We observe that in case ρ preserves particle numbers, i.e., $\rho\mathbb{N} = \mathbb{N}\rho$, then $\Gamma_\rho^{(k)}$ is entirely determined by $\gamma_\rho^{(1)}, \gamma_\rho^{(2)}, \ldots, \gamma_\rho^{(k)}$.

The cases $k = 1$ and $k = 2$ are obviously of special interest. We first discuss $k = 1$ and introduce the **pairing operator $\boldsymbol{\alpha_\rho}$** : $\mathfrak{h} \to \mathfrak{h}$ corresponding to ρ by

$$\big\langle g \,\big|\, \alpha_\rho(\mathsf{j}f) \,\big\rangle \;:=\; \mathrm{Tr}_{\mathfrak{F}}\big[\rho\, c(f)\, c(g)\big] , \tag{IV.8}$$

noting that α_ρ vanishes if ρ preserves particle numbers and further that

$$\alpha_\rho^* \;=\; -\mathsf{j}\,\alpha_\rho\,\mathsf{j} . \tag{IV.9}$$

The pairing operator is convenient for the representation of the reduced generalized 1-pdm $\Gamma_\rho^{(1)} : \mathfrak{h} \oplus \mathfrak{h} \to \mathfrak{h} \oplus \mathfrak{h}$ given by

$$\big\langle G \,\big|\, \Gamma_\rho^{(1)}\, F \big\rangle_{\mathfrak{h}\oplus\mathfrak{h}} \;= \mathrm{Tr}_{\mathfrak{F}}\big[\rho\, A^*(F)\, A(G)\big] , \tag{IV.10}$$

for all $F, G \in \mathfrak{h} \oplus \mathfrak{h}$. Viewed as an operator-valued 2×2-matrix acting on vectors $F = f_1 \oplus \mathsf{j}f_2$, the generalized 1-RDM $\Gamma_\rho^{(1)}$ appears as

$$\Gamma_\rho^{(1)} \;=\; \begin{pmatrix} \gamma_\rho^{(1)} & \alpha_\rho \\ \alpha_\rho^* & \mathbf{1} - \mathsf{j}\,\gamma_\rho^{(1)}\mathsf{j} \end{pmatrix} \;=\; \begin{pmatrix} \gamma_\rho^{(1)} & \alpha_\rho \\ -\mathsf{j}\alpha_\rho\mathsf{j} & \mathbf{1} - \mathsf{j}\,\gamma_\rho^{(1)}\mathsf{j} \end{pmatrix} , \tag{IV.11}$$

where we recall that the 1-pdm $\gamma_\rho^{(1)}$ is given by $\langle g|\, \gamma_\rho^{(1)} f\rangle \;=\; \mathrm{Tr}_{\mathfrak{F}}[\rho c^*(f) c(g)]$. Eq. (IV.11) is equivalent to

$$\mathsf{J}\,\Gamma_\rho^{(1)}\,\mathsf{J} \;= \mathbf{1} \;-\; \Gamma_\rho^{(1)} . \tag{IV.12}$$

Inserting $G = F$ in (IV.10) and using the anticommutation relations, it is easily checked that $0 \le \Gamma_\rho^{(1)} \le \mathbf{1}$ holds true which, in turn, is equivalent to

$$\begin{pmatrix} \gamma_\rho^{(1)} - (\gamma_\rho^{(1)})^2 - \alpha_\rho\alpha_\rho^* & \gamma_\rho^{(1)}\alpha_\rho - \alpha_\rho\mathsf{j}\,\gamma_\rho^{(1)}\mathsf{j} \\ [\gamma_\rho^{(1)}\alpha_\rho - \alpha_\rho\mathsf{j}\,\gamma_\rho^{(1)}\mathsf{j}]^* & \mathsf{j}\,[\gamma_\rho^{(1)} - (\gamma_\rho^{(1)})^2 - \alpha_\rho\alpha_\rho^*]\mathsf{j} \end{pmatrix} \;=\; \Gamma_\rho^{(1)} \;-\; \big(\Gamma_\rho^{(1)}\big)^2 \;\ge\; 0 . \tag{IV.13}$$

Note that this yields $\gamma_\rho^{(1)} - (\gamma_\rho^{(1)})^2 \ge 0$ and hence $0 \le \gamma_\rho^{(1)} \le \mathbf{1}$, as asserted in Lemma 1 (ii).

Furthermore, if $\rho \in \mathfrak{DM}_{\mathbb{N}}$ has finite particle number expectation then (IV.13) implies that the pairing operator $\alpha_\rho \in \mathcal{L}^2(\mathfrak{h})$ is Hilbert-Schmidt, with $\mathrm{Tr}_{\mathfrak{h}}[\alpha_\rho^* \alpha_\rho] \leq \mathrm{Tr}_{\mathfrak{h}}[\gamma_\rho^{(1)} - (\gamma_\rho^{(1)})^2]$, and that $\Gamma_\rho^{(1)} - (\Gamma_\rho^{(1)})^2 \in \mathcal{L}^1(\mathfrak{h} \oplus \mathfrak{h})$ is trace-class. In particular, $\Gamma_\rho^{(1)} - (\Gamma_\rho^{(1)})^2$ and hence also $\Gamma_\rho^{(1)}$ admits an expansion of the form $\Gamma_\rho^{(1)} = \sum_{i=1}^\infty \tilde{\lambda}_i \, |F_i\rangle\langle F_i|$, where $\tilde{\lambda}_i \in [0,1]$ are its eigenvalues and $\{F_i\}_{i=1}^\infty \subseteq \mathfrak{h} \oplus \mathfrak{h}$ is an orthonormal basis of eigenvectors $F_i = f_i' \oplus \mathfrak{j} f_i''$ of $\Gamma_\rho^{(1)}$. The invariance $\mathsf{J}\Gamma_\rho^{(1)}\mathsf{J} = \mathbf{1} - \Gamma_\rho^{(1)}$ implies that the eigenvalues and corresponding eigenvectors come in pairs λ_ℓ, F_ℓ and $1 - \lambda_\ell$, $\mathsf{J}F_\ell$. After changing the order of the eigenvalues, if necessary, we obtain that

$$\Gamma_\rho^{(1)} \;=\; \sum_{\ell=1}^\infty \Big(\lambda_\ell \, |F_\ell\rangle\langle F_\ell| \;+\; (1-\lambda_\ell) \, |\mathsf{J}F_\ell\rangle\langle \mathsf{J}F_\ell| \Big) \,, \tag{IV.14}$$

where $F_\ell = f_\ell' \oplus \mathfrak{j} f_\ell''$ and $\{F_\ell, \mathsf{J}F_\ell\}_{\ell=1}^\infty \subseteq \mathfrak{h} \oplus \mathfrak{h}$ is an orthonormal basis. If, additionally, ρ is particle-number preserving and so $\alpha_\rho \equiv 0$, then

$$\begin{aligned} \Gamma_\rho^{(1)} \;&= \gamma_\rho^{(1)} \oplus \big(\mathbf{1} - \mathfrak{j}\, \gamma_\rho^{(1)} \mathfrak{j} \big) \\ &= \sum_{\ell=1}^\infty \Big\{ \lambda_\ell \, |f_\ell \oplus 0\rangle\langle f_\ell \oplus 0| \;+\; (1-\lambda_\ell) \, |0 \oplus \mathfrak{j} f_\ell\rangle\langle 0 \oplus \mathfrak{j} f_\ell| \Big\} \,, \end{aligned} \tag{IV.15}$$

where λ_ℓ are the eigenvalues of $\gamma_\rho^{(1)}$ and $\{f_\ell\}_{\ell=1}^\infty \subseteq \mathfrak{h}$ is an orthonormal basis of its eigenvectors. Since $\sum_{\ell=1}^\infty \lambda_\ell = \mathrm{Tr}_{\mathfrak{h}}[\gamma_\rho^{(1)}] < \infty$, the sequence of eigenvalues including their multiplicities is summable.

Bogolubov Transformations For a density matrix $\rho \in \mathfrak{DM}_{\mathbb{N}}$ of finite particle number expectation the block-diagonal form (IV.15) of $\Gamma_\rho^{(1)}$ can always be obtained by conjugation $\mathbb{U}_W \rho \mathbb{U}_W^*$ of ρ by a (unitary) Bogolubov transformation $\mathbb{U}_W \in U(\mathfrak{F})$ on Fock space corresponding to a **Bogolubov linear map**, i.e., a unitary $W \in U(\mathfrak{h} \oplus \mathfrak{h})$ on $\mathfrak{h} \oplus \mathfrak{h}$, which additionally obeys $\mathsf{J}W = W\mathsf{J}$. The latter condition and the unitarity precisely ensure that the CAR (IV.4) are preserved under these transformations,

$$A(WF) \;= A^*(\mathsf{J}WF) \;=\; A^*(W\mathsf{J}F) \quad \text{and} \tag{IV.16}$$

$$\big\{ A(WF), A^*(WG) \big\} \;=\; \big\langle WF \big| WG \big\rangle_{\mathfrak{h}\oplus\mathfrak{h}} \;=\; \big\langle F \big| G \big\rangle_{\mathfrak{h}\oplus\mathfrak{h}} \,, \tag{IV.17}$$

The Bogolubov linear maps obviously form a subgroup

$$\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}} \;:= \big\{ W \in U(\mathfrak{h}\oplus\mathfrak{h}) \;\big|\; \mathsf{J}W \;=\; W\,\mathsf{J} \,\big\} \tag{IV.18}$$

of $U(\mathfrak{h}\oplus\mathfrak{h})$. Expressing W as a 2×2-matrix of operators, the Bogolubov linear maps can be alternatively characterized as

$$\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}} \;=\left\{ \begin{pmatrix} u & \mathsf{j}v\mathsf{j} \\ v & \mathsf{j}u\mathsf{j} \end{pmatrix} \;\in\; U(\mathfrak{h}\oplus\mathfrak{h}) \,\middle|\, \mathrm{Tr}_{\mathfrak{h}}\big[v^*v\big] \;<\; \infty \right\}, \tag{IV.19}$$

where the condition that v is of Hilbert-Schmidt class, $\mathrm{Tr}_{\mathfrak{h}}[v^*v]<\infty$, is known as the **Shale-Stinespring condition**. Each Bogolubov linear map $W\in\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ is unitarily implementable on Fock space which means that there exists a unitary $\mathbb{U}_W\in U(\mathfrak{F})$ such that, for all $F\in\mathfrak{h}\oplus\mathfrak{h}$,

$$\mathbb{U}_W\, A^*(F)\, \mathbb{U}_W^* \;= A^*(WF) \tag{IV.20}$$

and, in fact, $W\mapsto\mathbb{U}_W$ is a bijection $\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}\to\mathrm{Bog}_{\mathfrak{F}}$, where

$$\mathrm{Bog}_{\mathfrak{F}} \;:= \tag{IV.21}$$
$$\Big\{ \mathbb{U}\in U(\mathfrak{F}) \;\Big|\; \exists\, V\in\mathcal{B}(\mathfrak{h}\oplus\mathfrak{h})\;\forall\, F\in\mathfrak{h}\oplus\mathfrak{h}:\; \mathbb{U}\,A^*(F)\,\mathbb{U}^* \;=\; A^*(VF) \Big\}$$

is the subgroup $\mathbf{Bog}_{\mathfrak{F}}\subseteq U(\mathfrak{F})$ of **Bogolubov transformations**. The Shale-Stinespring condition ensures, that the vacuum vector remains in $\mathfrak{F}$ under the application of $\mathbb{U}_W$, and the transformed creation and annihilation operators $d^*(f):=\mathbb{U}_W c^*(f)\mathbb{U}_W^*$ and $d(f)$ constitute another Fock representation of the CAR with $\mathbb{U}_W\Omega\in\mathfrak{F}$ as the new vacuum vector.

We return to the 1-gRDM $\Gamma_\rho^{(1)}$ of a density matrix $\rho\in\mathfrak{DM}_{\mathbb{N}}$ of finite particle number expectation. These assume the form (IV.11) with nonvanishing pairing operator α_ρ unless ρ preserves particle numbers. In an orthonormal basis $\{F_\ell,\mathsf{J}F_\ell\}\subseteq\mathfrak{h}\oplus\mathfrak{h}$ of eigenvectors with eigenvalues λ_ℓ and $1-\lambda_\ell$, respectively, $\Gamma_\rho^{(1)}$ can be represented as in (IV.14). Starting from this one can construct a Bogolubov linear map $W\in\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ such that

$$W^*\,\Gamma_\rho^{(1)}\,W \;=\; \sum_{\ell=1}^{\infty}\Big\{\lambda_\ell\,|f_\ell\oplus 0\rangle\langle f_\ell\oplus 0| \;+\; (1-\lambda_\ell)\,|0\oplus\mathsf{j}f_\ell\rangle\langle 0\oplus\mathsf{j}f_\ell|\Big\}. \tag{IV.22}$$

Since, for all $F,G\in\mathfrak{h}\oplus\mathfrak{h}$,

$$\mathrm{Tr}_{\mathfrak{F}}\big[\rho\,A^*(WF)\,A(WG)\big] \;=\; \mathrm{Tr}_{\mathfrak{F}}\big[\mathbb{U}_W^*\,\rho\,\mathbb{U}_W\,A^*(F)\,A(G)\big]\,, \tag{IV.23}$$

we obtain that

$$\begin{aligned}\Gamma^{(1)}_{\mathbb{U}_W^*\rho\mathbb{U}_W} &= W^*\,\Gamma_\rho^{(1)}\,W \qquad\qquad (\text{IV.24})\\ &= \sum_{\ell=1}^{\infty}\Big\{\lambda_\ell\,|f_\ell\oplus 0\rangle\langle f_\ell\oplus 0| \;+\; (1-\lambda_\ell)\,|0\oplus\mathsf{j}f_\ell\rangle\langle 0\oplus\mathsf{j}f_\ell|\Big\}.\end{aligned}$$

In other words, the pairing operator $\alpha_{\mathbb{U}_W^*\rho\mathbb{U}_W} = 0$ of the transformed density matrix $\mathbb{U}_W^*\rho\mathbb{U}_W$ vanishes and $\gamma^{(1)}_{\mathbb{U}_W^*\rho\mathbb{U}_W} = \sum_{\ell=1}^{\infty} \lambda_\ell |f_\ell\rangle\langle f_\ell|$ where $\{f_\ell\} \in \mathfrak{h}$ is an orthonormal basis and $\lambda_\ell \in [0,1]$. Note that the vanishing $\alpha_{\mathbb{U}_W^*\rho\mathbb{U}_W} = 0$ of the pairing operator alone does not imply that $\mathbb{U}_W^*\rho\mathbb{U}_W$ is particle-number preserving. Further note that if $W = \left(\begin{smallmatrix} u & \mathsf{j}v\mathsf{j} \\ v & \mathsf{j}u\mathsf{j} \end{smallmatrix}\right) \in \mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ then

$$\begin{aligned} 0 \;\leq\; \gamma^{(1)}_{\mathbb{U}_W^*\rho\mathbb{U}_W} &= u^*\gamma_\rho^{(1)}u + v^*\big(\mathbf{1} - \mathsf{j}\gamma_\rho^{(1)}\mathsf{j}\big)v + v^*\alpha_\rho u + u^*\alpha_\rho v \\ &\leq u^*\gamma_\rho^{(1)}u + v^*v + v^*\alpha_\rho u + u^*\alpha_\rho v\,, \end{aligned} \tag{IV.25}$$

from which we conclude that the transformed density matrix $\mathbb{U}_W^*\rho\mathbb{U}_W$ has finite particle number expectation, as well, since

$$\|\gamma^{(1)}_{\mathbb{U}_W^*\rho\mathbb{U}_W}\|_{\mathcal{L}^1} \;\leq \|\gamma_\rho^{(1)}\|_{\mathcal{L}^1} + \|v\|_{\mathcal{L}^2}^2 + 2\|v\|_{\mathcal{L}^2}\,\|\alpha_\rho\|_{\mathcal{L}^2} \;<\; \infty\,. \tag{IV.26}$$

Inspired by these properties, we define by

$$\mathfrak{G}^{(1)} := \tag{IV.27}$$
$$\left\{\Gamma^{(1)} = \left(\begin{smallmatrix} \gamma^{(1)} & \alpha \\ \alpha^* & \mathbf{1}-\mathsf{j}\gamma^{(1)}\mathsf{j} \end{smallmatrix}\right) \in \mathcal{B}(\mathfrak{h}\oplus\mathfrak{h}) \;\middle|\; \Gamma^{(1)} = \mathsf{J}(\mathbf{1}-\Gamma^{(1)})\mathsf{J} \geq 0\,,\; \gamma^{(1)} \in \mathcal{L}^1(\mathfrak{h})\right\}$$

the set of **generalized one-particle density matrices 1-gpdm** and by

$$\mathfrak{g}^{(1)} := \left\{\gamma^{(1)} \in \mathcal{L}^1(\mathfrak{h}) \;\middle|\; 0 \leq \gamma^{(1)} \leq \mathbf{1}\right\} \tag{IV.28}$$

the set of **one-particle density matrices (1-pdm)**.

Representability of 1-gpdm We have just seen that any 1-gRDM of finite particle-number expectation value *necessarily* is a 1-gpdm in $\mathfrak{G}^{(1)}$. Representability asks for *sufficient* conditions for this relation. That is, a 1-gpdm $\Gamma^{(1)} \in \mathfrak{G}^{(1)}$ is called **representable**, if there exists a density matrix $\rho \in \mathfrak{DM}$ whose reduced generalized one-particle density matrix $\Gamma^{(1)}_\rho$ coincides with $\Gamma^{(1)}$, i.e., if $\Gamma^{(1)} = \Gamma^{(1)}_\rho$.

The following theorem gives an affirmative answer to question of representability of generalized 1-pdm.

Theorem 4. *Every generalized one-particle density matrix $\Gamma^{(1)} \in \mathfrak{G}^{(1)}$ is representable by a density matrix of finite particle number expectation value.*

Proof. Given $\Gamma^{(1)} \in \mathfrak{G}^{(1)}$ we can find a Bogolubov linear map $W \in \mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ such that

$$W^*\,\Gamma^{(1)}\,W \;=\; \begin{pmatrix} \gamma & 0 \\ 0 & \mathbf{1} - \mathsf{j}\,\gamma\mathsf{j} \end{pmatrix} \quad \text{and} \quad \gamma \;=\; \sum_{\ell=1}^{\infty} \lambda_\ell\, |f_\ell\rangle\langle f_\ell| \tag{IV.29}$$

assumes the form (IV.22). Here, $\{f_\ell\}_{\ell=1}^\infty \subseteq \mathfrak{h}$ is an orthonormal basis of eigenvectors of γ with corresponding eigenvalues $\lambda_\ell \in [0,1]$, which we assume w.l.o.g. to be arranged in descending order, $1 \geq \lambda_1 \geq \lambda_2 \geq \ldots \geq 0$. More specifically, we have that $1 = \lambda_1 = \ldots = \lambda_{K-1} > \lambda_K \geq \ldots \geq \lambda_L > \lambda_{L+1} = \lambda_{L+2} = \ldots = 0$, for unique $K \leq \mathrm{Tr}_\mathfrak{h}[\gamma] < \infty$ and $L \in \mathbb{Z}^+ \cup \{\infty\}$. Note that, for $K \leq \ell \leq L$, the eigenvalues $\lambda_\ell \in [\lambda_L, \lambda_K] \subseteq (0,1)$ are away from 0 and 1, and $\mu_\ell := \ln(1-\lambda_\ell) - \ln(\lambda_\ell) \in \mathbb{R}$ exists. Setting $n_k := c^*(f_k)\, c(f_k)$, for all $k \in \mathbb{Z}^+$, and $\mathbb{P}_1 := n_1\, n_2 \cdots n_{K-1}$, we define

$$\mathbb{h}_0 \;:=\; \sum_{\ell=K}^{L} \mu_\ell\, n_\ell\,, \;\; Z_0 \;:=\; \mathrm{Tr}_\mathfrak{F}[e^{-\mathbb{h}_0}]\,, \;\text{ and }\; \rho_0 \;:=\; \mathbb{P}_1\, Z_0^{-1}\, \exp[-\mathbb{h}_0]\,. \tag{IV.30}$$

Note that an orthonormal basis of $\mathfrak{F}$ of eigenvectors of n_k with eigenvalues $\nu_k \in \{0,1\}$ is given by $\Psi_{\underline{\nu}} := \prod_{\ell=1}^\infty [c^*(f_\ell)]^{\nu_\ell}\Omega$, where $\underline{\nu} = (\nu_\ell)_{\ell=1}^\infty \in \{0,1\}^{\mathbb{Z}^+}$ runs through all sequences of occupation numbers $\nu_\ell \in \{0,1\}$ of finite sum $|A(\underline{\nu})| < \infty$, with $A(\underline{\nu}) := \{\ell \in \mathbb{Z}^+ | \nu_\ell = 1\} \subseteq \mathbb{Z}^+$. That is, $n_k \Psi_{\underline{\nu}} = \nu_k \Psi_{\underline{\nu}}$, for any $k \in \mathbb{Z}^+$. Hence

$$Z_0 \;=\; \sum_{\underline{\nu}:|A(\underline{\nu})|<\infty} \big\langle \Psi_{\underline{\nu}} \big|\, e^{-\mathbb{h}_0} \Psi_{\underline{\nu}} \big\rangle \;=\; \prod_{\ell=K}^{L} \big(1 + e^{-\mu_\ell}\big) \;<\; \infty\,, \tag{IV.31}$$

since $\sum_{\ell=1}^\infty e^{-\mu_\ell} = \sum_{\ell=K}^L (1-\lambda_\ell)^{-1}\lambda_\ell \leq (1-\lambda_K)^{-1} \sum_{\ell=1}^\infty \lambda_\ell < \infty$. It follows that $\rho_0 \in \mathfrak{DM}$ is a density matrix, which is obviously particle-number preserving and, therefore, has vanishing pairing operator $\alpha_{\rho_0} = 0$. Moreover, if $\max\{k,\ell\} \geq K$ then

$$\langle f_\ell \,|\, \gamma^{(1)}_{\rho_0} f_k \rangle \;=\; \frac{\mathrm{Tr}_\mathfrak{F}[e^{-\mathbb{h}_0}\, c^*(f_k)\, c(f_\ell)]}{Z_0} \;=\; \frac{\delta_{k,\ell}\; e^{-\mu_k}}{1+e^{-\mu_k}} \;=\; \delta_{k,\ell}\; \lambda_k\,, \tag{IV.32}$$

while, for $\min\{k,\ell\} \leq K$, we observe that $\langle f_\ell | \gamma^{(1)}_{\rho_0} f_k\rangle = \delta_{k,\ell} = \delta_{k,\ell}\lambda_k$, as well. This implies that $W^*\Gamma^{(1)}W = \Gamma^{(1)}_{\rho_0}$ and thus

$$\Gamma^{(1)} \;= W\, \Gamma^{(1)}_{\rho_0}\, W^* \;=\; \Gamma^{(1)}_{\mathbb{U}_W\, \rho_0\, \mathbb{U}_W^*}\,. \tag{IV.33}$$

Since $\rho_0 \in \mathfrak{DM}$ is a density matrix, so is $\mathbb{U}_W\, \rho_0\, \mathbb{U}_W^* \in \mathfrak{DM}$. □

***N*-Representability of 1-pdm** Similar to the notion of representability of a generalized 1-pdm, we call a 1-pdm $\gamma^{(1)} \in \mathfrak{g}^{(1)}$ with $\mathrm{Tr}[\gamma^{(1)}] = N \in \mathbb{Z}^+$ ***N*-representable**, if there exists an N-particle density matrix $\rho \in \mathfrak{DM}^{(N)}$ such that $\gamma^{(1)} = \gamma^{(1)}_\rho$.

The N-representability of any 1-pdm has actually been proved by Lieb in [69], although this had not been its main purpose.

Theorem 5. *Let $N \in \mathbb{Z}^+$ with $N \geq 2$ and $\gamma^{(1)} \in \mathfrak{g}^{(1)}$ a one-particle density matrix of particle number expectation $\mathrm{Tr}[\gamma^{(1)}] = N$. Then $\gamma^{(1)}$ is N-representable.*

Proof. Given $\gamma^{(1)}$, the N-particle density matrix $\rho_{\mathrm{av}} \in \mathfrak{DM}^{(N)}$ in (IV.1) fulfills $\gamma^{(1)} = \gamma^{(1)}_{\rho_{\mathrm{av}}}$. □

Representability of generalized 2-pdm Let $N \in \mathbb{Z}^+$ with $N \geq 2$. As proven in Theorems 4 and 5 above, the maps $\mathfrak{DM}_{\mathbb{N}} \to \mathfrak{G}^{(1)}$, $\rho \mapsto \Gamma^{(1)}_\rho$ and $\mathfrak{DM}^{(N)} \to \{\gamma \in \mathfrak{g}^{(1)} | \mathrm{Tr}[\gamma] = N\}$, $\rho \mapsto \gamma^{(1)}_\rho$ are bijections. The simple characterizations of the sets $\mathfrak{G}^{(1)}$ and $\mathfrak{g}^{(1)}$ is an encouraging sign that the extension of the notion of representability to reduced generalized k-pdm for $k \geq 2$ leads to similarly simple characterizations.

Following this sign, we call a pair $(\Gamma^{(1)}, \Gamma^{(2)}) \in \mathcal{B}(\mathfrak{h}^2) \times \mathcal{B}(\mathfrak{h}^2 \otimes \mathfrak{h}^2)$ of bounded positive operators **representable**, if $\Gamma^{(1)} = \Gamma^{(1)}_\rho$ and $\Gamma^{(2)} = \Gamma^{(2)}_\rho$, for some density matrix $\rho \in \mathfrak{DM}_{\langle \mathbb{N}^2 \rangle < \infty}$ of finite particle number variance, where $\mathfrak{h}^2 := \mathfrak{h} \oplus \mathfrak{h}$.

Somewhat more restrictive, we call a pair $(\gamma^{(1)}, \gamma^{(2)}) \in \mathcal{B}(\mathfrak{h}) \times \mathcal{B}(\mathfrak{h} \otimes \mathfrak{h})$ of bounded positive operators **representable**, if $\gamma^{(1)} = \gamma^{(1)}_\rho$ and $\gamma^{(2)} = \gamma^{(2)}_\rho$, for some particle-number preserving density matrix $\rho \in \mathfrak{DM}_{\langle \mathbb{N}^2 \rangle < \infty}$ of finite particle number variance. If ρ can additionally be chosen to be an N-particle density matrix then $(\gamma^{(1)}, \gamma^{(2)})$, respectively, is called ***N*-representable**. Note that necessarily $\gamma^{(1)}$ results from $\gamma^{(2)}$ by taking a partial trace [see (II.28)] and $N = \mathrm{Tr}[\gamma^{(1)}]$ in this case.

With these definitions we obtain new characterizations of the total and the N-particle ground state energies as

$$E_{\mathrm{gs}} = \inf \Big\{ \mathcal{E}_Q\big(\gamma^{(1)}, \gamma^{(2)}\big) \;\Big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h}),\; (\gamma^{(1)}, \gamma^{(2)}) \in \mathcal{B}(\mathfrak{h}) \times \mathcal{B}(\mathfrak{h} \otimes \mathfrak{h}) \text{ is representable} \Big\}, \tag{IV.34}$$

$$E_{\mathrm{gs}}(N) = \inf \Big\{ \mathcal{E}_Q\big(\gamma^{(1)}, \gamma^{(2)}\big) \;\Big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h}),\; (\gamma^{(1)}, \gamma^{(2)}) \in \mathcal{B}(\mathfrak{h}) \times \mathcal{B}(\mathfrak{h} \otimes \mathfrak{h}) \text{ is } N\text{-representable} \Big\}. \tag{IV.35}$$

This characterization of the ground state energy was first given by Coleman [28], following a remark of Coulson [30]. It seems to yield a drastic simplification of the task of determining ground state energies and ground states of many-fermion systems, as the number of variables of the problem is reduced from N to 4. This is, however, too optimistic because the problem of restricting the variation in (IV.34) and (IV.35) to representable respectively N-representable pairs $(\gamma^{(1)}, \gamma^{(2)})$ is, perhaps, as difficult as solving the corresponding Schrödinger equation on Fock space altogether.

The requirement that the density matrix from which $(\gamma^{(1)}, \gamma^{(2)})$ derives is particle number preserving or even an N-particle density matrix adds considerably to the degree of difficulty of the problem, as is seen when comparing the proofs of Theorems 4 and 5 in case that $k = 1$. A characterization of the representability of $(\Gamma^{(1)}, \Gamma^{(2)}) \in \mathcal{B}(\mathfrak{h}^2) \times \mathcal{B}(\mathfrak{h}^2 \otimes \mathfrak{h}^2)$ would already be great progress.

Nevertheless, we now focus on particle number preserving density matrices ρ for which the reduced generalized 1-pdm $(\Gamma_\rho^{(1)}, \Gamma_\rho^{(2)})$ are completely determined by the 1-RDM $(\gamma_\rho^{(1)}, \gamma_\rho^{(2)})$. The difficulty described above has lead to what is known as the **representability problem** of quantum chemistry: *Specify a condition* $A : \mathcal{L}^1(\mathfrak{h}) \times \mathcal{L}^1(\mathfrak{h} \otimes \mathfrak{h}) \to \{\textit{true}, \textit{false}\}$ *such that* $(\gamma^{(1)}, \gamma^{(2)})$ *is representable if* $A(\gamma^{(1)}, \gamma^{(2)}) = \textit{true}$. The representability problem is considered open until today (at least by those who do not accept tautologies as its solution). It is known to be a hard problem in the sense of QMA complexity in computer science, as demonstrated by Liu, Christandl, and Verstraete in [75]. An overview on questions of reduced density matrices and their representability is given by Coleman and Yukalov in [29].

GPQ Condition and $T_{1;2}$ Condition While the representability problem is about the specification of a *sufficient* condition for the representability of a pair $(\gamma^{(1)}, \gamma^{(2)})$, research on conditions reduced one- and two-particle density matrices *necessarily* fulfill has been more successful in the past. Namely, if a condition $B : \mathcal{L}^1(\mathfrak{h}) \times \mathcal{L}^1(\mathfrak{h} \otimes \mathfrak{h}) \to \{\text{true}, \text{false}\}$ is such that $B(\gamma_\rho^{(1)}, \gamma_\rho^{(2)}) = \text{true}$, for any density matrix $\rho \in \mathfrak{DM}$ then it is immediate that

$$E_{\text{gs}} \geq \inf \Big\{ \mathcal{E}_Q\big(\gamma^{(1)}, \gamma^{(2)}\big) \;\Big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h})\,,\; B(\gamma^{(1)}, \gamma^{(2)}) = \text{true} \Big\}\,, \tag{IV.36}$$

$$E_{\text{gs}}(N) \geq \inf \Big\{ \mathcal{E}_Q\big(\gamma^{(1)}, \gamma^{(2)}\big) \;\Big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h})\,,\; \mathrm{Tr}[\gamma^{(1)}] = N\,,\; B(\gamma^{(1)}, \gamma^{(2)}) = \text{true} \Big\}\,. \tag{IV.37}$$

In practise, B is not a single condition but a list of conditions that $\gamma^{(1)}$ and $\gamma^{(2)}$ ought to fulfill, and (i) and (ii) in Lemma 1 are always part of this list. That is, it is understood that $\gamma^{(1)} \in \mathfrak{g}^{(1)}$ is a 1-pdm and obeys $0 \leq \gamma^{(1)} \leq \mathbf{1}_\mathfrak{h}$ and $\mathrm{Tr}[\gamma^{(1)}] < \infty$. Theorems 4 and 5 ensure that there are not more conditions on $\gamma^{(1)}$ alone, that do not involve $\gamma^{(2)}$.

Almost sixty years ago Coleman [28] and Garrod and Percus [45] specified three conditions, which a representable pair $(\gamma^{(1)}, \gamma^{(2)})$ of a one- and two-particle density matrix necessarily fulfill. These three conditions were orginally called "G", "P", and "Q", respectively, but we refer to them as a single condition which

we call the **GPQ condition**. We apply the scheme described in (IV.36) and (IV.37) above and introduce

$$E_{\mathrm{GPQ}}(N) \;:=\; \inf\big\{\mathcal{E}_Q(\gamma^{(1)},\gamma^{(2)}) \;\big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h})\,, \quad \mathrm{Tr}[\gamma^{(1)}] = N\,,\ (\gamma^{(1)},\gamma^{(2)}) \text{ fulfills GPQ}\big\}\,, \tag{IV.38}$$

observing that $E_{\mathrm{gs}}(N) \geq E_{\mathrm{GPQ}}(N)$. In [9], Knörr, Menge, and the author considered self-adjoint, but not necessarily positive, trace-class operators $\rho = \rho^* \in \mathcal{L}^1(\mathfrak{F})$ obeying $\mathrm{Tr}_{\mathfrak{F}}\big(|\rho|^{1/2}\mathbb{N}^2|\rho|^{1/2}\big) < \infty$. It is easy to see that, for these ρ, the operators $\Gamma^{(2)}_\rho$, given by (IV.6), define trace-class operators on $\mathfrak{h}^2 \otimes \mathfrak{h}^2$. In [9], the GPQ condition was proven to be is equivalent to the positivity of $\Gamma^{(2)}_\rho \geq 0$ on $\mathfrak{h}^2 \otimes \mathfrak{h}^2$. Furthermore, it was shown in [9] that the GPQ condition implies the fermion correlation inequality

$$\begin{aligned}\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[(P\otimes P)\gamma^{(2)}\big] \;&\geq \mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[(P\otimes P)(\mathbf{1}-\mathrm{Ex})(\gamma^{(1)}\otimes\gamma^{(1)})\big] \\ &\quad - \mathrm{Tr}_{\mathfrak{h}}[P\gamma^{(1)}]\;\min\Big\{1\,,\;9\,\mathrm{Tr}_{\mathfrak{h}}\big[P\big(\gamma^{(1)}-(\gamma^{(1)})^2\big)^{1/2}\big]\Big\}\,,\end{aligned} \tag{IV.39}$$

where $P = P^* = P^2 \in \mathcal{B}(\mathfrak{h})$ is an arbitrary orthogonal projection. This inequality is the key input for the proof in [3, 9] that, for large Coulomb systems, the difference of the Hartree–Fock energy and $E_{\mathrm{GPQ}}(N)$ is bounded by $o(Z^{5/3})$, which implies that the accuracy of the Hartree–Fock approximation is at least as good,

$$0 \;\leq\; E_{\mathrm{HF}}(Z) - E_{\mathrm{gs}}(N) \;\leq\; E_{\mathrm{HF}}(Z) - E_{\mathrm{GPQ}}(N) \;\leq\; o(Z^{(5/3)})\,. \tag{IV.40}$$

Since the exchange term is in magnitude greater than a universal multiple of $Z^{5/3}$, see (I.12)-(I.13), Eq. (IV.40) proves that the accuracy of the Hartree–Fock approximation is better than the smallest contribution to the Hartree–Fock energy.

In [39, 38], Erdahl found additional representability conditions he called T_1 and T_2. We refer to these as a single condition, the $\boldsymbol{T_{1;2}}$ **condition**. It arises from observables of the form $Q_4 := P_3^*P_3 + P_3P_3^*$, where P_3 is any polynomial in the self-dual field operators of degree three. Obviously, Q_4 is a nonnegative operator. Moreover, while both $P_3^*P_3$ and $P_3P_3^*$ are polynomials of degree six, their sum Q_4 is an anticommutator and hence a polynomial of degree four or less. Thus, $\mathrm{Tr}_{\mathfrak{F}}[\rho\,Q_4] \geq 0$ yields a condition the pair of reduced generalized 1-pdm and 2-pdm $(\Gamma^{(1)}_\rho, \Gamma^{(2)}_\rho)$ corresponding to ρ necessarily fulfills. We introduce

$$E_{\mathrm{GPQ:T}}(N) \;:=\; \inf\big\{\mathcal{E}_Q(\gamma^{(1)},\gamma^{(2)}) \;\big|\; (h\,\gamma^{(1)}) \in \mathcal{L}^1(\mathfrak{h})\,, \quad \mathrm{Tr}[\gamma^{(1)}] = N\,,\ (\gamma^{(1)},\gamma^{(2)}) \text{ fulfills GPQ and } T_{1;2}\big\}\,. \tag{IV.41}$$

Erdahl's theoretical discovery came into focus of quantum chemists some two decades later, when numerical simulations demonstrated, that, in test cases with small N, the accuracy of $E_{\mathrm{GPQ:T}}(N)$ is comparable to the accuray of *full CI (configuration interaction)* computations, i.e., the full solution of the N electron Schrödinger equation (projected onto a finite dimensional subspace, as part of the Galerkin approximation). These were carried out, e.g., by Erdahl and Mazziotti in [77], Zhao, Braams, Fukuda, Overton, and Percus in [99], Cances, Lewin, and Stoltz in [24], Braams, Percus, and Zhao in [20], and Naftchi-Ardebili, Hau, and Mazziotti in [78].

V Quadratic Hamiltonians and Quasifree Density Matrices

Quadratic Hamiltonians: We return to the definition of Bogolubov linear maps $\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ and Bogolubov transformations $\mathrm{Bog}_{\mathfrak{F}}$. The former consists of unitary linear maps $W \in U(\mathfrak{h}\oplus\mathfrak{h})$ on $\mathfrak{h}\oplus\mathfrak{h}$, which additionally obeys $\mathsf{J}W = W\mathsf{J}$, the latter are unitary operators $\mathbb{U}_W \in U(\mathfrak{F})$ obeying

$$\forall\, F \in \mathfrak{h}\oplus\mathfrak{h}: \qquad \mathbb{U}_W\, A^*(F)\, \mathbb{U}_W^* \;=\; A^*(WF)\,, \tag{V.1}$$

and the map $\mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}} \ni W \mapsto \mathbb{U}_W \in \mathrm{Bog}_{\mathfrak{F}}$ is a group isomorphism.

Next, we define the second quantization $\mathbb{Q}(T) \in \mathcal{B}(\mathcal{D}(N); \mathfrak{F}_f)$ of a bounded operator $T = T^* = \left(\begin{smallmatrix} a & b \\ b^* & 0 \end{smallmatrix}\right) \in \mathcal{B}[\mathfrak{h}\oplus\mathfrak{h}]$, with $a = a^*$ and $b = -\mathsf{j}b^*\mathsf{j}$ by

$$\mathbb{Q}(T) \;:=\; \sum_{i,j=1}^{\infty} \langle F_i|\, TF_j\rangle\; A^*(F_i)\, A(F_j)\,, \tag{V.2}$$

where $\{F_i\}_{i=1}^{\infty} \subseteq \mathfrak{h}\oplus\mathfrak{h}$ is an orthonormal basis. The definition of $\mathbb{Q}(T)$ is independent of the choice of this orthonormal basis. Under the assumption that $b \in \mathcal{L}^2(\mathfrak{h})$ is a Hilbert-Schmidt operator and $a \geq 0$ is nonnegative, $\mathbb{Q}(T)$ is self-adjoint and semibounded on the domain of the particle number operator. (Generally, a relative bound in form of the Hilbert-Schmidt property of $a^{-1/2}ba^{-1/2}$ should be sufficient, as this was shown to hold true for boson systems by Nam, Napiorkowski, and Solovej in [79].) We refer to $\mathbb{Q}(\boldsymbol{T})$ as the **quadratic Hamiltonian corresponding to $\boldsymbol{T}$** because it is of degree two in the self-dual field operators. An explicit computation (on finite vectors and then extension by continuity) yields

$$[\mathbb{Q}(T)\,,\, A^*(F)] \;=\; A^*\big(\widehat{T}F\big)\,, \tag{V.3}$$

which implies that

$$e^{i\mathbb{Q}(T)}\, A^*(F)\, e^{-i\mathbb{Q}(T)} \;=\; A^*\big(e^{i\widehat{T}}F\big)\,, \tag{V.4}$$

for any $F \in \mathfrak{h}\oplus\mathfrak{h}$, where $\widehat{T} = -\mathsf{J}\widehat{T}\mathsf{J} := \left(\begin{smallmatrix} a & 2b \\ 2b^* & -a \end{smallmatrix}\right) \in \mathcal{B}[\mathfrak{h}\oplus\mathfrak{h}]$.

Indeed, if we set $A_t^*(\widetilde{F}) := e^{it\mathbb{Q}(T)}A^*\big(e^{-it\widehat{T}}\widetilde{F}\big)e^{-it\mathbb{Q}(T)}$ for $t \in [0,1]$ then $\dot{A}_t^*(\widetilde{F}) = 0$, by (V.3), and hence $e^{i\mathbb{Q}(T)}A^*\big(e^{-i\widehat{T}}\widetilde{F}\big)e^{-i\mathbb{Q}(T)} = A_1^*(\widetilde{F}) = A_0^*(\widetilde{F}) = A^*(\widetilde{F})$ which directly yields (V.4) with $\widetilde{F} := e^{i\widehat{T}}F$.

Note that we cannot directly quantize $\frac{1}{2}\widehat{T}$ in the sense of (V.2), for if we replace T by $\frac{1}{2}\widehat{T}$ in (V.2), we obtain an expression $\frac{1}{2}\mathbb{Q}(\widehat{T})$, say, which fulfilled $\frac{1}{2}\mathbb{Q}(\widehat{T}) = \mathbb{Q}(T) + \frac{1}{2}\mathrm{Tr}[a]$ and would, hence, not exist in case that a is not trace-class. Further note that by the antilinearity of J, we have that $[i\widehat{T}]\mathsf{J} = -i\mathsf{J}T = \mathsf{J}[i\widehat{T}]$ and

hence $e^{-i\widehat{T}}\mathsf{J} = \mathsf{J}e^{-i\widehat{T}}$. Since $\mathbb{Q}(T)$ is self-adjoint, $e^{i\mathbb{Q}(T)} \in \mathrm{Bog}_{\mathfrak{F}}$ is a Bogolubov transformation with

$$\exp[i\mathbb{Q}(T)] \;=\; \mathbb{U}_{\exp[-i\widehat{T}]}\,. \tag{V.5}$$

In fact, all Bogolubov transformations can be written in this form or, at least, approximated in the strong topology. That is, we may identify the Bogolubov transformations with the family of unitary operators generated by i times self-adjoint quadratic Hamiltonians,

$$\mathrm{Bog}_{\mathfrak{F}} \;=\; \overline{\left\{ \exp[i\mathbb{Q}(T)] \;\middle|\; T = \left(\begin{smallmatrix} a & b \\ b^* & 0 \end{smallmatrix}\right),\; a \in \mathcal{B}(\mathfrak{h})\,,\; a \geq 0\,,\; b \in \mathcal{L}^2(\mathfrak{h}) \right\}}\,, \tag{V.6}$$

where the bar denotes closure in the strong operator topology.

Quasifree Density Matrices: It turns out that quadratic Hamiltonians play an important role not only for Bogolubov transformations, but also for density matrices. Recall from (IV.30) the definition of the density matrix $\rho_0 = \mathbb{P}_1 Z_0^{-1} \exp[-\mathbb{h}_0] \in \mathfrak{DM}$, where $\mathbb{P}_1 = n_1\, n_2 \cdots n_{K-1}$ and

$$\mathbb{h}_0 \;=\; \sum_{\ell=K}^{L} \mu_\ell\, c^*(f_\ell)\, c(f_\ell) \;=\; \mathbb{Q}(H_0) \tag{V.7}$$

is the quadratic Hamiltonian corresponding to

$$H_0 \;:=\; \begin{pmatrix} h_0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad h_0 \;:=\; \sum_{\ell=K}^{L} \mu_\ell\, |f_\ell\rangle\langle f_\ell|\,. \tag{V.8}$$

We now construct an approximation $\rho_\varepsilon \in \mathfrak{QDM}$ for ρ_0, such that $\rho_\varepsilon \to \rho_0$ in $\mathfrak{DM}_{\mathbb{N}}$, as $\varepsilon \to 0$. For $\varepsilon > 0$, we define

$$\tau_\ell(\varepsilon) \;:=\; \begin{cases} 1 - \varepsilon/K\,, & \ell < K\,, \\ \mu_\ell\,, & K \leq \ell \leq L\,, \\ \varepsilon\, e^{-\ell^2}\,, & \ell > L\,, \end{cases} \tag{V.9}$$

noting that $\tau_\ell(\varepsilon) \in (0,1)$, for all $\ell \in \mathbb{Z}^+$. Next, we set $H_\varepsilon := h_\varepsilon \oplus 0$, where $h_\varepsilon := \sum_{\ell=1}^{\infty} \tau_\ell(\varepsilon) |f_\ell\rangle\langle f_\ell|$. Finally, $Z_\varepsilon := \mathrm{Tr}_{\mathfrak{F}}\big(\exp[-\mathbb{Q}(H_\varepsilon)]\big)$ and $\rho_\varepsilon := Z_\varepsilon^{-1}\, \exp[-\mathbb{Q}(H_\varepsilon)]$.

We further recall from (IV.33) that ρ_0, after conjugation with the Bogolubov transformation $\mathbb{U}_W \in \mathrm{Bog}_{\mathfrak{F}}$, yields the density matrix $\widetilde{\rho}_0 := \mathbb{U}_W \rho_0 \mathbb{U}_W^*$ whose reduced generalized 1-pdm $\Gamma^{(1)}_{\widetilde{\rho}_0}$ equals the prescribed generalized 1-pdm $\Gamma^{(1)}$ from Theorem 4. If $T_\varepsilon = \left(\begin{smallmatrix} a_\varepsilon & b_\varepsilon \\ b_\varepsilon^* & 0 \end{smallmatrix}\right)$, $a_\varepsilon \in \mathcal{B}(\mathfrak{h})$, $a_\varepsilon \geq 0$, and $b_\varepsilon \in \mathcal{L}^2(\mathfrak{h})$, is such that $\mathbb{U}_W \Psi = \lim_{\varepsilon\to 0} \exp[i\mathbb{Q}(T_\varepsilon)]\Psi$, for all $\Psi \in \mathfrak{F}$, then in $\mathfrak{DM}_{\mathbb{N}}$

$$\widetilde{\rho}_\varepsilon \;:=\; \exp[i\mathbb{Q}(T_\varepsilon)]\, \rho_\varepsilon\, \exp[-i\mathbb{Q}(T_\varepsilon)] \;\to\; \widetilde{\rho}_0 \qquad \varepsilon \to 0\,, \tag{V.10}$$

due to the unitarity of $\mathbb{U}_W$ and the fact that $\rho_0 \in \mathfrak{DM}_{\mathbb{N}}$. On the other hand,

$$\widetilde{\rho}_\varepsilon \;=\; Z^{-1}\exp[-\mathbb{Q}(\widetilde{H}_\varepsilon)]\,, \quad \text{where} \quad \widetilde{H}_\varepsilon \;=\; e^{-i\widehat{T}}\,H_\varepsilon\,e^{i\widehat{T}}\,. \tag{V.11}$$

So, defining the set of **quasifree density matrices**

$$\mathfrak{QDM} \;:=\; \overline{\left\{ Z^{-1}\,\exp[-\mathbb{Q}(H_0)] \;\middle|\; H_0 \in \mathrm{qh}\,,\; Z := \exp[-\mathbb{Q}(H_0)] < \infty \right\}}\,, \tag{V.12}$$

where the bar indicates closure in $\mathfrak{DM}_{\mathbb{N}}$ and

$$\mathrm{qh} \;:=\; \left\{ \begin{pmatrix} a & b \\ b^* & 0 \end{pmatrix} \;\middle|\; a = a^* \geq 0\,,\; e^{-a} \in \mathcal{L}^1(\mathfrak{h})\,,\; b \in \mathcal{L}^2(\mathfrak{h}) \right\}\,, \tag{V.13}$$

we conclude from (V.10)-(V.11) that every 1-gpdm is the 1-gRDM of a semigroup generated by a quadratic Hamiltonian or a limit in $\mathfrak{DM}_{\mathbb{N}}$ thereof.

Note that $\mathfrak{QDM} \subseteq \mathfrak{DM}$. Further note that the closure in the definition (V.12) of quasifree density matrices is important because otherwise the orthogonal projection $\rho = |\Phi_{\underline{f}}\rangle\langle\Phi_{\underline{f}}|$ onto the Slater determinant $\Phi_{\underline{f}} = f_1 \wedge \cdots \wedge f_N$ of orthonormal orbitals $f_1, \ldots, f_N \in \mathfrak{h}$ would be excluded. The Slater determinant $\Phi_{\underline{f}}$, however, is the Bogolubov transform $\mathbb{U}_W\Omega$ of the vacuum vector, with $W = \left(\begin{smallmatrix} P^\perp & \mathrm{j}P\mathrm{j} \\ P & \mathrm{j}P^\perp\mathrm{j} \end{smallmatrix}\right)$, where $P = \sum_{n=1}^N |f_n\rangle\langle f_n|$ is the orthogonal projection onto the subspace spanned by $f_1, \ldots, f_n \in \mathfrak{h}$. Hence, the corresponding rank-one projection $\rho = |\mathbb{U}_W\Omega\rangle\langle\mathbb{U}_W\Omega|$ is a (pure) quasifree density matrix. The Bogolubov linear map $W \in \mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ is not of the form $\exp[-i\mathbb{Q}(T)]$, for any $T \in \mathrm{qh}$, but can be obtained as a strong limit of these.

With the definitions in Eqs. (V.12)-(V.13), we observe that the proof of Theorem 4 actually yields the following stronger statement.

Corollary 6. *Let $\Gamma^{(1)} \in \mathfrak{G}^{(1)}$ be a generalized one-particle density matrix. Then there exists a unique quasifree density matrix $\eta \in \mathfrak{QDM}$ such that $\Gamma^{(1)} = \Gamma^{(1)}_\eta$.*

We omit to comment on the uniqueness part of Corollary 6 but point out the following important consequence.

Corollary 7. *The requirement $\Gamma^{(1)}_\rho = \Gamma^{(1)}_\eta$ defines a map*

$$q : \mathfrak{DM} \;\to\; \mathfrak{QDM}\,,\;\; \rho \mapsto q(\rho) = \eta\,. \tag{V.14}$$

*For a density matrix $\rho \in \mathfrak{DM}$, its image $\boldsymbol{q(\rho)} = \eta \in \mathfrak{QDM}$ is called its **quasifree reduction**.*

For a quasifree state $\rho \in \mathfrak{QDM}$ the reduced generalized k-pdm can be explicitly computed in terms of its reduced generalized 1-pdm, as the following theorem asserts.

Theorem 8. *Let $\rho \in \mathfrak{DM}$ be a density matrix and denote $\langle M \rangle := \mathrm{Tr}_{\mathfrak{F}}[\rho M]$, for any $M \in \mathcal{B}(\mathfrak{F})$. Then the following statements are equivalent.*

(i) The density matrix $\rho \in \mathfrak{QDM}$ is quasifree.

(ii) For all $k \geq 2$, all truncated $2k$-point functions vanish, i.e., for all $F_1, F_2, \ldots, F_{2k} \in \mathfrak{h} \oplus \mathfrak{h}$,

$$\langle A_1 \, A_2 \cdots A_{2k} \rangle \;=\; \sum_{\pi \in \mathcal{P}_{2k}} (-1)^{\pi} \, \langle A_{\pi(1)} A_{\pi(2)} \rangle \, \langle A_{\pi(3)} A_{\pi(4)} \rangle \cdots \langle A_{\pi(2k-1)} A_{\pi(2k)} \rangle \,, \tag{V.15}$$

where $A_i := A^*(F_i)$, $\mathcal{P}_{2k}$ is the set of permutations $\pi : \{1, 2, \ldots, 2k\} \to \{1, 2, \ldots, 2k\}$ that obey $\pi(1) < \pi(3) < \ldots < \pi(2k-1)$ and $\pi(2j-1) < \pi(2j)$, for all $1 \leq j \leq k$, and $(-1)^{\pi}$ denotes its sign.

(iii) All truncated four-point functions vanish, i.e., for all $F_1, F_2, F_3, F_4 \in \mathfrak{h} \oplus \mathfrak{h}$,

$$\langle A_1 \, A_2 \, A_3 \, A_4 \rangle \;=\; \langle A_1 A_2 \rangle \langle A_3 A_4 \rangle - \langle A_1 A_3 \rangle \langle A_2 A_4 \rangle + \langle A_1 A_4 \rangle \langle A_2 A_3 \rangle \,, \tag{V.16}$$

where $A_i := A^*(F_i)$.

Characterization (ii) of quasifree density matrices in the above theorem is often taken as their definition. The somewhat surprising statement for a given density matrix, that already the vanishing of its truncated four-point functions implies its quasifreeness originates in a theorem of Marcinkiewicz [76] in (classical) probability theory. In the context of quantum physics, it was first proved by Robinson [87] and later generalized in [16, 17] for boson systems. The generalization to fermions can be traced back to work of Rajagopal and Sudarshan [85], see also the comment by Titulaer [98]. We refer to Salmhofer [88] for a modern presentation of truncated fermion correlation functions.

The quasifree reduction $q : \mathfrak{DM} \to \mathfrak{QDM}$ defined in Corollary 7 is a projection, i.e., an idempotent map $q^2 = q$ from the density matrices onto quasifree density matrices. Mauser and Gottlieb [49] observed that the image $q(\rho) \in \mathfrak{QDM}$ of $\rho \in \mathfrak{DM}$ under this projection is the closest element to ρ in $\mathfrak{QDM}$ in the sense that it minimizes the relative entropy among all quasifree density matrices, as the following Theorem asserts.

Theorem 9. *Let $\rho \in \mathfrak{DM}$ be a density matrix and $q(\rho) \in \mathfrak{QDM}$ its quasifree reduction. If the relative entropy*

$$S[\rho, q(\rho)] \;:=\; \mathrm{Tr}_{\mathfrak{F}}\big\{\rho\big(\log[\rho] - \log[q(\rho)]\big)\big\} \;<\; \infty \tag{V.17}$$

exists, then

$$S[\rho, q(\rho)] \;=\; \inf_{\eta \in \mathfrak{QDM}} \big\{S[\rho, \eta]\big\}\,. \tag{V.18}$$

Proof. Let $\eta \in \mathfrak{QDM}$ be a quasifree density matrix which, for simplicity, is assumed to be given as the exponential $\eta = Z^{-1}\,\exp[-\mathbb{Q}]$ of a quadratic Hamiltonian $\mathbb{Q} \equiv \mathbb{Q}(H_0)$, for some $H_0 \in \mathrm{qh}$, and that the von Neumann entropies $S[\rho] := -\mathrm{Tr}_{\mathfrak{F}}\{\rho\,\log[\rho]\}$, $S[\eta]$, and $S[q(\rho)]$ of ρ, its quasifree reduction, and η exist. Then $-\log[q(\rho)] = \mathbb{Q} + \log(Z)$ and hence

$$\begin{aligned} S[\rho, \eta] \;&= S[\rho] \,+\, \mathrm{Tr}_{\mathfrak{F}}\big\{\rho\big(-\log[\eta]\big)\big\} \;=\; \log(Z) \,+\, S[\rho] \,+\, \mathrm{Tr}_{\mathfrak{F}}\{\rho\,\mathbb{Q}\} \\ &= \log(Z) \,+\, S[\rho] \,+\, \mathrm{Tr}_{\mathfrak{F}}\{q[\rho]\,\mathbb{Q}\} \;=\; S[\rho] \,-\, \mathrm{Tr}_{\mathfrak{F}}\{q[\rho]\,\log[\eta]\}\,, \end{aligned} \tag{V.19}$$

since quadratic observables have the same expectation value w.r.t. a density matrix and its quasifree reduction. The same identity holds true, if we replace η by $q[\rho]$, and we obtain

$$S[\rho, \eta] - S[\rho, q(\rho)] \;= \mathrm{Tr}_{\mathfrak{F}}\big\{q(\rho)\big(\log[q(\rho)] - \log[\eta]\big)\big\} \;=\; S[q(\rho), \eta] \;\geq\; 0\,, \tag{V.20}$$

since relative entropy is nonnegative. □

We remark that the existence of the von Neumann entropies $S[\rho] := -\mathrm{Tr}_{\mathfrak{F}}\{\rho\log[\rho]\}$ of ρ and $q[\rho]$ is assumed in the proof of Theorem 9 only for convenience and is not implied by the finiteness of their relative entropy $S[\rho, q(\rho)]$. Note, however, that if $\rho \in \mathfrak{QDM}$ is quasifree and $S[\rho] < \infty$ then it assumes the simple form

$$S[\rho] \;=\; -\,\mathrm{Tr}_{\mathfrak{F}}\{\rho\,\log[\rho]\} \;=\; S^{(1)}[\Gamma^{(1)}_{\rho}] \;:=\; -\mathrm{Tr}_{\mathfrak{h}\oplus\mathfrak{h}}\big\{\Gamma^{(1)}_{\rho}\,\log[\Gamma^{(1)}_{\rho}]\big\}\,. \tag{V.21}$$

This is not hard to check for a quasifree density matrix $\rho_0 \in \mathfrak{QDM}$ of the form $\rho_0 = \mathbb{P}_1 Z_0^{-1}\exp[-\mathbb{h}_0]$, as in (IV.30) by explicit computation. The general identity (V.21) then follows from the invariance of $S[\rho]$ and $S^{(1)}[\Gamma^{(1)}]$ under unitary transformations and the application of a suitable Bogolubov linear map $W \in \mathrm{Bog}_{\mathfrak{h}\oplus\mathfrak{h}}$ to $\Gamma^{(1)}_{\rho}$ to transform it to $\Gamma^{(1)}_{\rho_0}$ and the corresponding Bogolubov transformation $\mathbb{U}_W \in \mathrm{Bog}_{\mathfrak{F}}$ to ρ to transform it to ρ_0. See also [12].

VI Generalized Hartree-Fock Approximation and Generalizations of Lieb's Variational Principle

Generalized Hartree–Fock Approximation: Since quasifree density matrices are, in particular, density matrices, we immediately observe that the **generalized Hartree–Fock energy**

$$\boldsymbol{E_{\mathbf{gHF}}} \;:=\; \inf\left\{\mathrm{Tr}_{\mathfrak{F}}(\rho\,\mathbb{H}_\mu)\;\middle|\;\rho\;\in\;\mathfrak{QDM}\,,\;\langle\mathbb{H}\rangle_\rho<\infty\right\} \tag{VI.1}$$

defines an upper bound $E_{\mathrm{gHF}} \geq E_{\mathrm{gs}}$ on the total ground state energy E_{gs} defined in (II.15). For a quasifree density matrix $\rho \in \mathfrak{QDM}$, the energy expectation value $\mathrm{Tr}_{\mathfrak{F}}(\rho\,\mathbb{H}_\mu) = \mathcal{E}_{\mathrm{gHF}}(\Gamma_\rho^{(1)})$ depends only on its reduced generalized 1-pdm $\Gamma_\rho^{(1)} = \left(\begin{smallmatrix}\gamma_\rho^{(1)} & \alpha_\rho \\ \alpha_\rho^* & \mathbf{1}-\mathsf{J}\gamma_\rho^{(1)}\mathsf{J}\end{smallmatrix}\right) \in \mathfrak{G}^{(1)}$, where

$$\begin{aligned}\mathcal{E}_{\mathrm{gHF}}(\Gamma_\rho^{(1)}) \;&:=\mathrm{Tr}_{\mathfrak{h}}\big[h_\mu\,\gamma_\rho^{(1)}\big] \;+\; \frac{1}{2}\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\,(\mathbf{1}-\mathrm{Ex})(\gamma_\rho^{(1)}\otimes\gamma_\rho^{(1)})\big]\\ &+\;\frac{1}{2}\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\,\mathrm{Ex}\,(\alpha_\rho^*\otimes\alpha_\rho)\big]\,,\end{aligned} \tag{VI.2}$$

where $h_\mu := h - \mu\mathbf{1}$. Moreover, since $\mathfrak{QDM} \ni \rho \mapsto \Gamma_\rho^{(1)} \in \mathfrak{G}^{(1)}$ is a bijection, we obtain

$$E_{\mathrm{gHF}} \;=\; \inf\left\{\mathcal{E}_{\mathrm{gHF}}(\Gamma^{(1)})\;\middle|\;\Gamma^{(1)}\;\in\;\mathfrak{G}^{(1)}\right\}. \tag{VI.3}$$

Note that if minimizers $\Gamma_{\mathrm{gHF}}^{(1)}$ exist then they necessarily fulfill a stationarity condition, which in generalized Hartree–Fock theory also takes the form of a self-consistent equation

$$\Gamma_{\mathrm{gHF}}^{(1)} \;=\; \mathbf{1}_\neg\big(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}]\big)\,, \tag{VI.4}$$

where $h_{\mathrm{gHF}}[\Gamma^{(1)}]$ is again an effective Hamiltonian on $\mathfrak{h}\oplus\mathfrak{h}$ and $\mathbf{1}_\neg(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}])$ is a certain projection onto the eigenspaces of $h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}]$ of negative and zero eigenvalues such that $\mathbf{1}(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}]<0) \leq \mathbf{1}_\neg(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}]) \leq \mathbf{1}(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}] \leq 0)$. The precise form of $\mathbf{1}_\neg$ is difficult to determine because of the requirement $\mathbf{1}-\Gamma_{\mathrm{gHF}}^{(1)} = \mathsf{J}\Gamma_{\mathrm{gHF}}^{(1)}\mathsf{J}$ which $\Gamma_{\mathrm{gHF}}^{(1)}$ and, therefore, also $\mathbf{1}_\neg(h_{\mathrm{gHF}}[\Gamma_{\mathrm{gHF}}^{(1)}])$ necessarily fulfills.

It is possible, however, to use the (yet another) generalization of generalized Hartree–Fock theory to positive temperatures $1/\beta > 0$, which is not reviewed here, and obtain a minimizer $\Gamma_{\mathrm{gHF}}^{(1)}$ by the zero temperature limit $\beta\to\infty$ of a

family of minimizers $\big(\Gamma_\beta^{(1)}\big)_{\beta\in\mathbb{R}^+}$ for inverse temperature $1/\beta$. Fox fixed β, the minimizer $\Gamma_\beta^{(1)}$ necessarily fulfills the self-consistent equation

$$\Gamma_\beta^{(1)} \;=\; F_\beta\big(h_\beta[\Gamma_\beta^{(1)}]\big)\,, \tag{VI.5}$$

where $h_\beta[\Gamma]$ is a suitable effective Hamiltonian, itself depending on β, and $F_\beta(x) = (1+e^{\beta x})^{-1}$ is the Fermi function.

We remark that the generalized Hartree–Fock theory for positive temperature derives from a variational principle, namely the minimization of the **Hartree–Fock pressure functional $-\mathcal{P}_\beta$** by

$$-\mathcal{P}_\beta(\Gamma^{(1)}) \;:=\; \mathcal{E}_{\mathrm{gHF}}(\Gamma^{(1)}) \;-\; \beta^{-1}\,S^{(1)}(\Gamma^{(1)})\,. \tag{VI.6}$$

Lieb, Solovej, and the author have demonstrated in [12] that it fulfills

$$\mathcal{P}_\beta(\Gamma^{(1)}) \;\le\; \beta^{-1}\,\log\big[\mathrm{Tr}_{\mathfrak{F}}\big\{\,\exp[-\beta\mathbb{H}_\mu]\big\}\big]\,, \tag{VI.7}$$

for any generalized 1-pdm $\Gamma^{(1)} \in \mathfrak{G}^{(1)}$ and, hence, yields a lower bound to the pressure (in the sense of statistical mechanics), in analogy to $\mathcal{E}_{\mathrm{gHF}}(\Gamma^{(1)})$ being an upper bound to the total ground state energy E_{gs}.

Repulsive Potentials: If $\mathfrak{h} = L^2(M,d\nu)$ for a measure space $(M,d\nu)$, and $V(x,y) \ge 0$ is a repulsive potential, i.e., a nonnegative multiplication operator on $\mathfrak{h}\otimes\mathfrak{h}$, then

$$\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\,\mathrm{Ex}\,(\alpha_\rho^*\otimes\alpha_\rho)\big] \;=\; \int V(x,y)\,|\alpha_\rho(x,y)|^2\,d\nu(x)\,d\nu(y) \;\ge\; 0\,. \tag{VI.8}$$

In other words: For repulsive pair potentials the pairing operator yields a nonnegative contribution to the energy. Now, if $\Gamma^{(1)} = \left(\begin{smallmatrix}\gamma^{(1)} & \alpha\\ \alpha^* & \mathbf{1}-\mathrm{j}\gamma^{(1)}\mathrm{j}\end{smallmatrix}\right) \in \mathfrak{G}^{(1)}$ is a 1-gpdm, so is $\widetilde{\Gamma}^{(1)} := \left(\begin{smallmatrix}\mathbf{1} & 0\\ 0 & -\mathbf{1}\end{smallmatrix}\right)\Gamma^{(1)}\left(\begin{smallmatrix}\mathbf{1} & 0\\ 0 & -\mathbf{1}\end{smallmatrix}\right) = \left(\begin{smallmatrix}\gamma^{(1)} & -\alpha\\ -\alpha^* & \mathbf{1}-\mathrm{j}\gamma^{(1)}\mathrm{j}\end{smallmatrix}\right) \in \mathfrak{G}^{(1)}$, and by the convexity of $\mathfrak{G}^{(1)}$ we conclude that $\widehat{\Gamma}^{(1)} = \frac{1}{2}\big(\Gamma^{(1)}+\widetilde{\Gamma}^{(1)}\big) = \left(\begin{smallmatrix}\gamma^{(1)} & 0\\ 0 & \mathbf{1}-\mathrm{j}\gamma^{(1)}\mathrm{j}\end{smallmatrix}\right) \in \mathfrak{G}^{(1)}$ is a 1-gpdm, too. Its energy expectation value, however, is

$$\begin{aligned}\mathcal{E}_{\mathrm{gHF}}(\widehat{\Gamma}^{(1)}) \;&= \mathrm{Tr}_{\mathfrak{h}}\big[h_\mu\,\gamma^{(1)}\big] \;+\; \frac{1}{2}\mathrm{Tr}_{\mathfrak{h}\otimes\mathfrak{h}}\big[V\,(\mathbf{1}-\mathrm{Ex})(\gamma^{(1)}\otimes\gamma^{(1)})\big]\\ &\le \mathcal{E}_{\mathrm{gHF}}(\Gamma^{(1)})\,.\end{aligned} \tag{VI.9}$$

It follows that for repulsive pair potentials, the generalized Hartree-Fock energy agrees with the (total) Hartree-Fock energy and does not improve the approximation,

$$E_{\mathrm{gHF}} \;=\; E_{\mathrm{HF}} \;=\; \inf\big\{\mathcal{E}_{\mathrm{HF}}(\gamma)\;\big|\;\gamma\;\in\mathcal{L}^1(\mathfrak{h})\,,\;0\le\gamma\le\mathbf{1}\big\}\,, \tag{VI.10}$$

where the **total Hartree–Fock energy** is defined as $\boldsymbol{E_{\mathbf{HF}}} := \inf_{N\in\mathbb{Z}^+} E_{\mathrm{HF}}(N)$. We stress that the total Hartree–Fock energy, as a function of the chemical potential μ, is the Legrende transform of the Hartree-Fock energy $E_{\mathrm{HF}}(N)$ for N particles. This does not have a direct consequence for $E_{\mathrm{HF}}(N)$.

Attractive Potentials: If $\mathfrak{h} = L^2(M, d\nu)$ and the pair potential $V : M \times M \to \mathbb{R}$ is strictly negative in some subset of $M \times M$ then it may happen that $E_{\mathrm{gHF}} < E_{\mathrm{HF}}$ and the generalized Hartree–Fock approximation is, indeed, better than the original Hartree–Fock approximation and **pairing** occurs, i.e., all minimizers $\Gamma^{(1)} = \left(\begin{smallmatrix} \gamma^{(1)} & \alpha \\ \alpha^* & \mathbf{1}-\mathsf{j}\gamma^{(1)}\mathsf{j} \end{smallmatrix}\right) \in \mathfrak{G}^{(1)}$ have a nonvanishing pairing operator $\alpha \neq 0$. In this case the Hartree–Fock equations indicating the stationarity of the energy functional at the minimum turn into BCS-type equations for which those found by Bardeen, Cooper, and Schrieffer in [15] for the description of superconductivity are a special case. Because of similarity, the stationarity condition is usually called *BCS equation*, and the latter have been systematically analyzed for translational invariant systems under the additional assumption, or constraint, that only translation invariant states enter the energy functional [56], see Section VII.

There is no general criterion for the occurence of *pairing*, but in case the fermions in the model are electrons or other spin-$\frac{1}{2}$ particles, the one-particle Hilbert space is of the form $\mathfrak{h} = \hat{\mathfrak{h}} \otimes \mathbb{C}^2$ with $\hat{\mathfrak{h}} = L^2(M, d\nu)$, the interaction potential V is spin-independent and purely attractive, $V \leq 0$, and the operators $h = \hat{h} \otimes \mathbf{1}$ and $V = (-\widehat{V}) \otimes (\mathbf{1} \otimes \mathbf{1})$ are *real*, i.e., $\mathsf{j} = \hat{\mathsf{j}} \otimes \mathbf{1}$, $\hat{\mathsf{j}}\hat{h} = \hat{h}\hat{\mathsf{j}}$, and $(\hat{\mathsf{j}} \otimes \hat{\mathsf{j}})\widehat{V} = \widehat{V}(\hat{\mathsf{j}} \otimes \hat{\mathsf{j}})$, an explicit characterization of pairing was given by Fröhlich, Jonsson, and the author in [7]: Under these assumptions, the energy minimizing 1-gpdm always takes the form

$$\Gamma^{(1)} \equiv \Gamma^{(1)}[\hat{\gamma}] := \begin{pmatrix} \hat{\gamma} & 0 & 0 & \sqrt{\hat{\gamma}-\hat{\gamma}^2} \\ 0 & \hat{\gamma} & -\sqrt{\hat{\gamma}-\hat{\gamma}^2} & 0 \\ 0 & -\sqrt{\hat{\gamma}-\hat{\gamma}^2} & \mathbf{1}-\hat{\gamma} & 0 \\ \sqrt{\hat{\gamma}-\hat{\gamma}^2} & 0 & 0 & \mathbf{1}-\hat{\gamma} \end{pmatrix}, \tag{VI.1}$$

where the auxiliary 1-pdm $\hat{\gamma} \in \mathcal{L}^1(\hat{\mathfrak{h}})$, $0 \leq \hat{\gamma} \leq \mathbf{1}_{\hat{\mathfrak{h}}}$, on $\hat{\mathfrak{h}}$ minimizes the resulting auxiliary functional

$$\widehat{\mathcal{E}}_{\mathrm{aux}}(\hat{\gamma}) := \frac{1}{2}\mathcal{E}_{\mathrm{gHF}}(\Gamma^{(1)}[\hat{\gamma}]) = \mathrm{Tr}_{\hat{\mathfrak{h}}}[\hat{h}\,\hat{\gamma}] \tag{VI.2}$$

$$- \frac{1}{2}\iint \widehat{V}(x,y)\Big\{\rho_{\hat{\gamma}}(x)\,\rho_{\hat{\gamma}}(y) - |\hat{\gamma}(x,y)|^2 + \big|\sqrt{\hat{\gamma}-\hat{\gamma}^2}(x,y)\big|^2\Big\}\, d\nu(x)\, d\nu(y)\,.$$

Note that the minimizer is real in the sense that $\hat{\mathrm{j}}\hat{\gamma} = \hat{\gamma}\hat{\mathrm{j}}$ and $\mathrm{j}\gamma = \gamma\mathrm{j}$. Further note that the pairing operator entering $\Gamma^{(1)}[\hat{\gamma}]$ assumes the form

$$\alpha \;=\; \sqrt{\hat{\gamma} - \hat{\gamma}^2} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \tag{VI.3}$$

where the second 2×2-matrix factor $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ ensures the antisymmetry condition $\alpha^* = -\alpha = -\mathrm{j}\alpha\mathrm{j}$, which an admissible pairing operator necessarily fulfills according to (IV.9).

The physical system to which [7] was applied is a star consisting of neutrons, which are spin-$\frac{1}{2}$ fermions that attract each other by gravity. While it is generally important to prove statements about minimization problems without requiring the actual existence of a minimizer, [7] left this existence question unresolved. Lenzmann and Lewin, however, proved in [64] the existence of a minimizer for these neutron stars under natural conditions.

Dirac–Fock Equations: Shortly after the discovery of *nonrelativistic* quantum mechanics and the formulation of the Hartree–Fock approximation, a *relativistic* analogue, the **Dirac–Fock (DF) equations**, was formulated by Swirles [97]. The proof of existence of solutions the Dirac–Fock equations pose a considerably more difficult problem as compared to proving this for the Hartree–Fock equations, due to the unboundeness of the energy functional from below, direct methods from the calculus of variations do not really apply, as was pointed out by Chaix, Iracane, and Lions in [26, 27] who introduced and studied the **Bogolubov–Dirac–Fock (BDF) model**. An effective renormalization and then control on the unboundedness below of the BDF model, i.e., of the Hartree–Fock energy functional for electrons and positrons for small particle number and coupling constant was first obtained by Barbaroux, Helffer, Siedentop, and the author in [5]. The existence of solutions for Coulomb systems was shown by Esteban and Séré [40] and by Paturel [83]. In the case of atoms, Barbaroux, Farkas, Helffer, and Siedentop and Barbaroux, Esteban, and Séré related the Dirac–Fock equations to the Hartree-Fock equations of the electron-positron field in [14, 13]. For the same model, Hainzl, Lewin, and Séré proved in [52, 53] the existence of a minimizer and its uniqueness for the BDF model, but in contrast to [5] Hainzl et al. chose the projection on the free Dirac sea as a reference and extended their results later to atoms and molecules with small particle number and small coupling constants. Huber and Siedentop proved [60], in turn, that the Dirac–Fock equations for atoms possess solutions if, among other smallness conditions on coupling constants, the particle number N is such that the shells of the corresponding hydrogen-like Dirac operator are exactly filled.

Generalization of Lieb's Variational Principle: We come back to Lieb's argument which establishes his variational principle. The N-particle density matrix $\rho_{\rm av}$ in (IV.1) results from averaging the pure quasifree density matrices $|\Phi(\underline{g}^{(\theta)})\rangle\langle\Phi(\underline{g}^{(\theta)})|$ over all possible values of θ. Hence, there exists at least one choice of θ such that the energy expectation value $\langle\Phi(\underline{g}^{(\theta)})|\,\mathbb{H}\Phi(\underline{g}^{(\Theta)})\rangle$ of this pure quasifree density matrix is smaller or equal to the energy expectation $\mathrm{Tr}_{\mathfrak{F}}[\rho_{\rm av}\,\mathbb{H}]$ of $\rho_{\rm av}$.

Derezinski, Napiorkowski, and Solovej [31] and, simultaneously, Breteaux, Knoerr, Menge, and the author generalized this statement in [6] and demonstrated that the generalized Hartree–Fock energy $E_{\rm gHF}$ can be approximated by energy expectation values of *pure* quasifree density matrices to arbitrary accuracy. Moreover, as we additionally point out here, if $E_{\rm gHF}$ is a minimum then there is also a pure quasifree density matrix among the minimizers.

The generalization does not only extend the set of density matrices, over which the Hartree–Fock energy functional is being varied, but also allows for any semibounded self-adjoint Hamiltonian with no additional repulsiveness assumption on the pair potential and not even on the form of the Hamiltonian, that it be a sum of a one-body term and a pair interaction.

Theorem 10. *Suppose that $\mathbb{H}_\mu = \mathbb{H}_\mu^*$ is semibounded. Then $E_{\rm gHF} = \widehat{E}_{\rm gHF}$, where*

$$\widehat{E}_{\rm gHF} \;:=\; \inf\left\{\langle\Omega\,|\,\mathbb{U}^*\,\mathbb{H}_\mu\,\mathbb{U}\Omega\rangle \;\middle|\; \mathbb{U}\in\mathrm{Bog}_{\mathfrak{F}}\right\}. \tag{VI.4}$$

Moreover, if there is a quasifree density matrix $\rho_{\rm gHF}\in\mathfrak{QDM}\cap\mathfrak{DM}_{\mathbb{N}}$ of finite particle number such that $\mathrm{Tr}_{\mathfrak{F}}[\rho_{\rm gHF}\mathbb{H}_\mu] = E_{\rm gHF}$, then there exists a Bogolubov transformation $\mathbb{U}_{\rm gHF}\in\mathrm{Bog}_{\mathfrak{F}}$ such that $E_{\rm gHF} = \langle\Omega|\,\mathbb{U}^*\mathbb{H}_\mu\mathbb{U}\Omega\rangle$.

Proof. We only prove the second part of the theorem and assume that $\rho_{\rm gHF}\in\mathfrak{QDM}\cap\mathfrak{DM}_{\mathbb{N}}$ is a quasifree density matrix of finite particle number expectation value and with $\mathrm{Tr}_{\mathfrak{F}}[\rho_{\rm gHF}\mathbb{H}_\mu] = E_{\rm gHF}$. The requirement of finiteness of $\langle\mathbb{N}\rangle_{\rho_{\rm gHF}}$ can be relaxed, but we do not carry this out here. We can find a Bogolubov transformation $\widetilde{\mathbb{U}}\in\mathrm{Bog}_{\mathfrak{F}}$ such that $\rho_{\rm gHF}$ takes the form $\rho_{\rm gHF} = \widetilde{\mathbb{U}}^*\rho_0\widetilde{\mathbb{U}}$, where $\rho_0 = \mathbb{P}_1 Z^{-1}\exp[-\mathbb{Q}(H_0)]$, with $\mathbb{P}_1 = n_1\,n_2\cdots n_{K-1}$ and $\mathbb{Q}(H_0) = \sum_{\ell=K}^{L}\mu_\ell\,c^*(f_\ell)\,c(f_\ell)$, as in (IV.30) and in (V.7), respectively.

Since $\mathbb{H}_\mu$ is semibounded, $\widetilde{\mathbb{H}} := \mathbb{H}_\mu - E_{\rm gs}\geq 0$, as a quadratic form, $\widetilde{E}_{\rm gHF} := E_{\rm gHF} - E_{\rm gs}\geq 0$, and $\widetilde{E}_{\rm gHF} = \mathrm{Tr}_{\mathfrak{F}}[\rho_{\rm gHF}\widetilde{\mathbb{H}}] = \mathrm{Tr}_{\mathfrak{F}}[\rho_{\rm gHF}^{1/2}\widetilde{\mathbb{H}}\rho_{\rm gHF}^{1/2}]$. It follows that

$$\widetilde{E}_{\rm gHF} \;=\mathrm{Tr}_{\mathfrak{F}}\big[\rho_{\rm gHF}^{1/2}\,\widetilde{\mathbb{H}}\,\rho_{\rm gHF}^{1/2}\big] \;=\sum_{\underline{\nu}:|A(\underline{\nu})|<\infty}\big\langle\rho_{\rm gHF}^{1/2}\widetilde{\mathbb{U}}^*\Psi_{\underline{\nu}}\,\big|\,\widetilde{\mathbb{H}}\,\rho_{\rm gHF}^{1/2}\widetilde{\mathbb{U}}^*\Psi_{\underline{\nu}}\big\rangle \tag{VI.5}$$

$$=\sum_{\underline{\nu}:|A(\underline{\nu})|<\infty}\big\langle\rho_0^{1/2}\Psi_{\underline{\nu}}\,\big|\,\widetilde{\mathbb{U}}^*\widetilde{\mathbb{H}}\widetilde{\mathbb{U}}\,\rho_0^{1/2}\Psi_{\underline{\nu}}\big\rangle \;=\sum_{\underline{\nu}\in\mathcal{A}}\big\langle\rho_0^{1/2}\Psi_{\underline{\nu}}\,\big|\,\widetilde{\mathbb{U}}^*\widetilde{\mathbb{H}}\widetilde{\mathbb{U}}\,\rho_0^{1/2}\Psi_{\underline{\nu}}\big\rangle\,,$$

where the orthonormal basis $\big\{\Psi_{\underline{\nu}}\,\big|\underline{\nu}\in\{0,1\}^{\mathbb{Z}^+},\,|A(\underline{\nu})|<\infty\big\}\subseteq\mathfrak{F}$ is introduced in (IV.31), and the convergence of the series is guaranteed by the positivity of each term. Moreover, the summation can be restricted to the subset $\mathcal{A}:=\{\underline{\nu}\in\{0,1\}^{\mathbb{Z}^+}|\,|A(\underline{\nu})|<\infty,\rho_0^{1/2}\Psi_{\underline{\nu}}\neq 0\}$ of indices $\underline{\nu}$, for which $\rho_0^{1/2}\Psi_{\underline{\nu}}$ is nonvanishing. The latter vectors and the set $\mathcal{A}$ can, however, be determined explicitly. Indeed, $\rho_0^{1/2}\Psi_{\underline{\nu}}\neq 0$ only if $A(\underline{\nu})\supseteq\{1,2,\ldots,K-1\}$, and in this case, up to a sign, we have that

$$\begin{aligned}\rho_0^{1/2}\Psi_{\underline{\nu}} &= c_1^*\cdots c_{K-1}^*\prod_{\ell\in A(\underline{\nu})\cap\{K,\ldots,L\}}\Big(\sqrt{\lambda_\ell\,(1-\lambda_\ell)^{-1}}\;c_\ell^*\Big)\Omega\\ &= \big\|\rho_0^{1/2}\Psi_{\underline{\nu}}\big\|\prod_{\ell\in A(\underline{\nu})\cap\{1,\ldots,L\}}c_\ell^*\Omega\,.\end{aligned}\tag{VI.6}$$

It follows that $\mathcal{A}:=\big\{\underline{\nu}\in\{0,1\}^{\mathbb{Z}^+}\big|\,|A(\underline{\nu})|<\infty,A(\underline{\nu})\subseteq\{1,\ldots,L\}\big\}$ and that

$$\widetilde{E}_{\mathrm{gHF}}\;=\sum_{\underline{\nu}\in\mathcal{A}}\|\rho_0^{1/2}\Psi_{\underline{\nu}}\|^2\,\langle\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\,|\,\widetilde{\mathbb{H}}\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\rangle\,.\tag{VI.7}$$

Since $\Psi_{\underline{\nu}}$ is a Slater determinant, for each $\underline{\nu}\in\mathcal{A}$, the pure density matrix $|\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\rangle\langle\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}|\in\mathfrak{QDM}$ is quasifree and, hence, $\langle\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}|\widetilde{\mathbb{H}}\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\rangle\geq\widetilde{E}_{\mathrm{gHF}}$. Moreover, since $\|\rho_0^{1/2}\Psi_{\underline{\nu}}\|^2>0$, for all $\underline{\nu}\in\mathcal{A}$ and $\sum_{\underline{\nu}\in\mathcal{A}}\|\rho_0^{1/2}\Psi_{\underline{\nu}}\|^2=\mathrm{Tr}_{\mathfrak{F}}[\rho_0]=1$, Eq. (VI.7) implies that

$$\forall\,\underline{\nu}\in\mathcal{A}:\qquad\langle\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\,|\,\widetilde{\mathbb{H}}\widetilde{\mathbb{U}}\Psi_{\underline{\nu}}\rangle\;=\;\widetilde{E}_{\mathrm{gHF}}\tag{VI.8}$$

and thus the assertion. □

VII Symmetries and Restricted Hartree–Fock Approximation

In this final section we discuss symmetries of the quantum system under consideration. We assume that the Hamiltonian $\mathbb{H} = \mathbb{h} + \frac{1}{2}\mathbb{V}$ is given in second quantized form as in (II.8) with $\mathbb{h}$ and $\mathbb{V}$ as in (II.21)-(II.22) and to obey stability of matter, i.e., that $\mathbb{H}_\mu = \mathbb{H} - \mu\mathbb{N}$ is semibounded for sufficiently small $\mu < 0$ and hence $\mathbb{H}_\mu + E_0 \geq 1$, for sufficiently large $E_0 > 0$.

A family $\mathcal{S}$ of unitary operators $U \in \mathcal{S}$ is called a **symmetry** of $\mathbb{H}$ if $U(\mathbb{H} + E_0)^{-1} = (\mathbb{H} + E_0)^{-1}U$, for all $U \in \mathcal{S}$. Given a symmetry $\mathcal{S}$, we define the **restricted generalized Hartree–Fock (gHF) energy** to be

$$\boldsymbol{E_{\mathbf{gHF}}(\mathcal{S})} := \tag{VII.1}$$
$$\inf\left\{ \mathrm{Tr}_{\mathfrak{F}}(\rho\, \mathbb{H}_\mu) \;\middle|\; \rho \in \mathfrak{QDM},\ \langle \mathbb{H} \rangle_\rho < \infty,\ \forall U \in \mathcal{S}:\ U\rho = \rho U \right\}.$$

Obviously, $E_{\mathrm{gHF}}(\mathcal{S}) \geq E_{\mathrm{gHF}}$, and the approximation made by the restricted gHF energy is not better, and potentially worse, than the one without restriction. The importance of the restricted gHF approxmiation, however, lies in its improved accessability to explicit computation. Translation invariant generalized 1-pdm, for example, can be diagonalized by Fourier transform, or rotationally invariant generalized 1-pdm have a natural decomposition in terms of spherical harmonics.

- If $E_{\mathrm{gHF}}(\mathcal{S}) = E_{\mathrm{gHF}}$ then the symmetry $\mathcal{S}$ is called **preserved**.
- If $E_{\mathrm{gHF}}(\mathcal{S}) > E_{\mathrm{gHF}}$ then the symmetry $\mathcal{S}$ is called **broken**.

It turns out that both cases of preserved symmetry and broken symmetry occur in different models. The reason is the hidden concavity of Hartree–Fock functionals, which is used in the proof of Theorem 3 and which leads to instabilities at the minimum of the restricted functional. We discuss symmetries on various examples of physical interest.

Closed Shell Theorem in Unrestricted Hartree–Fock Theory and Rotation of Atoms: We first discuss rotation symmetry and come back to the Hartree–Fock approximation as originally introduced for atoms. The periodic table of the elements is usually described in terms of angular momentum shells, which contain the electron states. This picture implicitly assumes that the electron orbitals are eigenfunctions of the angular momentum operators L^2 and L_z. Indeed, the Hamiltonian and the Hartree–Fock functional of an atom is invariant under rotations about the origin, where the atomic nucleus is located. Its minimizers, however, do generally not possess this rotational symmetry unless we study the restricted theory. Indeed, Griesemer and Hantsch show in [50] that, without restriction by symmetries, the HF minimizer of an atom with two electrons and a small nuclear

charge Z breaks rotational symmetry, while the HF minimizer becomes rotationally invariant for N electrons that fill up the lowest angular momentum shells (e.g., $N = q$, where $q = 2$ is the number of spin states), as $Z \gg N$ becomes sufficiently large.

This phenomenon is also reflected by the *closed shell theorem* in [11]: If a HF minimizer $\gamma_{\mathrm{HF}} \in \mathfrak{g}^{(1)}$ for a Coulomb system of N electrons exists, then it is the rank-N orthogonal projection onto the smallest N eigenvalues $e_1 \leq e_2 \leq \ldots \leq e_N$ of the corresponding effective Hamiltonian $h_{\mathrm{HF}}[\gamma_{\mathrm{HF}}]$, as in (III.16), and the lowest spectral point of $h_{\mathrm{HF}}[\gamma_{\mathrm{HF}}]$ greater or equal than e_N is *strictly bigger* than e_N,

$$e_{N+1} \;:=\; \inf\Big\{\sigma\big(h_{\mathrm{HF}}[\gamma_{\mathrm{HF}}]\big) \setminus \{e_1, e_2, \ldots, e_N\}\Big\} \;>\; e_N\,. \tag{VII.2}$$

The interpretation of this statement is that, in Hartree–Fock approximation, atoms and molecules never possess an open shell because the highest energy level is always fully occupied. In particular, rare earth elements with one loosely bound valence electron in a degenerate high momentum shell do no occur in Hartree–Fock theory. Therefore, the Hartree–Fock approximation for a single Lithium atom, say, does not yield orbitals which are products of a radial function and a spherical harmonic.

Particle Number Conservation: The strongly continuous one-parameter group $\mathcal{N} = \big(\exp[-it\mathbb{N}]\big)_{t\in\mathbb{R}}$ of unitary operators generated by the particle number operator $\mathbb{N}$ is a symmetry of all Hamiltonians $\mathbb{H}$ of the form (II.8), as these conserve particle number.

In Section VI it is demonstrated that in case of a repulsive potential, choosing a vanishing pairing operator is always favorable for the energy minimization, and the particle number symmetry is always preserved. For attractive potentials, this is not always the case and, depending on the model, the particle number symmetry is sometimes preserved, sometimes it is broken.

Translation Invariance in $\mathbb{R}^3$: Three-dimensional systems are translation invariant, if (the resolvent of) $\mathbb{H}$ commutes with all $U_{\vec{a}} \in \mathcal{T}_{\mathbb{R}^3}$, where $\mathcal{T}_{\mathbb{R}^3} = \{U_{\vec{a}} | \vec{a} \in \mathbb{R}^3\}$ and $U_{\vec{a}} = \exp[-i\vec{a}\cdot\vec{\mathbb{p}}]$ is the translation by $\vec{a} \in \mathbb{R}^3$. These translations are generated by $\vec{\mathbb{p}} = (\mathbb{p}_1, \mathbb{p}_2, \mathbb{p}_3)$, where $\mathbb{p}_\nu = \sum_{j,k=1}^{\infty} \langle f_j | (-i\partial_\nu)\rangle c^*(f_j)c(f_k)$ is the second quantization of the momentum operator $-i\partial_\nu$ in the ν^{th} coordinate direction.

From a physics point of view, it would be desirable to define these translational invariant systems with the (single-fermion) configuration space $\mathbb{R}^3$ as described above. This would necessitate general states, rather than density matrices, and ultimately require an operator algebraic framework which we cannot provide here.

Translation Invariance on a large Torus: To circumvent the problem related to the thermodynamic (i.e., infinite volume) limit it is customary to replace the configuration space $\mathbb{R}^3$ by a torus $\Lambda := (\mathbb{R}/L\mathbb{Z})^3$ of large, but finite, sidelength $L \gg 1$. The Hamiltonian $\mathbb{H}$ then commutes with translations $U_{\vec{a}} = \exp[-i\vec{a} \cdot \vec{\mathbb{p}}]$ by $\vec{a} \in \Lambda$ modulo L in $\mathcal{T}_\Lambda := \{U_{\vec{a}} | \vec{a} \in \Lambda\}$. The resulting model is called *Fermi Jellium* or *Fermi gas*, and one is interested in the limit $L \to \infty$ and in the energy per unit volume $e_{\rm gs} := \lim_{L\to\infty}\{L^{-3}\, E_{\rm gs}\}$. As the ground state energy (at fixed μ) is an *extensive* quantity, so are the generalized Hartree–Fock energy $E_{\rm gHF}$ and the restricted generalized Hartree–Fock energy $E_{\rm gHF}(\mathcal{T}_\Lambda)$. For this reason, we define the respective energies $e_{\rm gHF} := \lim_{L\to\infty}\{L^{-3}\, E_{\rm gHF}\}$ and $e_{\rm gHF}(\mathcal{T}_\Lambda) := \lim_{L\to\infty}\{L^{-3}\, E_{\rm gHF}(\mathcal{T}_\Lambda)\}$ per unit volume.

More than fifty years ago, Overhauser considered the above model with a repulsive interaction, for which the pairing operator vanishes and the generalized Hartree–Fock energy agrees with the original total Hartree–Fock energy. At high density, the paramagnetic state represented by a Slater determinant of plane waves occupying for both spin-up and spin-down electrons all momenta $k \in \Lambda^*$ below the Fermi energy, i.e., for which $\omega(k) \leq \mu$, is the natural tranlation invariant HF minimizer and yields $e_{\rm gHF}(\mathcal{T}_\Lambda)$. He demonstrated in [80, 81, 82], however, that a lower energy $e_{\rm gHF} < e_{\rm gHF}(\mathcal{T}_\Lambda)$ is produced by Slater determinants which are not translation invariant but represent a spin wave. The precise Hartree–Fock minimizer breaking the translation invariance is not known explicitly, but in a recent paper [47] Gontier, Hainzl, and Lewin estimated the difference $e_{\rm gHF}(\mathcal{T}_\Lambda) - e_{\rm gHF} > 0$ of the energies and proved that it is exponentially small in the interaction coupling. Thus, although the restricted HF energy is higher than the HF energy without restriction, the two terms agree to any order in powers of the coupling constant.

The BCS Model - Spin Invariance: We further introduce global spin transformations which rotate the spin variables $\mathbb{C}^2$ at each point in space by the *same* unitary transformations $S \in SU(2)$. As the Hamiltonian is invariant under such global spin rotations, this defines an additional symmetry $SU(2)$ of the system. (One variant of) The **BCS model** is now defined to be the restricted generalized Hartree–Fock energy $E_{\rm BCS}(\mathcal{T}_\Lambda \times SU(2))$. In the simplest model case, $E_{\rm BCS}(\mathcal{T}_\Lambda \times SU(2))$ can be explicitly computed thanks to the restriction of the variation to translation-invariant generalized 1-pdm which are in the same spin singlet state at any point in Λ. Additionally choosing j to be complex conjugation in Fourier space, the generalized Hartree–Fock energy functional is varied only over $\Gamma^{(1)} \in$

$\mathfrak{G}^{(1)}$ of the form

$$\Gamma^{(1)}(k,k') \;=\; \delta_{k,k'} \begin{pmatrix} \hat\gamma(k) \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) & \hat\alpha(k) \otimes \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \\ \hat\alpha(k)^* \otimes \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) & [\mathbf{1} - \hat\gamma(k)] \otimes \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \end{pmatrix}, \tag{VII.1}$$

where $\hat\gamma \in L^1(\Lambda^*; \mathbb{R}_0^+)$ and $\hat\alpha \in L^2(\Lambda^*)$ and $\Lambda^* = \frac{2\pi}{L}\mathbb{Z}^3$. Inserting this into the energy functional at zero temperature gives

$$\begin{aligned} \mathcal{E}_{\mathrm{BCS}}(\Gamma^{(1)}) \;=\; & \sum_{k\in\Lambda^*} \big(\omega(k) - \mu\big)\, \hat\gamma(k) \;+\; \frac{1}{2}\|\widehat{V}\|_1 \|\hat\gamma\|_1^2 \\ & -\; \frac{1}{2}\int_\Lambda V(x)\, |\gamma(x)|^2\, d^3x \;+\; \frac{1}{2}\int_\Lambda V(x)\, |\alpha(x)|^2\, d^3x\,, \end{aligned} \tag{VII.2}$$

where V, γ, and α are the inverse Fourier transform of $\widehat{V}$, $\hat\gamma$, and $\hat\alpha$, respectively. The potential V is assumed to be negative (attractive) for some part of Λ in order not to rule out nonvanishing $\alpha \neq 0$ to begin with. This model and its variants, questions of existence and uniqueness of its minimizers, the characterization of the resulting minimizers by the BCS gap equation and the analysis of its solution for zero and positive temperatures have been analyzed and physically interpreted by Hainzl and Seiringer and others in a remarkable series of papers [51, 42, 55, 54, 43, 21, 22], see [56] for a review. Spin symmetry breaking and a phase transition between a ferromagentic and a paramagnetic phase has been recently proved for a Hartree–Fock model like (VII.2) under the additional assumption of the absence $\alpha = 0$ of pairing, i.e., restriction to conserved particle numbers, by Gontier and Lewin in [48].

The Hubbard Model at Half-Filling: An example of a translation and spin invariant model, for which both the translation invariance and the spin invariance are broken and the generalized Hartree–Fock minimizers can be explicitly computed is the Hubbard model at half-filling. Bach et al. proved in [12] that these symmetries are indeed broken and determined all Hartree-Fock mimimizers explicitly.

Periodic structures: If the configuration space $\mathbb{R}^3$ is again a torus $\Lambda_L := (\mathbb{R}/L\mathbb{Z})^3$ for some large integer $L \gg 1$ then the Hamiltonian often commutes only with integer translations $\vec{a}$ contained in the subgroup $\mathbb{Z}_L^3 = (\mathbb{Z}/L\mathbb{Z})^3 \subset \Lambda$, leading to the symmetry $\mathcal{S}_{\mathbb{Z}_L^3}$ of $\mathbb{H}$. A typical example is a system of the form $\mathbb{H} = \mathbb{h} + \frac{1}{2}\mathbb{V}$ with $\mathbb{V}$ having the full translation symmetry $\mathcal{S}_\Lambda$ but $\mathbb{h}$ having only the smaller symmetry $\mathcal{S}_{\mathbb{Z}_L^3}$ due to the presence of a periodic external potential.

The closed shell theorem described above does not only hold for Coulomb systems, but for general N fermion systems with a repulsive interaction potential (for

which the generalized and the original Hartree–Fock approximation coincide). In general, it may fail, however, in case of restricted Hartree–Fock minimizers. For periodic systems, the existence of minmizer was established by Catto, Le Bris, and Lions in [25] where the 1-pdm are restricted to those which are invariant under (integral) lattice translations. Ghimenti and Lewin have later shown in [46] a kind of closed shell theorem and proved that the minimizer is a projection onto the smallest energies of the corresponding Hartree–Fock effective operator.

Acknowledgement: I thank Mathieu Lewin and Konstantin Merz for numerous very helpful corrections and suggestions.

References

[1] A. Anantharaman and E. Canés. Existence of minimizers for Kohn–Sham models in quantum chemistry. *Annales de l'I.H.P. Analyse non linéaire*, 26:2425–2455, 2009.

[2] H. Araki. On quasifree states of CAR and Bogoliubov automorphisms. *Publ. RIMS Kyoto*, 6:385–442, 1970/71.

[3] V. Bach. Error bound for the Hartree-Fock energy of atoms and molecules. *Commun. Math. Phys.* , 147:527–548, 1992.

[4] V. Bach. Accuracy of mean field approximations for atoms and molecules. *Commun. Math. Phys.* , 155:295–310, 1993.

[5] V. Bach, J.-M. Barbaroux, B. Helffer, and H. Siedentop. On the stability of the relativistic electron-positron field. *Commun. Math. Phys.*, 201:445–460, 1999.

[6] V. Bach, S. Breteaux, H.-K. Knörr, and E. Menge. Generalized one-particle density matrices and quasifree states. *J. Math. Phys.*, 55:012101, 2014. dx.doi.org/10.1063/1.4853875.

[7] V. Bach, J. Fröhlich, and L. Jonsson. Bogolubov-Hartree-Fock mean field theory for neutron stars and other systems with attractive interactions. *J. Math. Phys.*, 50:102102, 2009. doi:10.1063/1.3225565.

[8] V. Bach and A. Hach. On the ultraviolet limit of the Pauli-Fierz Hamiltonian in the Lieb-Loss model. *Ann. Henri Poincaré*, 2021.

[9] V. Bach, H.K. Knörr, and E. Menge. Fermion correlation inequalites derived from g- and p-conditions. *Doc. Math.*, 17:449–470, 2012.

[10] V. Bach, H.K. Knörr, and E. Menge. Representability conditions by Grassmann integration. *Adv. Theor. Math. Phys.*, 19(6):1141–1182, 2015.

[11] V. Bach, E. H. Lieb, M. Loss, and J. P. Solovej. There are no unfilled shells in Hartree-Fock theory. *Phys. Rev. Lett.* , 72(19):2981–2983, 1994.

[12] V. Bach, E. H. Lieb, and J. P. Solovej. Generalized Hartree-Fock theory and the Hubbard model. *J. Stat. Phys.* , 76:3–90, 1994.

[13] J.-M. Barbaroux, M. Esteban, and E. Séré. Some connections between Dirac–Fock and electron-positron Hartree–Fock. *Ann. Henri Poincaré*, 6(1):85–102, 2005.

[14] J.-M. Barbaroux, W. Farkas, B. Helffer, and H. Siedentop. On the Hartree–Fock equations of the electron-positron field. *Comm. Math. Phys.*, 255(1):131–159, 2005.

[15] J. Bardeen, L.N. Cooper, and J.R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108(5):1175–1204, 1957.

[16] K. Baumann. When is a field theory a generalized free field? *Commun. Math. Phys.*, 43:221–223, 1975.

[17] K. Baumann and G. Hegerfeldt. A noncommutative Marcinkiewicz theorem. *Publ. RIMS, Kyoto Univ.*, 21:191–204, 1985.

[18] M. Born, W. Heisenberg, and P. Jordan. Zur Quantenmechanik. II. *Z. Phys.*, 35(8-9):557–615, 1926.

[19] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annal. Phys.* , 84:457, 1927.

[20] B. Braams, J. Percus, and Z. Zhao. The t1 and t2 representability conditions. In D. Mazziotti, editor, *Reduced-Density-Matrix Mechanics*, volume 134 of *Advances in Chemical Physics*, chapter 3, pages 93–101. Wiley-Interscience, 2007.

[21] G. Bräunlich, C. Hainzl, and R. Seiringer. Translation-invariant quasi-free states for fermionic systems and the BCS approximation. *Rev. Math. Phys.*, 26(7):1450012, 2014.

[22] G. Bräunlich, C. Hainzl, and R. Seiringer. Bogolubov–Hartree–Fock theory for strongly interacting fermions in the low density limit. *Math. Phys. Anal. Geom.*, 19:13, 2016.

[23] E. Cancès and C. Le Bris. On the convergence of SCF algorithms for the Hartree–Fock equations. *ESAIM: M2AN*, 34(4):749–774, 2000.

[24] E. Cancès, M. Lewin, and G. Stoltz. The electronic ground state energy problem: a new reduced density matrix approach. *J. Chem. Phys.*, 125:064101–064106, 2006.

[25] I. Catto, C. Le Bris, and P.-L. Lions. On the thermodynamic limit for Hartree–Fock type models. *Ann. Inst. H. Poincaré Anal. Non Lineaire*, 18(6):687–760, 2001.

[26] P. Chaix and D. Iracane. From quantum electrodynamics to mean-field theory: I. The Bogoliubov-Dirac-Fock formalism. *J. Phys. B.*, 22(23):3791–3814, December 1989.

[27] P. Chaix, D. Iracane, and P.L. Lions. From quantum electrodynamics to mean-field theory: II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation. *J. Phys. B.*, 22(23):3815–3828, December 1989.

[28] A. J. Coleman. Structure of fermion density matrices. *Rev. Mod. Phys.* , 35(3):668–689, 1963.

[29] A.J. Coleman and V.I. Yukalov. *Reduced Density Matrices - Coulson's Challenge*. Lecture Notes in Chemistry. Springer, Berlin, Heidelberg, 2000.

[30] C. A. Coulson. Present state of molecular structure calculations. *Rev. Mod. Phys.*, 32(2):170–177, 1960.

[31] J. Derezinski, M. Napiorkowski, and J.P. Solovej. On the minimization of Hamiltonians over pure Gaussian states. In *Complex Quantum Systems. Analysis of Large Coulomb Systems*, volume 24 of *Lecture Notes Series*. Inst. for Math. Studies, National University of Singapore, 2013.

[32] P. A. M. Dirac. Note on exchange phenomena in the Thomas-Fermi atom. *Proc. Cambridge Philos. Soc.*, 26:376–385, 1931.

[33] P.A.M. Dirac. The quantum theory of the emission and absorption of radiation. *Proc. Royal Soc. London Series*, 114(A):243, 1927.

[34] P.A.M. Dirac. The quantum theory of the electron. *Proc. Royal Soc. London*, 117(A):610–624, 1928.

[35] P.A.M. Dirac. The quantum theory of the electron II. *Proc. Royal Soc. London*, 118(A):351–361, 1928.

[36] F. Dyson and A. Lenard. Stability of matter I. *J. Math. Phys.* , 8:423–434, 1967.

[37] F. Dyson and A. Lenard. Stability of matter II. *J. Math. Phys.* , 9:698–711, 1967.

[38] R.M. Erdahl. On the structure of the diagonal conditions. *Int. J. Quant. Chem.*, 13:731–736, 1978.

[39] R.M. Erdahl. Representability. *Int. J. Quant. Chem.*, 13:697–718, 1978.

[40] M. Esteban and E. Séré. Solutions of the Dirac-Fock equations for atoms and molecules. *Commun. Math. Phys.*, 203(3):499–530, 1999.

[41] V. Fock. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* , 61:126–148, 1930.

[42] R. Frank, C. Hainzl, S. Naboko, and R. Seiringer. The critical temperature for the BCS equation at weak coupling. *J. Geom. Anal.*, 17(4):549–567, 2007.

[43] A. Freiji, C. Hainzl, and R. Seiringer. The gap equation for spin-polarized fermions. *J. Math. Phys.*, 53:012101, 2012.

[44] G. Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Arch. Ration. Mech. Anal.*, 169:35–71, 2003.

[45] C. Garrod and J. K. Percus. Reduction of the N-particle variational problem. *J. Math. Phys*, 5(12):1756–1776, 1964.

[46] M. Ghimenti and M. Lewin. Properties of periodic Hartree-Fock minimizers. *Calc. Var. Partial Diff. Eq.*, 35:39–56, 2009.

[47] D. Gontier, C. Hainzl, and M. Lewin. Lower bound on the Hartree–Fock energy of the electron gas. *Phys. Rev. A*, 99:052501, 2019.

[48] D. Gontier and M. Lewin. Spin symmetry breaking in the translation-invariant Hartree–Fock uniform electron gas. *SIAM J. Math. Anal.*, 51(4), 2019.

[49] A. Gottlieb and N. Mauser. Properties of nonfreeness: An entropy measure of electron correlation. *Intl. J. Q. Inform.*, 05(06):815–827, 2007.

[50] M. Griesemer and F. Hantsch. Unique solutions to Hartree-Fock equations for closed shell atoms. *Arch. Ration. Mech. Anal.*, 203(3):883–900, 2012.

[51] C. Hainzl, E. Hamza, R. Seiringer, and J.P. Solovej. The BCS functional for general pair interactions. *Commun. Math. Phys.*, 281(2):349–367, 2008.

[52] C. Hainzl, M. Lewin, and E. Séré. Existence of a stable polarized vacuum in the Bogoliubov–Dirac–Fock approximation. *Commun. Math. Phys.*, 257(3):275–283, 2005.

[53] C. Hainzl, M. Lewin, and E. Séré. Existence of atoms and molecules in the mean-field approximation of no-photon quantum electrodynamics. *Arch. Ration. Mech. Anal.*, 192(3):453–499, 2009.

[54] C. Hainzl and R. Seiringer. The BCS critical temperature for potentials with negative scattering length. *Lett. Math. Phys.*, 84(2-3):99–107, 2008.

[55] C. Hainzl and R. Seiringer. Critical temperature and energy gap for the BCS equation. *Phys. Rev. B*, 77:184517, 2008.

[56] C. Hainzl and R. Seiringer. The BCS functional of superconductivity and its mathematical properties. *J. Math. Phys.*, 57:021101, 2016.

[57] D. Hartree. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. *Proc. Cambridge Philos. Soc.* , 24:89–132, 1928.

[58] W. Heisenberg. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Physik*, 33:879–893, 1925.

[59] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B 864–871, 1964.

[60] M. Huber and H. Siedentop. Solutions of the Dirac–Fock equations and the energy of the electron-positron field. *Arch. Ration. Mech. Anal.*, 184:1–22, 2007.

[61] W. Hughes. *An Atomic Energy Lower Bound that Gives Scott's Correction*. PhD thesis, Princeton, Department of Mathematics, 1986.

[62] V. Ja. Ivrii and I. M. Sigal. Asymptotics of the ground state energies of large Coulomb systems. *Ann. Math.* , 138(2):243–335, 1993.

[63] M. Ernzerhof J.P. Perdew, K. Burke. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[64] E. Lenzmann and M. Lewin. Minimiziers for the Hartree-Fock-Bogoliubov theory of neutron stars and white dwarfs. *Duke Math. J.*, 152(2):257–315, 2010.

[65] M. Levy. Universal variational functionals of electron densities, first order density matrices, and natural spin-orbitals and solution of the V-representability problem. *Proc. Natl. Acad. Sci. USA*, 76:6062–6065, 1979.

[66] M. Lewin. Geometric methods for nonlinear many-body quantum systems. *J. Funct. Anal.*, 260:3535–3595, 2011.

[67] M. Lewin. Existence of Hartree–Fock excited states for atoms and molecules. *Lett. Math. Phys.*, 108:985–1006, 2018.

[68] M. Lewin and S. Paul. A numerical perspective on Hartree–Fock–Bogoliubov theory. *ESAIM: M2AN*, 48(1):53–86, 2014.

[69] E. H. Lieb. Variational principle for many-fermion systems. *Phys. Rev. Lett.* , 46(7):457–459, 1981.

[70] E. H. Lieb and B. Simon. The Hartree-Fock theory for Coulomb systems. *Commun. Math. Phys.* , 53:185–194, 1977.

[71] E. H. Lieb and B. Simon. The Thomas-Fermi theory of atoms, molecules and solids. *Adv. Math.* , 23:22–116, 1977.

[72] E.H. Lieb. Density functional for Coulomb systems. *Int. J. Quantum Chem.*, 24:243–277, 1983.

[73] Elliott H. Lieb and Walter E. Thirring. Bound for the kinetic energy of Fermions which proves the stability of matter. *Phys. Rev. Lett.* , 35(11):687–689, September 1975.

[74] P.-L. Lions. Solutions of Hartree–Fock equations for Coulomb systems. *Commun. Math. Phys.*, 109:33–97, 1987.

[75] Y.-K. Liu, M. Christandl, and F. Verstraete. Quantum computational complexity of the n-representability problem: QMA complete. *Phys. Rev. Lett.*, 98:110503, 2007.

[76] J. Marcinkiewicz. Sur une propriété de la loi de Gauß. *Math. Z.*, 44:612–618, 1939.

[77] D. Mazziotti and R. Erdahl. Uncertainty relations and reduced density matrices: Mapping many-body quantum mechanics onto four particles. *Phys. Rev. A*, 63:042113, 2001.

[78] K. Naftchi-Ardebili, N.W. Hau, and D.A. Mazziotti. Rank restriction for the variational calculation of two-electron reduced density matrices of many-electron atoms and molecules. *Phys. Rev. A*, 84:052506, 2011.

[79] P.T. Nam, M. Napiórkowski, and J.P. Solovej. Diagonalization of bosonic quadratic Hamiltonians by Bogoliubov transformations. *J. Func. Anal.*, 270(11):4043–4368, 2016.

[80] A.W. Overhauser. Giant spin density waves. *Phys. Rev. Lett.*, 4(9):462–465, 1960.

[81] A.W. Overhauser. Spin density waves in an electron gas. *Phys. Rev.*, 128(3):1437–1452, 1962.

[82] A.W. Overhauser. Exchange and correlation instabilities of simple metals. *Phys. Rev.*, 167(3):691–698, 1968.

[83] E. Paturel. Solutions of the Dirac–Fock equations without projector. *Ann. Henri Poincaré*, 1(6):1123–1157, 2000.

[84] W. Pauli. *Probleme der modernen Physik*. S. Hirzel, Leipzig, 1928.

[85] A.K. Rajagopal and E.C.G. Sudarshan. Some generalizations of the Marcinkiewicz theorem and its implications to certain approximation schemes in many-particle physics. *Phys. Rev. A*, 10(5):1852–1857, 1974.

[86] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I–IV*. Academic Press, San Diego, 2 edition, 1980.

[87] D.W. Robinson. A theorem concerning the positive metric. *Commun. Math. Phys.*, 1:89–94, 1965.

[88] M. Salmhofer. Continuous renormalization for fermions and Fermi liquid theory. *J. Stat. Phys.*, 134(5):941–952, 2009.

[89] E. Schrödinger. Quantisierung als Eigenwertproblem. *Ann. d. Physik*, 79:489, 1926.

[90] H. K. H. Siedentop and R. Weikard. On the leading energy correction for the statistical model of the atom: Interacting case. *Commun. Math. Phys.* , 112:471–490, 1987.

[91] H. K. H. Siedentop and R. Weikard. On the leading correction of the Thomas-Fermi model: Lower bound – with an appendix by A. M. K. Müller. *Invent. Math.* , 97:159–193, 1989.

[92] J. C. Slater. A simplification of the Hartree-Fock method. *Phys. Rev.* , 81:385–390, 1951.

[93] John C. Slater. A note on Hartree's method. *Phys. Rev.* , 35:210–211, 1930.

[94] J.P. Solovej. Many body quantum mechanics. March 2014.

[95] J.P. Solovej, T. Østergaard Sørensen, and W. Spitzer. The relativistic Scott correction for atoms and molecules. *Commun. Pure Appl. Math.*, 63(1):39–118, 2010.

[96] J.P. Solovej and W. Spitzer. A new coherent states approach to semiclassics which gives Scott's correction. *Commun. Math. Phys.*, 241(2-3):383–420, 2003.

[97] B. Swirles. The relativistic self-consistent field. *Proc. Roy. Soc. A*, 152:625–649, 1935.

[98] M. Titulaer. Marcinkiewicz's theorem and approximation schemes for many-particle correlation functions: Comment on a paper by Rajagopal and Sudarshan. *Phys. Rev. A*, 11(6):2204–2208, 1975.

[99] Z. Zhao, B.J. Braams, M. Fukuda, M.L. Overton, and J.K. Percus. The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions. *J. Chem. Phys.*, 120:2095––2104, 2004.