Adv. Theor. Math. Phys. **5** (2001) 1169–1205

Derivation of the nonlinear Schrödinger equation from a many body Coulomb system

László Erdős
School of Mathematics, GeorgiaTech
Atlanta, GA 30332
lerdos@math.gatech.edu

Horng-Tzer Yau
Courant Institute, New York University
251 Mercer Str. New York, NY 10012
yau@cims.nyu.edu

Abstract

We consider the time evolution of N bosonic particles interacting via a mean field Coulomb potential. Suppose the initial state is a product wavefunction. We show that at any finite time the correlation functions factorize in the limit $N \to \infty$. Furthermore, the limiting one particle density matrix satisfies the nonlinear Hartree equation. The key ingredients are the uniqueness of the BBGKY hierarchy for the correlation functions and a new apriori estimate for the many-body Schrödinger equations.

---

e-print archive: http://xxx.lanl.gov/math-ph/0111042

AMS 2000 Subject Classification: 35Q55, 45F15, 81Q05, 81V70

Keywords: Nonlinear Hartree equation, Mean field Coulomb system of bosons, BBGKY hierarchy

1 Introduction

It is well-known that the Schrödinger equation governs the dynamics of non-relativistic quantum systems. The basic existence and uniqueness of the Schrödinger equation were studied extensively for many systems, including the important Coulomb systems. Despite these efforts, there are few cases that the solutions to the many-body Schrödinger equation can be described in reasonable details. One such case is a system of N weakly interacting bosons, or more precisely, N bosons interacting with a pair potential whose coupling constant is of order N^{-1}.

For such a system, we expect that the potential acting on any single particle is approximately generated by the average density of all other particles. Therefore, there are no correlations generated by the dynamics and all correlations of the state at a later time are almost all due to the initial correlations. In the simplest scenario, i.e., when the initial state has no correlation, the correlations at a later time should be negligible as well. We consider initial states of N bosons that are described by products of one-particle wave functions. Therefore, our goal is to show that, if the initial N-body wave function is a product of a one-particle wave function, then it is well-approximated by a product of some one-particle wave function at later times as well. Assuming this is correct, it is easy to see that the one particle wave function satisfies a nonlinear Schrödinger equation whose nonlinearity is given by the convolution of the potential with the mass density of the one particle wave function, i.e., $V * |\psi|^2$. This equation is typically called the *Hartree equation.* In the special case when the pair interaction is Coulomb, it is called the Schrödinger-Poisson system. As usual for the Schrödinger equation, we can always cast it in the form of an equation for the density matrix. We shall use the same name for both setups.

This problem has a long history and has been considered by many authors. For regular pair-potentials, it was solved by Hepp [5] in the context of field operators and by Spohn [9] using density matrices. Hepp's method requires differentiability of the potential, while Spohn uses only the boundedness of the potential. Ginibre and Velo [3] have greatly extended these results to include singular potentials (including Coulomb). However they worked in the second quantized framework and the quasi-free initial state used in [3] describes special classes of excitations around the product state. In particular, this method does not apply to wave-functions with a fixed number of particles.

The basic tool to analyze the many-body dynamics is the BBGKY hierarchy. We notice that the time derivative of the k-particle correlation function (more precisely, k-particle density matrix) of an N-body state is given by the $k+1$ particle correlation function for each $k = 1, 2, \dots N-1$. This gives a hierarchy of coupled evolution equation for all k correlation functions up to N, called the *Schrödinger hierarchy.* Without further limiting procedure, this system is just a rewriting of the original Schrödinger equation. If we take the *formal* $N \to \infty$ limit, it converges to an infinite system of equation, called the *BBGKY hierarchy*, for the k-particle density matrices for $k = 1, 2, \dots$. For a product initial state this infinite hierarchy has a trivial solution of product form built up from the solution of the Hartree equation. Although the N-body hierarchy has a unique solution, the infinite hierarchy may, in principle, admit more than one solution. Thus our task is to establish the convergence of the Schrödinger hierarchy to the BBGKY hierarchy as $N \to \infty$, and prove the uniqueness of the solution to the BBGKY hierarchy.

Both the BBGKY and the Schrödinger hierarchy can be put into systems of integral equations. For bounded potential, these systems can be solved by iteration which converges in the trace norm of the density matrices. Using this idea on the Schrödinger hierarchy, Spohn proved that the k-point correlation functions of the N body system are given by convergent power series with an error estimate uniformly in N. Thus one can take the limit $N \to \infty$ and one obtains that the k-point correlation functions are indeed given by the products of the solution to the nonlinear Hartree equation.

Bardos et. al. have followed a different route [1]. They showed that any w^*-limit point of the solutions to the N-body hierarchy satisfies the infinite BBGKY hierarchy. This proof requires the potential to be bounded below and in $L^2 + L^\infty$. If, in addition, the potential is bounded, then the sequence of correlation functions is shown to be w^*-precompact in the trace norm. The uniqueness of the solution to the infinite hierarchy is established by controlling the trace norm similarly to [9]. In particular, this result establishes the convergence to the BBGKY hierarchy for the repulsive Coulomb systems which was not covered in Spohn's work.

The main question is the uniqueness of the BBGKY hierarchy in the case of singular potentials. Furthermore, the convergence in the attractive Coulomb case should be resolved since it describes the important gravitational system. Technically, all the uniqueness methods rely heavily on the boundedness of the potential via the estimate $\mathrm{Tr}|V\gamma| \leq \|V\|_\infty \mathrm{Tr}|\gamma|$ for a density matrix γ. The Coulomb potential, i.e., $V(x) = \pm|x|^{-1}$, is unbounded, but Hardy's inequality allows us to control it in H^1-norm, i.e.,

$$\mathrm{Tr}\, \left|V\gamma\right| \leq C\big(\, \mathrm{Tr}\, \nabla\gamma\nabla + \mathrm{Tr}\, \gamma \,\big)\ .$$

Therefore the solution to the BBGKY hierarchy should be unique under the right Sobolev norm. Since we have correlation functions of arbitrary number of particles, the Sobolev space we choose is the iterative Sobolev space (see Section 3 for the precise definition) which is weaker than the usual one but sufficient for the uniqueness. This poses the problem of estimating the iterative Sobolev norm for correlation functions. This estimate can only be obtained from the original N-body Schrödinger equation for which we know only the conservation of mass and energy. Notice that in [9] and [1] the only estimates obtained are trace norm and H^1 norm bounds, which are consequences of the conservation of mass and energy for the N-body Schrödinger equation.

Since we can only establish the uniqueness of the BBGKY hierarchy in the iterative Sobolev norm, we need to establish an apriori estimate that the k-correlation functions are bounded in such a norm. It should be emphasized that to establish such estimates directly for the correlation functions of a Coulomb system is very difficult due to the $|x|^{-1}$ singularity. However, it is feasible for a cutoff Coulomb systems with an N dependent cutoff. The main idea here is to take the energy to a higher power and deduce from there the estimate on the iterative Sobolev norm. Notice that the third power of the potential is already not in L^1. Hence it does not even define a meaningful operator in the usual sense. This is part of the technical reason we need to perform a truncation for the Coulomb singularity.

We now need to control the original evolution and the cutoff dynamics. Here we use the conservation of the L^2 norm to control the difference of these two evolutions. Since to control an N-body system in general produces a factor N, our cutoff has to be rather small. The various restrictions on the cutoff scale finally give the choice of order $o(N^{-1/2})$.

This work was partly inspired by the work of [1] and partly by the question posed by J. Yngvason regarding the derivation of the time dependent Gross-Pitaevskii equation from the many-body Schrödinger equation. Unfortunately, our method, as it stands, still cannot prove the convergence to the Gross-Pitaevskii equation. Another possible motivation for studying high density bosons with Coulomb interaction is that electrically charged ions may have bosonic statistics.

This work was supported by NSF grants DMS-9970323 (L.E.) and DMS-0072098 (H.-T. Y.) The discussions with C. Bardos, F. Golse, N. Mauser and J. Yngvason are gratefully acknowledged. The results of this work and [1] were jointly outlined in [2].

2 Many-body Schrödinger evolution with mean field potential

Consider a system of N-bosons weakly interacting with a Coulomb potential. The dynamics of such a system is governed by the Schrödinger equation with the Hamiltonian

$$H_N = -\frac{1}{2}\sum_{\ell=1}^{N} \Delta_{x_\ell} + \frac{1}{N}\sum_{\ell<j} V(x_\ell - x_j)$$

defined on $\bigotimes_{\ell=1}^N L^2(\mathbf{R}^3)$, where $V(x) = \pm\mu|x|^{-1}$ with $\mu > 0$. The wave functions are symmetric functions of N variables. Since the Hamiltonian is symmetric, the wave function at the time t will be symmetric as long as the initial data is symmetric.

The Schrödinger equation is given by

$$i\partial_t \Psi_{N,t} = H_N \Psi_{N,t}, \tag{2.1}$$

with the initial data specified at the time $t = 0$. The equation can be solved explicitly by $\Psi_{N,t} = e^{-itH_N}\Psi_{N,0}$. Let $\gamma_{N,t} = \pi_{\Psi_{N,t}}$ be the projection operator in $L^2(\mathbf{R}^3)$ associated with the wave function $\Psi_{N,t}$. The Schrödinger equation is equivalent to the operator equation

$$i\partial_t \gamma_{N,t} = [H_N, \gamma_{N,t}] \, . \tag{2.2}$$

This is a more general setup since it allows a general density matrix not coming from a wave function. The n-point density matrix of $\gamma_{N,t}$ is defined by

$$\begin{aligned}\gamma_{N,t}^{(n)}(x_1,\ldots x_n; x_1',\ldots x_n') := \int & \Psi_{N,t}(x_1,\ldots x_n, x_{n+1},\ldots x_N) \\ & \times\overline{\Psi_{N,t}}(x_1',\ldots x_n', x_{n+1},\ldots x_N)\mathrm{d}x_{n+1}\ldots \mathrm{d}x_N\end{aligned} \tag{2.3}$$

if $n \leq N$ and $\gamma_{N,t}^{(n)} := 0$ otherwise. The normalization is $(n \leq N)$

$$\mathrm{Tr}\, \gamma_{N,t}^{(n)} = 1 \, . \tag{2.4}$$

It is a simple calculation that the n-point density matrices of the solution to the Schrödinger equation satisfy the finite hierarchy

$$\begin{aligned}\gamma_{N,t}^{(k)}&(x_1,\dots x_k;x_1',\dots x_k') \qquad (2.5)\\ &= \mathcal{U}_{N,k}(t)\gamma_{N,0}^{(k)}(x_1,\dots x_k;x_1',\dots x_k') + (-i)\sum_{\ell=1}^{k}\int_0^t \mathrm{d}s\,\mathcal{U}_{N,k}(t-s)\\ &\Bigg[\int \mathrm{d}x_{k+1}\Big(V(x_\ell - x_{k+1}) - V(x_\ell' - x_{k+1})\Big)\\ &\times\gamma_{N,s}^{(k+1)}(x_1,\dots x_k,x_{k+1};x_1',\dots x_k',x_{k+1})\Bigg]\end{aligned}$$

where $\mathcal{U}_{N,k}(t)\gamma = e^{-itH_N^{(k)}}\gamma e^{itH_N^{(k)}}$ and

$$H_N^{(k)} = -\frac{1}{2}\sum_{\ell=1}^{k}\Delta_\ell + \frac{1}{N}\sum_{\ell<j}^{k} V(x_\ell - x_j)\ .$$

If we take the limit $N\to\infty$ and neglect all lower order terms, this system converges to the BBGKY hierarchy given by the following infinite system of equations for density matrices

$$\begin{aligned}\gamma_t^{(k)}&(x_1,\dots x_k;x_1',\dots x_k') \qquad (2.6)\\ &= \mathcal{U}_k(t)\gamma_0^{(k)}(x_1,\dots x_k;x_1',\dots x_k') + (-i)\sum_{\ell=1}^{k}\int_0^t \mathrm{d}s\,\mathcal{U}_k(t-s)\\ &\Bigg[\int \Big(V(x_\ell - x_{k+1}) - V(x_\ell' - x_{k+1})\Big)\\ &\times\gamma_s^{(k+1)}(x_1,\dots x_k,x_{k+1};x_1',\dots x_k';,x_{k+1})\mathrm{d}x_{k+1}\Bigg]\end{aligned}$$

where $\mathcal{U}_k(t)\gamma := \left(\prod_{j=1}^k e^{-it\Delta_j/2}\right)\gamma\left(\prod_{j=1}^k e^{-it\Delta_j/2}\right)$, $k=1,2,\dots$. This is the integral form of an infinite system of evolution equations with initial data $\Gamma_0 := \left(\gamma_0^{(1)},\gamma_0^{(2)},\dots\right)$.

Suppose the initial data is of product form and given by

$$\Psi_{N,0}(x_1,\dots x_N) := \prod_{j=1}^{N}\psi_0(x_j) \qquad (2.7)$$

for some initial one-particle wave function ψ_0. We always assume that ψ_0 is normalized in $L^2(\mathbf{R}^3)$, i.e. $\|\psi_0\|_2 = 1$. For this initial data, the k-point density matrix, $\gamma_{N,0}^{(k)}$ is, simply the tensor product $\bigotimes_1^k \gamma_{\psi_0}$ of the one-point density matrix γ_{ψ_0}, which is the projection matrix onto ψ_0. One can check that if the initial data to the BBGKY hierarchy is of product form, then there is a special solution to the BBGKY hierarchy which is of product form and the one-particle density matrix satisfies the nonlinear Schrödinger equation

$$i\partial_t \gamma_t = \Big[-\frac{1}{2}\Delta_x + \Big(\int V(\cdot - z)\gamma_t(z,z)\mathrm{d}z \Big) \, , \, \gamma_t \Big] \, . \tag{2.8}$$

If we denote the kernel by $\gamma_t(x,x')$, then (2.8) is equivalent to

$$\begin{aligned} i\partial_t \gamma_t(x,x') = & -\frac{1}{2}\Big[\Delta_x - \Delta_{x'}\Big]\gamma_t(x,x') \\ & + \int \mathrm{d}z \Big[V(|x-z|) - V(x'-z) \Big] \gamma_t(z,z)\gamma_t(x,x') \, . \end{aligned} \tag{2.9}$$

We can put this equation into a more familiar form. If $\gamma_0(x,x') := \psi_0(x)\overline{\psi_0(x')}$, then $\gamma_t(x,x') := \psi_t(x)\overline{\psi_t(x')}$ where ψ_t satisfies

$$i\partial_t \psi_t(x) = -\frac{1}{2}\Delta_x \psi_t(x) + \Big(\int V(x-z)|\psi_t(z)|^2 \mathrm{d}z \Big) \psi_t(x) \tag{2.10}$$

with initial data $\psi_{t=0} = \psi_0$. This equation was studied extensively by Ginibre and Velo. In particular, the equation preserves the L^2 norm and the energy and we have [4]

$$\sup_{t \geq 0} \|\psi_t\|_{H^1} < \infty \tag{2.11}$$

if the H^1 norm of the initial condition $\|\psi_0\|_{H^1}$ is finite. Although the equation is nonlinear, the normalization condition $\|\psi_0\|_2 = 1$ can be assumed without loss of generality at the expense of changing μ.

Therefore, if we can justify the limiting procedure and prove the uniqueness of the BBGKY hierarchy, the evolution of the weakly interacting N-bosons can be understood by a one-body nonlinear Schrödinger equation. We first describe the topology for the limiting procedure.

Denote by H the Hilbert space $L^2(\mathbf{R}^3)$ and let $\mathcal{L}(H)$ be the set of bounded operators with operator norm $\|\cdot\|$. We let $\mathcal{L}^1(H)$ be the set of trace class operators with the norm $\|\gamma\|_1 := \mathrm{Tr}|\gamma|$. The set of density matrices, $\widehat{\mathcal{L}}^1$, is defined as the subset of nonnegative self-adjoint trace class operators. Let $\mathcal{K} := \mathcal{K}(H)$ be the set of compact operators equipped with the operator norm. It is well-known (Theorem VI.26 in Vol.I. of [7]) that the dual space of the compact operators is the space of trace class operators, i.e., $(\mathcal{K}, \|\cdot\|)^* = (\mathcal{L}^1, \|\cdot\|_1)$. This gives rise to the variational characterization of the trace norm:

$$\|A\|_1 = \sup_{K \in \mathcal{K}(H)\,:\,\|K\|=1} \Big| \operatorname{Tr} AK \Big| . \tag{2.12}$$

The w^*-topology on $\mathcal{L}^1$ is induced by a family of seminorms $A \to |\operatorname{Tr} AK|$ that are indexed by the family of compact operators $K \in \mathcal{K}$.

Denote by $H^{\otimes k} := \bigotimes_{i=1}^k H = \bigotimes_{i=1}^k L^2(\mathbf{R}^3)$ the k-tensor product of $L^2(\mathbf{R}^3)$. We can define the trace norm on $H^{\otimes k}$ and extend the duality from $k=1$ to all k. Define the space

$$\mathcal{C} := \Big\{ \Gamma = (\gamma^{(1)}, \gamma^{(2)}, \dots) \ : \ \gamma^{(k)} \in \mathcal{L}^1(H^{\otimes k}) \Big\} = \prod_{k=1}^{\infty} \mathcal{L}^1(H^{\otimes k}) .$$

We equip this set with the product of the w^*-topologies on each component. Thus the convergence in $\mathcal{C}$ is characterized by the following property: $\Gamma_n \to \Gamma$ as $n \to \infty$, if for each k we have $\gamma_n^{(k)} \to \gamma^{(k)}$ in w^* sense in $\mathcal{L}^1(H^{\otimes k})$.

Let $\langle t \rangle := (1+t^2)^{1/2}$. We define the set

$$L^\infty(\mathbf{R}_+, \langle t \rangle^{-1} \mathrm{d}t, \mathcal{C}) := \prod_{k=1}^{\infty} L^\infty\Big(\mathbf{R}_+, \langle t \rangle^{-1}\mathrm{d}t, \mathcal{L}^1(H^{\otimes k})\Big)$$

i.e., the set of functions $\gamma^{(k)}(t) : \mathbf{R}_+ \to \mathcal{L}^1(H^{\otimes (k)})$ with $\sup_t \langle t \rangle^{-1} \|\gamma^{(k)}(t)\|_1 < \infty$. We equip this set with the product of the w^*-topologies on each factor. On each factor the w^*-topology is given by seminorms

$$\gamma^{(k)}(t) \to \Big| \int_0^\infty \operatorname{Tr}\Big[\, K(t) \gamma^{(k)}(t) \, \Big] \, \mathrm{d}t \Big|$$

where $K(t) \in L^1(\mathbf{R}_+, \langle t \rangle \mathrm{d}t, \mathcal{K}(H^{\otimes k}))$, i.e. $\int_0^\infty \|K(t)\| \langle t \rangle \mathrm{d}t < \infty$.

Let

$$\Gamma_{N,t} := \Big(\gamma_{N,t}^{(1)}, \gamma_{N,t}^{(2)}, \dots \Big) \in \mathcal{C} . \tag{2.13}$$

For any one-particle wave function ψ we define

$$\gamma_\psi^{(n)}(x_1, \dots x_n; x_1', \dots x_n') := \prod_{j=1}^n \gamma_\psi(x_j, x_j') .$$

Suppose $\psi_t \in H^1(\mathbf{R}^3)$ is a solution to the nonlinear Schrödinger equation (2.10), then the collection of density matrices

$$\Gamma_{\psi_t} := \left(\gamma^{(1)}_{\psi_t}, \gamma^{(2)}_{\psi_t}, \dots\right) \in \mathcal{C} \tag{2.14}$$

is a solution to (2.6). We now state the main theorem:

Theorem 2.1. *Let $V(x) = \pm\mu|x|^{-1}$ be the repulsive or attractive Coulomb potential with some $\mu > 0$. Assume that $\psi_0 \in H^2(\mathbf{R}^3)$ and let ψ_t be the solution of (2.10). Let $\Gamma_{N,t}$ be the solution to (2.5) with initial condition $\Gamma_{N,0} := \Gamma_{\psi_0}$. As $N \to \infty$, we have*

$$\Gamma_{N,t} \to \Gamma_{\psi_t}$$

in the w^* topology of $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{C})$. In other words, for each k

$$\gamma^{(k)}_{N,t} \to \gamma^{(k)}_{\psi_t}$$

in the weak topology of $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{L}^1(H^{\otimes k}))$.*

Remark 1: Note that our theorem requires the initial one-particle wave function to be in H^2 although the natural space for the nonlinear Schrödinger equation (2.10) is H^1.

Remark 2: Our proof works for more general potentials as well, for example one can consider $V(x) = A(x)|x|^{-1} + B(x)$ with $A, B \in \mathcal{S}(\mathbf{R}^3)$ Schwarz class. The control on high derivatives is necessary because of the commutator estimate (5.13).

We close this section by an important remark. The equations of the BBGKY hierarchy are written for the kernels of the density matrices, however we interpret them as equations for density matrices even though we sometimes use the traditional kernel notation. This point of view allows us to circumvent the problem that apriori the kernels are only functions defined almost everywhere, hence setting $x'_{k+1} = x_{k+1}$ in the interaction term requires an extra argument. In this paper we interpret the integral

$$\int V(x_\ell - x_{k+1})\gamma^{(k+1)}(x_1, \dots x_k, x_{k+1}; x'_1, \dots x'_k;, x_{k+1})\mathrm{d}x_{k+1}$$

as the partial trace

$$\mathrm{Tr}_{x_{k+1}} \widetilde{V}\gamma^{(k+1)}$$

where $\widetilde{V}$ is the multiplication operator by $V(x_\ell - x_{k+1})$. We will always verify that $\widetilde{V}\gamma^{(k+1)}$ is a trace class operator on $L^2(\mathbf{R}^{3(k+1)})$, hence the partial trace is well defined. In the Appendix we give the precise definition of the partial trace and we collect a few basic information.

Convention: Throughout the paper the letter C will refer to various constants depending only on μ.

3 Sobolev spaces of density matrices

We recall that $\mathcal{L}(H)$, $\mathcal{L}^1(H)$ and $\mathcal{K}(H)$ denote the set of bounded, trace class and compact operators on the Hilbert space H, respectively. Also, $\widehat{\mathcal{L}}^1 \subset \mathcal{L}^1(H)$ denote the set of density matrices. We have the standard inequality

$$\|AB\|_1 \leq \|A\| \, \|B\|_1 \tag{3.1}$$

and the variational characterization of the trace norm (2.12). We also recall that $|A+B| \leq |A|+|B|$ is *not* true in general, but $\text{Tr}|A+B| \leq \text{Tr}|A|+\text{Tr}|B|$ is true, following from (2.12), (3.1). Moreover, if we have two Hilbert spaces, H_1, H_2, then the following inequality is valid for partial traces

$$\text{Tr}_1 \Big| \text{Tr}_2 \, A \Big| \leq \text{Tr}_{1,2} \, |A| \, . \tag{3.2}$$

In the Appendix we give the proof of this inequality and we also collect a few basic facts about partial traces.

We define the analogue of the H^1 Sobolev norm for trace class operators as follows:

$$\|T\|_{\mathcal{H}^1} := \text{Tr} \, |STS|$$

where $S := (I - \Delta)^{1/2}$. We introduce the space of operators $\mathcal{H}^1 = \mathcal{H}^1(H)$ as:

$$T \in \mathcal{H}^1 \iff \|T\|_{\mathcal{H}^1} < \infty$$

and let $\widehat{\mathcal{H}}^1 := \mathcal{H}^1 \cap \widehat{\mathcal{L}}^1$. Notice that $\|\gamma\|_{\mathcal{H}^1} = \text{Tr} \, S\gamma S = \text{Tr}\gamma + \text{Tr}(-\Delta)\gamma$ for $\gamma \in \widehat{\mathcal{H}}^1$.

We define a weak* topology on $\mathcal{H}^1$ and for this purpose we identify this space with a dual space. We let

$$\mathcal{A} := \Big\{ SKS \; : \; K \in \mathcal{K} \Big\}$$

and we equip this space with the norm

$$\|T\|_{\mathcal{A}} := \|S^{-1}TS^{-1}\| \, .$$

Lemma 3.1. *With the notations above*

$$(\mathcal{A}, \|\cdot\|_{\mathcal{A}})^* = (\mathcal{H}^1, \|\cdot\|_{\mathcal{H}^1}) .$$

The proof of this Lemma is found in the Appendix and it is similar to the standard proof of $(\mathcal{K}, \|\cdot\|)^* = (\mathcal{L}^1, \|\cdot\|_1)$ outlined in [7]. Using this duality, we equip $\mathcal{H}^1$ with a w^*-topology, induced by the seminorms $T \to |\mathrm{Tr}\; AT|$ indexed by elements $A \in \mathcal{A}$.

Lemma 3.2. *Let $\gamma_n \in \widehat{\mathcal{H}}^1$ be a sequence of uniformly bounded density matrices, i.e.*

$$C := \limsup_n \|\gamma_n\|_{\mathcal{H}^1} = \limsup_n Tr\, S\gamma_n S < \infty .$$

Then one can extract a w^*-convergent subsequence, $\gamma_{n_k} \to \gamma$, $\gamma \in \widehat{\mathcal{H}}^1$, and any w^*-limit point γ of the sequence $\{\gamma_n\}$ satisfies $\|\gamma\|_{\mathcal{H}^1} = Tr\, S\gamma S \leq C$

The convergence in Lemma 3.2 follows from standard application of the Banach-Alaouglu theorem. The positivity of the limit can be checked by testing with the projection operators $P_f := |f\rangle\langle f| \in \mathcal{A}$ for all $f \in H$.

Q.E.D.

Define the following norms for operators $\gamma^{(k)} \in \mathcal{L}(H^{\otimes k})$

$$\|\gamma^{(k)}\|_{\mathcal{H}^{1,(k)}} := \mathrm{Tr}\Big| S_{x_1} \dots S_{x_k} \gamma^{(k)} S_{x_k} \dots S_{x_1} \Big| .$$

This norm is equivalent to the norm

$$\mathrm{Tr}|\gamma^{(k)}| + \mathrm{Tr}\Big|\nabla_{x_1}\nabla_{x_2}\dots\nabla_{x_k}\gamma^{(k)}\nabla_{x_k}\dots\nabla_{x_1}\Big|$$

and for density matrices we have the identity

$$\|\gamma^{(k)}\|_{\mathcal{H}^{1,(k)}} = \mathrm{Tr}(I-\Delta_{x_1})(I-\Delta_{x_2})\dots(I-\Delta_{x_k})\gamma^{(k)} .$$

Here we denote by $\mathcal{H}^{1,(k)} = \mathcal{H}^1(H^{\otimes k})$ the set of operators with finite $\mathcal{H}^{1,(k)}$ norm. Since higher derivatives on the same variable are not allowed, these norms are weaker than the operator analogue of the traditional higher order $W^{k,1}$ Sobolev norms.

The w^*-topology on the space $\mathcal{H}^{1,(k)}$ is given by the seminorms indexed by the set

$$\mathcal{A}^{(k)} := \Big\{ S_{x_1}\dots S_{x_k}\; K\; S_{x_k}\dots S_{x_1} \;:\; K \in \mathcal{K}(H^{\otimes k}) \Big\}$$

with norm

$$\|T\|_{\mathcal{A}^{(k)}} := \left\|(S_{x_1}\dots S_{x_k})^{-1}T(S_{x_1}\dots S_{x_k})^{-1}\right\|$$

and

$$\left(\mathcal{A}^{(k)}, \|\cdot\|_{\mathcal{A}^{(k)}}\right)^* = \left(\mathcal{H}^{1,(k)}, \|\cdot\|_{\mathcal{H}^{1,(k)}}\right)$$

analogously to the one variable case in Lemma 3.1. The analog of Lemma 3.2 also holds.

Define the set

$$\mathcal{D} := \left\{\Gamma = (\gamma^{(1)}, \gamma^{(2)}, \dots) \;:\; \gamma^{(k)} \in \mathcal{H}^1(H^{\otimes k})\right\} = \prod_{k=1}^{\infty} \mathcal{H}^1(H^{\otimes k})\,.$$

This set is equipped with the product topology τ generated by the w^*-topology on each $\mathcal{H}^1(H^{\otimes k})$. The convergence in $\mathcal{D}$ is characterized by the following property: $\Gamma_n \to \Gamma$ in $(\mathcal{D}, \tau)$ as $n \to \infty$, if for each k we have $\gamma_n^{(k)} \to \gamma^{(k)}$ in w^* sense in $\mathcal{H}^{1,(k)}$.

Finally, we will consider time dependent density matrices. The set $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{H}^{1,(k)})$ consists of functions $\gamma^{(k)}(t) : \mathbf{R}_+ \to \mathcal{H}^{1,(k)}$ with $\sup_t \langle t\rangle^{-1}\|\gamma^{(k)}(t)\|_{\mathcal{H}^{1,(k)}} < \infty$. We define the set $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{D}) = \prod_{k=1}^{\infty} L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{H}^{1,(k)})$ and we equip it with the product of the w^*-topologies on each factor. On each factor the w^*-topology is given by seminorms

$$\gamma^{(k)}(t) \to \left|\int_0^\infty \mathrm{Tr}\Big[\, A(t)\gamma^{(k)}(t)\,\Big]\,\mathrm{d}t\right|$$

where $A(t) \in L^1(\mathbf{R}_+, \langle t\rangle\mathrm{d}t, \mathcal{A}^{(k)})$, i.e. $\int_0^\infty \|A(t)\|_{\mathcal{A}^{(k)}}\langle t\rangle\mathrm{d}t < \infty$.

4 Structure of the proof

In this section we give the main propositions and show how our theorem follows from them. The proofs of the propositions are found in the subsequent sections.

4.1 Cutoff of the Coulomb singularity.

We now introduce an N dependent cutoff for the Coulomb potential. Decompose $V = V_1 + V_2$ where

$$V_1(x) = \theta(\sqrt{N}\varepsilon^{-1}x)V(x)$$

with a smooth cutoff function $0 \leq \theta \leq 1$ with $\theta \equiv 1$ on outside of $B(0,2)$ and $\theta \equiv 0$ inside $B(0,1)$. Here $B(x,r)$ denotes the ball of radius r centered at x. Notice the potentials, V_1 and V_2, depend on ε and N which dependence is not explicitly labelled.

We define the cutoff Hamiltonian

$$H_{N,\varepsilon} = -\frac{1}{2}\sum_{\ell=1}^{N} \Delta_{x_\ell} + \frac{1}{N}\sum_{\ell<j} V_1(x_\ell - x_j)$$

and the remaining part of the potential

$$W := W_{N,\varepsilon} = \frac{1}{N}\sum_{\ell<j} V_2(x_\ell - x_j)$$

so that $H_N = H_{N,\varepsilon} + W$.

The density matrices $\gamma_{N,t}^{\varepsilon,(k)}$ of the cutoff dynamics again satisfy the N-body Schrödinger hierarchy (2.5) if V is replaced with V_1, i.e.

$$\begin{aligned}
&\gamma_{N,t}^{\varepsilon,(k)}(x_1,\dots x_k; x_1',\dots x_k') \qquad (4.1)\\
&= \mathcal{U}_{N,\varepsilon,k}(t)\gamma_{N,0}^{(k)}(x_1,\dots x_k; x_1',\dots x_k') + (-i)\sum_{\ell=1}^{k}\int_0^t \mathrm{d}s\, \mathcal{U}_{N,\varepsilon,k}(t-s)\\
&\quad \Big[\int \mathrm{d}x_{k+1}\Big(V_1(x_\ell - x_{k+1}) - V_1(x_\ell' - x_{k+1})\Big)\\
&\quad \times\gamma_{N,s}^{\varepsilon,(k+1)}(x_1,\dots x_k, x_{k+1}; x_1',\dots x_k', x_{k+1})\Big]
\end{aligned}$$

where $\mathcal{U}_{N,\varepsilon,k}(t)\gamma = e^{-itH_{N,\varepsilon}^{(k)}}\gamma e^{itH_{N,\varepsilon}^{(k)}}$ and

$$H_{N,\varepsilon}^{(k)} = -\frac{1}{2}\sum_{\ell=1}^{k} \Delta_\ell + \frac{1}{N}\sum_{\ell<j}^{k} V_1(x_\ell - x_j)\ .$$

We now consider the evolution of the cutoff dynamics with smooth initial data. We introduce the notation

$$L := \frac{1}{N}\sum_{\ell=1}^{N}(I - \Delta_{x_\ell})\ .$$

We are ready to state the main propositions on the BBGKY hierarchy:

Proposition 4.1. *[Apriori bound] Assume that the initial data of the N-body Schrödinger hierarchy (4.1)*

$$\Gamma_{N,0} := \left(\gamma_{N,0}^{(1)}, \gamma_{N,0}^{(2)}, \dots \gamma_{N,0}^{(N)}, 0, 0, \dots\right) \in \mathcal{C} \tag{4.2}$$

forms a consistent sequence of normalized density matrices, i.e. $\gamma_{N,0}^{(k)} = Tr_{k+1}\, \gamma_{N,0}^{(k+1)}$, $k = 1, 2, \dots, N-1$ *and* $Tr\, \gamma_{N,0}^{(N)} = 1$. *Assume that*

$$Tr\, L^k \gamma_{N,0}^{(N)} \le \nu_0^k, \qquad k = 1, 2, \dots N \tag{4.3}$$

for some constant $\nu_0 > 0$ *independent of* N. *Let* $\varepsilon > 0$ *be any fixed constant and*

$$\Gamma_{N,t}^{\varepsilon} := \left(\gamma_{N,t}^{\varepsilon,(1)}, \gamma_{N,t}^{\varepsilon,(2)}, \dots\right) \tag{4.4}$$

be the solution to the N-body Schrödinger hierarchy (4.1) with initial data $\Gamma_{N,0}$ *and Hamiltonian* $H_{N,\varepsilon}$. *Then there exists a* $\nu > 0$ *depending only on* ν_0, *and for any* k *fixed there is a constant* $N(\varepsilon, k)$ *such that* $\Gamma_{N,t}^{\varepsilon}$ *satisfies the a priori estimate*

$$\|\gamma_{N,t}^{\varepsilon,(k)}\|_{\mathcal{H}^{1,(k)}} \le \nu^k \tag{4.5}$$

for $N \ge N(\varepsilon, k)$. *Therefore,* $\Gamma_{N,t}^{\varepsilon}$ *forms a pre-compact sequence in* $L^\infty(\mathbf{R}_+, \mathcal{D})$ *(with the* w^* *topology) by Lemma 3.2.*

For $\nu > 0$ define the norm for $\Gamma = \left(\gamma^{(1)}, \gamma^{(2)}, \dots\right) \in \mathcal{D}$

$$\|\Gamma\|_\nu := \sum_{k=1}^{\infty} \nu^{-k} \|\gamma^{(k)}\|_{\mathcal{H}^{1,(k)}} \,.$$

Proposition 4.2. *[Convergence to the infinite hierarchy] Assume the initial data (4.2) satisfies the estimates* (4.3). *Let* $\varepsilon > 0$ *be any fixed constant. Then any* w^* *limit point* $\Gamma_{\infty,t} \in \mathcal{D}$ *of the sequence* $\{\Gamma_{N,t}^{\varepsilon}\}_{N=1,2,\dots} \subset \mathcal{D}$ *satisfies the infinite BBGKY hierarchy* (2.6) *and for some* ν *large enough* $\|\Gamma_{\infty,t}\|_\nu$ *is uniformly bounded for all* t.

At this stage we do not know that the limit point $\Gamma_{\infty,t} = \left(\gamma_{\infty,t}^{(1)}, \gamma_{\infty,t}^{(2)}, \dots\right)$ forms a consistent family of density matrices. This will follow from the uniqueness of the solution to the nonlinear Schrödinger equation (2.10) and the uniqueness result below.

Proposition 4.3. *[Uniqueness of the infinite BBGKY hierarchy] Let* $T > 0$ *be any fixed time. Then for any* $\nu > 0$ *the infinite BBGKY hierarchy (2.6) has at most one solution in the set* $L^\infty\big([0, T], (\mathcal{D}, \|\cdot\|_\nu)\big)$.

4.2 Approximation by smooth initial data

For any $\kappa > 0$ let $\psi_0^\kappa := e^{\kappa\Delta}\psi_0$. This smoothed initial data satisfies the following bound

$$\|\psi_0 - \psi_0^\kappa\| \le C\kappa \|\psi_0\|_{H^2} \tag{4.6}$$

using a simple estimate in Fourier space

$$\|\psi_0 - \psi_0^\kappa\|^2 = \int (e^{-\kappa p^2} - 1)^2 |\widehat{\psi}_0(p)|^2 \mathrm{d}p \le C\kappa^2 \int p^4 |\widehat{\psi}_0(p)|^2 \mathrm{d}p \, .$$

For $\delta > 0$, let

$$\Psi_{N,0}^\delta(x_1, \ldots x_N) := \prod_{j=1}^N \psi_0^{\delta/N}(x_j)$$

denote the δ/N regularized initial wave function. Let $\gamma_{N,0}^{\delta,(N)}$ be the projection matrix onto $\Psi_{N,0}^\delta$. The main reason to regularize the initial data is the following bound, whose proof is postponed.

Proposition 4.4. *There exists a constant C such that for any $\delta > 0$ and $k \ge 0$,*

$$Tr\Big[L^k \gamma_{N,0}^{\delta,(N)}\Big] = \langle \Psi_{N,0}^\delta, L^k \Psi_{N,0}^\delta \rangle \le \Big(C\|\psi_0\|_{H^2}^2\Big)^k \tag{4.7}$$

whenever N is sufficiently large depending on δ and k.

The proof of this Proposition will be given in Sect 5. From now on we assume that the one-particle wave function of the initial data is $\psi_0 \in H^2(\mathbf{R}^3)$. We can now apply the previous setup of the $\varepsilon/\sqrt{N}$ cutoff dynamics (Section 4.1) to the regularized initial data. Let

$$\Psi_{N,t}^\delta = e^{-itH_N}\Psi_{N,0}^\delta \, , \quad \text{and} \quad \Psi_{N,t}^{\delta,\varepsilon} := e^{-itH_{N,\varepsilon}}\Psi_{N,0}^\delta$$

be the solutions to the Schrödinger equation (2.1) with the original Hamiltonian and with the regularized Hamiltonian, $H_{N,\varepsilon}$, respectively. Denote the corresponding density matrices by $\gamma_{N,t}^{\delta,(k)}$ and $\gamma_{N,t}^{\delta,\varepsilon,(k)}$. From Propositions 4.1-4.4 and from the fact that Γ_{ψ_t} (2.14) solves the infinite BBGKY hierarchy (2.6), we have the following corollary. Notice that the BBGKY hierarchy is independent of the cutoff $\varepsilon/\sqrt{N}$ and the initial data for the BBGKY hierarchy is independent of the smoothing parameter δ/N. Both regulatizations disappear after we have taken the limit $N \to \infty$.

Corollary 4.5. *Suppose that $\psi_0 \in H^2(\mathbf{R}^3)$ and let ψ_t be the solution to (2.10). There exists a constant ν depending on $\|\psi_0\|_{H^2}$ such that for any k and $\delta > 0$*

$$\|\gamma_{N,t}^{\delta,\varepsilon,(k)}\|_{\mathcal{H}^{1,(k)}} \leq \nu^k \tag{4.8}$$

for sufficiently large $N \geq N(\delta, \varepsilon, k)$. Furthermore,

$$\Gamma_{N,t}^{\delta,\varepsilon} \to \Gamma_{\psi_t}$$

as $N \to \infty$, in the w^*-topology of $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{D})$. In other words, for each k

$$\gamma_{N,t}^{\delta,\varepsilon,(k)} \to \gamma_{\psi_t}^{(k)}$$

in the weak topology of $L^\infty(\mathbf{R}_+, \langle t\rangle^{-1}\mathrm{d}t, \mathcal{H}^1(H^{\otimes k}))$ as $N \to \infty$.*

We now control the deviation due to the cutoff and smoothing. The next Lemma shows that the N-body wave functions for the δ/N regularized and the original initial data are close for all time.

Lemma 4.6. *Suppose the initial data $\psi_0 \in H^2(\mathbf{R}^3)$. Then we have*

$$\lim_{\delta\to 0} \limsup_{N\to\infty} \sup_t \|\Psi_{N,t} - \Psi_{N,t}^{\delta}\| = 0 \; . \tag{4.9}$$

Proof. By unitarity, it is sufficient to consider $t = 0$;

$$\|\Psi_{N,0} - \Psi_{N,0}^{\delta}\| \leq \sum_{j=1}^{N} \|\psi_0\|^{j-1} \, \|\psi_0 - \psi_0^{\delta/N}\| \, \left\|\psi_0^{\delta/N}\right\|^{N-j-1} \leq \delta\|\psi_0\|_{H^2}$$

using (4.6) and $\|\psi_0^{\kappa}\| \leq \|\psi_0\| = 1$. **Q.E.D.**

Finally, the following lemma controls the difference between the original dynamics and the $\varepsilon/\sqrt{N}$-regularized dynamics.

Lemma 4.7. *For any ε fixed, we have*

$$\sup_{\delta} \limsup_{N\to\infty} \|\Psi_{N,t}^{\delta,\varepsilon} - \Psi_{N,t}^{\delta}\|^2 \leq C(\psi_0)\varepsilon t \; ,$$

where $C(\psi_0)$ depends on the H^2-norm of ψ_0.

The proof of this Lemma will be given in Section 8. Combining these Lemmas, we have

$$\lim_{\varepsilon\to 0} \lim_{\delta\to 0} \limsup_{N\to\infty} \sup_t \langle t\rangle^{-1} \|\Psi_{N,t}^{\delta,\varepsilon} - \Psi_{N,t}\| = 0 \; .$$

Thus we have for any $k \geq 1$

$$\lim_{\varepsilon\to 0}\lim_{\delta\to 0}\limsup_{N\to\infty}\sup_t \langle t\rangle^{-1}\|\gamma_{N,t}^{\delta,\varepsilon,(k)} - \gamma_{N,t}^{(k)}\|_1 = 0 \, .$$

Since the trace norm is stronger than the w^*-topology in $\mathcal{L}^1(H^{\otimes(k)})$, Theorem 2.1 follows from Corollary 4.5 and Lemmas 4.6 and 4.7. Notice that the limit in Corollary 4.5 was controlled in the w^* Sobolev norm. The control in Lemmas 4.6 and 4.7 was in strong sense but without control on the derivatives. This is why the final result is only in the w^*-topology of trace class operators. **Q.E.D.**

5 Apriori bounds

In this section we first prove Proposition 4.4 then Proposition 4.1. Recall that Proposition 4.4 guarantees that the smoothed initial data satisfies the apriori bound (4.3) for the initial condition in Proposition 4.1.

Proof of Proposition 4.4: Let $\mathbf{n} = (n_1, n_2, \dots n_\ell)$ be a sequence of positive integers, let $k(\mathbf{n}) = n_1 + \dots + n_\ell$ and $\ell(\mathbf{n}) = \ell$. We also define $L_j := \frac{1}{N}(I - \Delta_{x_j})$, clearly $L = \sum_{j=1}^N L_j$. The following identity is easily checked

$$L^k = \sum_{\ell=1}^k \sum_{\mathbf{n}\,:\,k(\mathbf{n})=k\,:\,\ell(\mathbf{n})=\ell} \quad \sum_{i_1,\dots i_\ell=1\,:\,disjoint}^N L_{i_1}^{n_1} L_{i_2}^{n_2} \dots L_{i_\ell}^{n_\ell} \, . \tag{5.1}$$

We have (recall $\kappa = \delta/N$)

$$\begin{aligned}\left\langle \psi_0^\kappa, (I-\Delta)^n \psi_0^\kappa \right\rangle &= \int (1+p^2)^n e^{-\kappa p^2} |\widehat{\psi}_0(p)|^2 \mathrm{d}p \\ &\leq 2^n \|\widehat{\psi}_0\|^2 + 2^n \int p^{2n} e^{-\kappa p^2} |\widehat{\psi}_0(p)|^2 \mathrm{d}p \\ &\leq C^n (1 + n!\, \kappa^{2-n}) \|\psi_0\|_{H^2}^2 \, . \end{aligned} \tag{5.2}$$

Therefore

$$\begin{aligned}
&\sum_{i_1,\dots i_\ell=1\,:\,disjoint}^{N} \left\langle \Psi^\delta_{N,0}, L^{n_1}_{i_1} L^{n_2}_{i_2}\dots L^{n_\ell}_{i_\ell}\Psi^\delta_{N,0}\right\rangle \qquad (5.3)\\
=\;& \frac{N!}{(N-\ell)!N^k}\prod_{j=1}^{\ell}\left\langle \psi_0^\kappa, (I-\Delta)^{n_j}\psi_0^\kappa\right\rangle\\
\le\;& N^{\ell-k}\Big(C\|\psi_0\|^2_{H^2}\Big)^k\prod_{j=1}^{\ell}(1+n_j!\,\kappa^{2-n_j})\\
=\;& \Big(C\|\psi_0\|^2_{H^2}\Big)^k\prod_{j=1}^{\ell}\Big(N^{1-n_j}+n_j!\,N^{-1}\delta^{2-n_j}\Big)\\
\le\;& \Big(C\|\psi_0\|^2_{H^2}\Big)^k\times\begin{cases}1 & \text{if } \ell=k \quad (\text{i.e. } n_1=n_2=\ldots=1)\\ k!\,\langle\delta^{-k}\rangle N^{-1} & \text{if } \ell<k\end{cases}\quad.
\end{aligned}$$

Hence from this estimate, (5.1) and from the exponential bound C^k on number of possible $\mathbf{n}$ sequences with $k(\mathbf{n})=k$ we obtain

$$\left\langle \Psi^\delta_{N,0}, L^k\Psi^\delta_{N,0}\right\rangle \le \Big(C\|\psi_0\|^2_{H^2}\Big)^k\Big(1+k!\,\langle\delta^{-k}\rangle N^{-1}\Big) \qquad (5.4)$$

and (4.7) follows. This completes the proof of Proposition 4.4. **Q.E.D.**

Proof of Proposition 4.1. We define

$$Q_i = \frac{1}{N}(\beta^2-\tfrac{1}{2}\Delta_{x_i})\,,\qquad Q:=\sum_{i=1}^{N}Q_i \qquad (5.5)$$

where the constant $\beta\ge 2$ will be chosen later. We also introduce $U_{ij} := N^{-2}V_1(x_i-x_j)$ (for $i\ne j$), $U_{ii}=0$ and $U:=\sum_{1\le i<j\le N}U_{ij}$. Thus

$$H_{N,\varepsilon}+\beta^2N = N(Q+U)\,.$$

We also define

$$Q^{(k)} := \sum_{i_1,i_2,\dots i_k=1\,:\,disjoint}^{N} Q_{i_1}Q_{i_2}\dots Q_{i_k}\,.$$

Note that $Q, Q^{(k)}$ and U depend on N, but this fact is suppressed in the notation. The key Proposition is the following result to compare Q^k with $H^k_{N,\varepsilon}$:

Proposition 5.1. *There exist a constant β (depending only on μ) such that for any k and any $N \geq N(\varepsilon, k)$ we have the following operator inequalities*

$$C_1^k Q^k \leq (Q+U)^k \leq C_2^k Q^k \tag{5.6}$$

with some positive constants C_1, C_2.

First we show how this Proposition implies Proposition 4.1. The assumptions (4.3) and (5.6) imply that

$$\mathrm{Tr}\ \gamma_{N,0}^{(N)}(H_{N,\varepsilon} + \beta^2 N)^k \leq C_3^k(\nu_0^k + \beta^{2k})N^k$$

for $k = 1, 2, \ldots N$ if N is large enough. Since any function of the Hamiltonian is conserved along the time evolution, we have

$$\mathrm{Tr}\ \gamma_{N,t}^{\varepsilon,(N)}(H_{N,\varepsilon} + \beta^2 N)^k \leq C_3^k(\nu_0^k + \beta^{2k})N^k \ .$$

By (5.6) and $Q^{(k)} \leq Q^k$ we obtain

$$\mathrm{Tr}\ \gamma_{N,t}^{\varepsilon,(N)} Q^{(k)} \leq C_4^k(\nu_0^k + \beta^{2k}) \ .$$

Using the total symmetry of the N-body density matrix and $\beta \geq 1$,

$$\|\gamma_{N,t}^{\varepsilon,(k)}\|_{\mathcal{H}^{1,(k)}} \leq 2^k\ \mathrm{Tr}\ \gamma_{N,t}^{\varepsilon,(N)} Q^{(k)}$$

and (4.5) follows. The pre-compactness of the sequence $\Gamma_{N,t}^{\varepsilon}$ follows from the multivariable analog of Lemma 3.2 and a diagonal selection procedure.

Q.E.D.

Proof of Proposition 5.1: We proceed by a step two induction on k. The statement (5.6) for $k = 0$ is trivial. For $k = 1$ the bounds follow from

$$|U| \leq C\beta^{-1}Q \tag{5.7}$$

which is a consequence of Hardy's inequality in three dimensions (Uncertainty Principle Lemma in Vol.II. Section X.2. of [7])

$$\frac{1}{4|x|^2} \leq -\Delta_x \ , \qquad x \in \mathbf{R}^3 \ . \tag{5.8}$$

Assuming (5.6) is valid for some k, we show it for $k+2$. Writing $(Q + U)^{k+2} = (Q+U)(Q+U)^k(Q+U)$ we have

$$C_1^k(Q+U)Q^k(Q+U) \leq (Q+U)^{k+2} \leq C_2^k(Q+U)Q^k(Q+U) \ .$$

In order to compare UQ^kU with Q^{k+2} we will prove the following lemma:

Lemma 5.2. *For $0 \le k \le N/2$, $0 \le a < 3$*

$$UQ^kU \le \left(C\beta^{-2} + C(k)\Big[\beta^{2k}N^{-k} + N^{1-a/2}\varepsilon^{-2k-2+a} + N^{-1}\Big]\right)Q^{k+2} . \tag{5.9}$$

with constants depending on a and we can set $C(0) = 0$.

Remark: A similar estimate holds for potentials much more singular than the Coulomb potential. With essentially the same proof, one can establish it for $V(x) \sim |x|^{-2+\eta}$ for any $\eta > 0$.

From the Schwarz inequality $A^*B + B^*A \le A^*A + B^*B$, with $A^* = \sqrt{2}UQ^{k/2}$, and $B = -\frac{1}{\sqrt{2}}Q^{k/2+1}$, we have

$$-\Big(UQ^kQ + QQ^kU\Big) \le 2UQ^kU + \frac{1}{2}QQ^kQ .$$

Together with Lemma 5.2 with some $2 < a < 3$ and β large enough, the lower bound in (5.6) follows from

$$(Q+U)Q^k(Q+U) \ge \frac{1}{2}QQ^kQ - UQ^kU \ge \frac{1}{4}Q^{k+2}$$

as $N \ge N(\varepsilon, k)$. So $C_1 = 1/2$.

For the upper bound, we again let $2 < a < 3$. Similar arguments lead to

$$(Q+U)Q^k(Q+U) \le 2Q^{k+2} + 2UQ^kU \le CQ^{k+2}$$

for N large enough, which completes the proof of Lemma 5.1 **Q.E.D.**

Proof of Lemma 5.2. From Schwarz inequality, we obtain

$$UQ^kU \le N^2 \sum_{i,j=1}^{N} |U_{ij}|\, Q^k\, |U_{ij}| . \tag{5.10}$$

For $k = 0$, (5.9) follows from

$$|U_{ij}|^2 \le CN^{-3}Q_j \le C\beta^{-2}N^{-2}Q_iQ_j , \tag{5.11}$$

which is a consequence of Hardy's inequality (5.8).

For $k \geq 1$ and for each fixed i, j we write $Q = Q_i + Q_j + Q_{\widehat{ij}}$. By the weighted Minkowski inequality, we have for some constant C

$$Q^k \leq (Ck)^k (Q_i^k + Q_j^k) + 2Q_{\widehat{ij}}^k \,. \tag{5.12}$$

Let $D_i := \sqrt{-1} N^{-1/2} \nabla_{x_i}$. We have from Leibniz rule and Schwarz inequality that there is a constant $C(k)$ such that

$$|U_{ij}|\; D_i^{2k} \;|U_{ij}| \leq C D_i^k \;|U_{ij}|^2 D_i^k + C(k) \sum_{m=1}^{k} D_i^{k-m} \left| (D_i^m |U_{ij}|) \right|^2 D_i^{k-m} \,. \tag{5.13}$$

To estimate the derivative of U_{ij} we notice that

$$|\nabla_x^m V_1(x-y)|^2 \leq C^m \left(\sqrt{N}/\varepsilon \right)^{2m+2-a} |x-y|^{-a} \tag{5.14}$$

for $a \leq 2m+2$. We will choose $2 < a < 3$. The singularity of $|x-y|^{-a}$ is controlled by the following lemma whose proof is postponed:

Lemma 5.3. *Let $W \in L^1(\mathbf{R}^3)$ be any nonnegative potential, then*

$$W(x-y) \leq C \|W\|_{L^1} (I - \Delta_x)(I - \Delta_y) \tag{5.15}$$

on $L^2(\mathbf{R}^3 \times \mathbf{R}^3)$.

Combining this lemma with (5.14) we obtain ($m \geq 1$)

$$\left| (D_i^m |U_{ij}|) \right|^2 \leq C^m N^{-1-a/2} \varepsilon^{-2m-2+a} Q_i Q_j \,. \tag{5.16}$$

We have $Q_i^k \leq C(k) \left[D_i^{2k} + (\beta^2 N^{-1})^k \right]$. Hence from (5.11), (5.13), and (5.16) we obtain

$$|U_{ij}|\; Q_i^k \;|U_{ij}| \leq C(k) N^{-2} \Big((\beta^2 N^{-1})^k + N^{1-a/2} \varepsilon^{-2k-2+a} + N^{-1} \Big) Q_j (Q_i^k + Q_i) \,.$$

Using $Q = \sum_j Q_j$, $\sum_i Q_i^k \leq Q^k$ and $Q \geq I$,

$$N^2 \sum_{i,j=1}^{N} |U_{ij}|\; Q_i^k \;|U_{ij}| \leq C(k) \Big((\beta^2 N^{-1})^k + N^{1-a/2} \varepsilon^{-2k-2+a} + N^{-1} \Big) Q^{k+1} \,. \tag{5.17}$$

Finally, we use that U_{ij} and $Q_{\widehat{ij}}$ commute and the estimate (5.11) to bound

$$N^2 \sum_{i,j=1}^{N} |U_{ij}|\, Q_{\widehat{ij}}^k\, |U_{ij}| \le N^2 \sum_{i,j=1}^{N} C\beta^{-2}N^{-2}Q_iQ_jQ_{\widehat{ij}}^k \le C\beta^{-2}Q^{k+2}\,. \quad (5.18)$$

Hence from (5.12), (5.17), (5.18) and $Q \ge I$ the statement (5.9) follows. This finished the proof of Lemma 5.2 **Q.E.D.**

Proof of Lemma 5.3. Let

$$H(x) := \int_{\mathbf{R}^3} \frac{e^{ipx}}{\sqrt{1+p^2}}\, \mathrm{d}p\,, \quad (5.19)$$

then $H(x) \sim |x|^{-2}$ around the origin, smooth outside and it decays faster than any polynomial at infinity. Clearly (5.15) is equivalent to

$$\frac{1}{\sqrt{(I-\Delta_x)(I-\Delta_y)}} W(x-y) \frac{1}{\sqrt{(I-\Delta_x)(I-\Delta_y)}} \le C\|W\|_{L^1}\,. \quad (5.20)$$

Let $f \in L^2(\mathbf{R}^3 \times \mathbf{R}^3)$, then

$$\left\langle f, \frac{1}{\sqrt{(I-\Delta_x)(I-\Delta_y)}} W(x-y) \frac{1}{\sqrt{(I-\Delta_x)(I-\Delta_y)}} f \right\rangle \quad (5.21)$$

$$= \int \mathrm{d}x\mathrm{d}x'\mathrm{d}x''\mathrm{d}y\mathrm{d}y'\mathrm{d}y'' \overline{f(x,y)} H(x-x')H(y-y')W(x'-y')$$
$$\times H(x'-x'')H(y'-y'')f(x'',y'')$$
$$\le 2\int \mathrm{d}x\mathrm{d}x'\mathrm{d}x''\mathrm{d}y\mathrm{d}y'\mathrm{d}y'' |f(x,y)|^2 \left[\frac{|x-x'|}{|x'-x''|}\frac{|y-y'|}{|y'-y''|}\right]^{3/4}$$
$$\times \Big| H(x-x')H(y-y')W(x'-y')H(x'-x'')H(y'-y'') \Big|$$

by symmetry of (x,y) and (x'',y'') and a Schwarz inequality. By the properties of H

$$\int \frac{|H(u)|}{|u|^{3/4}} \mathrm{d}u \le C$$

hence x'', y'' integrals can be performed. We also define $G(u) := |u|^{3/4}|H(u)|$ and notice that $\|G\|_{L^2} \le C$. Finally, by Young's inequality:

$$\sup_{x,y} \int G(x-x')|W(x'-y')|G(y'-y)\mathrm{d}x'\mathrm{d}y' \le \|G\|_{L^2}^2\|W\|_{L^1}$$

we obtain (5.20). **Q.E.D.**

6 Proof of the convergence

In this section we prove Proposition 4.2. The uniform boundedness of $\|\Gamma_{\infty,t}\|_\nu$ follows from (4.5) and Lemma 3.2. To check that the limit satisfies (2.6), it is sufficient test the equation against trace class operators $\mathcal{O} \in \mathcal{L}^1(H^{\otimes k})$ on a fixed time interval $[0, T]$. We note that $\Gamma^\varepsilon_{N,t}$ satisfies the finite Schrödinger hierarchy with the cutoff potential (4.1).

Since $\mathcal{O}$ is compact, we have

$$\lim_{N\to\infty} \text{Tr}\ \mathcal{O}\Big(\gamma^{\varepsilon,(k)}_{N,t} - \gamma^{(k)}_{\infty,t}\Big) = 0\ .$$

This shows the convergence of the left side of (4.1) to that of (2.6).

For the convergence of the first term on the right side of (4.1), we now show that $\mathcal{U}_{N,k,\varepsilon}(t)$ converges to $\mathcal{U}_k(t)$. By one-step Duhamel expansion we have

$$e^{isH^{(k)}_{N,\varepsilon}} - \Big(\prod_{j=1}^k e^{is\Delta_j}\Big) = \frac{-i}{N}\sum_{\ell<j}^k \int_0^s e^{i(s-u)H^{(k)}_{N,\varepsilon}} V_1(x_\ell - x_j)\Big(\prod_{j=1}^k e^{iu\Delta_j}\Big)\ \text{d}u\ . \tag{6.1}$$

From the triangle inequality, the unitarity of $e^{i(s-u)H^{(k)}_{N,\varepsilon}}$ and $|V_1| \le C\varepsilon^{-1}N^{1/2}$, we can bound the operator norm of the left side by

$$\begin{aligned}
&\Big\| e^{isH^{(k)}_{N,\varepsilon}} - \Big(\prod_{j=1}^k e^{is\Delta_j}\Big)\Big\| \\
&\qquad \le \frac{1}{N}\sum_{\ell<j}^k \int_0^s \Big\| e^{i(s-u)H^{(k)}_{N,\varepsilon}} V_1(x_\ell - x_j)\Big(\prod_{j=1}^k e^{iu\Delta_j}\Big)\Big\|\ \text{d}u \\
&\qquad \le \frac{Ck^2 s}{\varepsilon\sqrt{N}}\ .
\end{aligned} \tag{6.2}$$

Using (6.2) the convergence of $\gamma^{(k)}_{N,0}$, the fact that $\text{Tr}\gamma^{(k)}_{\infty,0} \le \text{Tr}\gamma^{(k)}_{N,0} = 1$ and $\|\mathcal{O}\| < \infty$, we thus have

$$\lim_{N\to\infty} \Big|\ \text{Tr}\ \mathcal{O}\Big[\mathcal{U}_{N,k,\varepsilon}(t)\gamma^{(k)}_{N,0} - \mathcal{U}_k(t)\gamma^{(k)}_{\infty,0}\Big]\ \Big| = 0\ .$$

We now show the convergence of the second term on the right hand side of (4.1), i.e.,

$$\lim_{N\to\infty}\int_0^t \text{Tr}\ \mathcal{O}\Bigg(\mathcal{U}_{N,k,\varepsilon}(t-s)\Bigg[\int \text{d}x_{k+1}\Big(V_1(x_\ell - x_{k+1}) - V_1(x'_\ell - x_{k+1})\Big)\gamma^{\varepsilon,(k+1)}_{N,s}\Bigg]$$

$$-\mathcal{U}_k(t-s)\left[\int \mathrm{d}x_{k+1}\Big(V(x_\ell - x_{k+1}) - V(x'_\ell - x_{k+1})\Big)\gamma^{(k+1)}_{\infty,s}\right]\Bigg)\mathrm{d}s = 0$$

for any $\ell \leq k$. Decompose this difference into three terms $(I.)+(II.)+(III.)$ where

$$\begin{aligned}(I.) &= \int_0^t \mathrm{Tr}\,\Big(\mathcal{O}\mathcal{U}_{N,k,\varepsilon}(t-s) - \mathcal{O}\mathcal{U}_k(t-s)\Big) \\ &\quad\times\left[\int \mathrm{d}x_{k+1}\Big(V_1(x_\ell - x_{k+1}) - V_1(x'_\ell - x_{k+1})\Big)\gamma^{\varepsilon,(k+1)}_{N,s}\right]\mathrm{d}s\,,\end{aligned} \tag{6.3}$$

$$\begin{aligned}(II.) &= \int_0^t \mathrm{Tr}\,\mathcal{O}\mathcal{U}_k(t-s)\left[\int \mathrm{d}x_{k+1}\Big(V(x_\ell - x_{k+1}) - V(x'_\ell - x_{k+1})\Big)\right. \\ &\quad\left.\times\Big(\gamma^{\varepsilon,(k+1)}_{N,s} - \gamma^{(k+1)}_{\infty,s}\Big)\right]\mathrm{d}s\,,\end{aligned} \tag{6.4}$$

$$\begin{aligned}(III.) &= -\int_0^t \mathrm{Tr}\,\mathcal{O}\mathcal{U}_k(t-s) \\ &\quad\times\left[\int \mathrm{d}x_{k+1}\Big(V_2(x_\ell - x_{k+1}) - V_2(x'_\ell - x_{k+1})\Big)\gamma^{\varepsilon,(k+1)}_{N,s}\right]\mathrm{d}s\,.\end{aligned} \tag{6.5}$$

We estimate these three terms separately.

The first term is bounded by

$$\begin{aligned}|(I.)| &\leq t\sup_{s\leq t}\Big\|\mathcal{O}\mathcal{U}_{N,k,\varepsilon}(t-s) - \mathcal{O}\mathcal{U}_k(t-s)\Big\| \\ &\quad\times\sup_{s\leq t}\mathrm{Tr}\left|\int \mathrm{d}x_{k+1}\Big(V_1(x_\ell - x_{k+1}) - V_1(x'_\ell - x_{k+1})\Big)\gamma^{\varepsilon,(k+1)}_{N,s}\right|\,.\end{aligned}$$

The norm $\Big\|\mathcal{O}\mathcal{U}_{N,k,\varepsilon}(t-s) - \mathcal{O}\mathcal{U}_k(t-s)\Big\|$ is bounded by

$$\begin{aligned}\Big\|\mathcal{O}\mathcal{U}_{N,k,\varepsilon}(t-s) - \mathcal{O}\mathcal{U}_k(t-s)\Big\| &\leq \|\mathcal{O}\|\sup_{s\leq t}\left\|e^{isH^{(k)}_{N,\varepsilon}} - \Big(\prod_{j=1}^k e^{is\Delta_j}\Big)\right\| \\ &\leq \frac{Ck^2\langle t\rangle^2}{\varepsilon\sqrt{N}}\|\mathcal{O}\|\end{aligned}$$

where we have used (6.2) in the last inequality. To bound the trace term, from the triangle inequality we reduce it to estimate

$$\mathrm{Tr}\left|\int \mathrm{d}x_{k+1}V_1(x_\ell - x_{k+1})\gamma^{\varepsilon,(k+1)}_{N,s}\right|$$

for $s \leq t$. Recall that we interpret the x_{k+1} integration as a partial trace over this variable. We can now use (3.2) and bound this term by

$$\mathrm{Tr}\left|V_1(x_\ell - x_{k+1})\gamma_{N,s}^{\varepsilon,(k+1)}\right|$$

where the trace is now over all $k+1$ variables. We define

$$S^{(k)} := S_{x_1}S_{x_2}\dots S_{x_k} .$$

From the inequality (5.8) we know that

$$|V_1(x_\ell - x_{k+1})|^2 \leq |V(x_\ell - x_{k+1})|^2 \leq CS_{x_{k+1}}^2 \leq C[S^{(k+1)}]^2 . \tag{6.6}$$

Hence for any self-adjoint $\gamma \in \mathcal{H}^{1,(k+1)}$ we have

$$\mathrm{Tr}\Big|V_1(x_\ell - x_{k+1})\gamma\Big| = \mathrm{Tr}\sqrt{\gamma V_1^2\gamma} \leq C\mathrm{Tr}\sqrt{\gamma[S^{(k+1)}]^2\gamma} = C\mathrm{Tr}\sqrt{S^{(k+1)}\gamma^2 S^{(k+1)}} \tag{6.7}$$

using the cyclicity for the trace of the square root (9.1). Since the square root is monotonic in operator sense and $S^{(k+1)} \geq I$, we obtain

$$C\mathrm{Tr}\sqrt{S^{(k+1)}\gamma^2 S^{(k+1)}} \leq C\mathrm{Tr}\sqrt{S^{(k+1)}\gamma[S^{(k+1)}]^2\gamma S^{(k+1)}} = C\|\gamma\|_{\mathcal{H}^{1,(k+1)}} , \tag{6.8}$$

where the last equality is the definition of the H^1 norm. Therefore we showed that

$$\mathrm{Tr}\left|\int \mathrm{d}x_{k+1}V_1(x_\ell - x_{k+1})\gamma_{N,s}^{\varepsilon,(k+1)}\right| \leq C\|\gamma_{N,s}^{\varepsilon,(k+1)}\|_{\mathcal{H}^{1,(k+1)}}$$

which is uniformly bounded by (4.5), hence the first term (I.) vanishes as $N \to \infty$.

The first half of the second term $(II.)$ is written as

$$\mathrm{Tr}\int_0^t \mathcal{OU}_k(t-s)V(x_\ell - x_{k+1})\Big(\gamma_{N,s}^{\varepsilon,(k+1)} - \gamma_{\infty,s}^{(k+1)}\Big)\mathrm{d}s ,$$

where the trace refers to all $k+1$ variables. In order to check that $(II.)$ vanishes in the limit $N \to \infty$, it suffices to show that $A(s) := [S^{(k+1)}]^{-1}\mathcal{OU}_k(t-s)V_{\ell,k+1}[S^{(k+1)}]^{-1}$ is in the space $L^\infty([0,t],\mathcal{K}^{(k+1)})$, i.e. that it is compact and is uniformly bounded for all $0 \leq s \leq t$. Here $V_{\ell,k+1}$ is the multiplication operator by $V(x_\ell - x_{k+1})$.

For uniform boundedness, we only have to show that $\mathcal{U}_k(t-s)V_{\ell,k+1}S_{k+1}^{-1}$ is uniformly bounded, since $\mathcal{O}$ is bounded. Here $S_{k+1} := S_{x_{k+1}}$ for brevity. From the unitarity of $\mathcal{U}_k(t-s)$, it suffices to prove that

$$\left\|V_{\ell,k+1}S_{k+1}^{-1}\right\| \leq C\,. \tag{6.9}$$

This follows from the inequality (6.6).

For the compactness, it is sufficient to show that $S_{k+1}^{-1}\mathcal{O}\mathcal{U}_k(t-s)V_{\ell,k+1}S_{k+1}^{-1}$ is actually Hilbert-Schmidt for all s. Recall that $\mathcal{O}$ is a compact operator on $H^{\otimes k}$, i.e., in the space of the first k variables, and the unitary map $\mathcal{U}_k(t-s)$ acts trivially on the $(k+1)$-th variable. Hence we can define $\mathcal{O}_s := \mathcal{O}\mathcal{U}_k(t-s)$, which is a Hilbert-Schmidt operator on $H^{\otimes k}$. To emphasize that $\mathcal{O}\mathcal{U}_k(t-s)$ is viewed as an operator on $H^{\otimes(k+1)}$, we will write $\mathcal{O}_s \otimes I_{k+1}$. The identity I_{k+1} acts on the $(k+1)$-th variable.

We introduce $X := (x_1, \dots x_k)$ and $Y = (y_1, \dots y_k)$, and we compute the Hilbert-Schmidt norm of $S_{k+1}^{-1}\mathcal{O}\mathcal{U}_k(t-s)V_{\ell,k+1}S_{k+1}^{-1}$:

$$\begin{aligned}
&\mathrm{Tr}(\mathcal{O}_s \otimes I_{k+1})S_{k+1}^{-1}V_{\ell,k+1}S_{k+1}^{-2}V_{\ell,k+1}S_{k+1}^{-1}(\mathcal{O}_s \otimes I_{k+1}) \\
&\quad= \mathrm{Tr}(\mathcal{O}_s \otimes I_{k+1})V_{\ell,k+1}S_{k+1}^{-2}V_{\ell,k+1}S_{k+1}^{-2}(\mathcal{O}_s \otimes I_{k+1}) \\
&\quad= \int \mathcal{O}_s(X;Y)V(y_\ell - x_{k+1})\frac{1}{I-\Delta}(x_{k+1}-y_{k+1})V(y_{k+1}-y_\ell) \\
&\qquad\quad \times \frac{1}{I-\Delta}(y_{k+1}-x_{k+1})\mathcal{O}_s(Y;X)\mathrm{d}X\mathrm{d}Y\mathrm{d}x_{k+1}\mathrm{d}y_{k+1} \\
&\quad\leq C\int |\mathcal{O}_s(X;Y)|^2|V(y_\ell - x_{k+1})|^2 \\
&\qquad\quad \times \left|\frac{1}{I-\Delta}(y_{k+1}-x_{k+1})\right|^2 \mathrm{d}X\mathrm{d}Y\mathrm{d}x_{k+1}\mathrm{d}y_{k+1} \\
&\quad\leq C\int |\mathcal{O}_s(X;Y)|^2\mathrm{d}X\mathrm{d}Y < \infty\,,
\end{aligned} \tag{6.10}$$

since $\mathcal{O}_s$ is Hilbert-Schmidt. We used a Schwarz inequality in the third line and the fact that

$$\int \left|\frac{1}{I-\Delta}(z)\right|^2 \mathrm{d}z = \int \frac{\mathrm{d}p}{(1+p^2)^2} < \infty\,.$$

where $(I-\Delta)^{-1}(x-y)$ denotes the translation invariant kernel of the operator $(I-\Delta)^{-1}$. This completes the estimate of the second term $(II.)$.

Finally the third term is estimated as

$$(III.) \leq 2t\|\mathcal{O}\|\sup_{s\leq t}\,\mathrm{Tr}\left|\int \mathrm{d}x_{k+1}V_2(x_1-x_{k+1})\gamma_{N,s}^{\varepsilon,(k+1)}\right|\,.$$

From (3.2) and Lemma 5.3 we can bound the trace term by

$$\mathrm{Tr}\Big|V_2\gamma\Big| = \mathrm{Tr}\sqrt{\gamma V_2^2\gamma} \le C\varepsilon^{1/2}N^{-1/4}\mathrm{Tr}\sqrt{\gamma S_1^2 S_{k+1}^2\gamma}$$

where we introduced $\gamma := \gamma_{N,s}^{\varepsilon,(k+1)}$ for brevity and we computed $\|V_2^2\|_{L^1} \le C\varepsilon N^{-1/2}$. Clearly

$$\mathrm{Tr}\sqrt{\gamma S_1^2 S_{k+1}^2\gamma} \le \mathrm{Tr}\sqrt{\gamma[S^{(k+1)}]^2\gamma} \le C\|\gamma\|_{\mathcal{H}^{1,(k+1)}}$$

as in (6.7)-(6.8), which is uniformly bounded in N by (4.5). Therefore $(III.)$ goes to zero as $N \to \infty$, and we have proved Proposition 4.2. **Q.E.D.**

7 Uniqueness of the infinite hierarchy

Proof of Lemma 4.3. Since the equation is linear, it is sufficient to show that the solution Γ_t up to time T is identically zero if the initial condition is $\Gamma_0 = 0$ and if for some K the apriori bound $\|\Gamma_t\|_\nu \le K$ holds for $t \le T$. Notice that it is sufficient to show that uniqueness holds for a short time, $t \le T(\nu)$, then one can extend it up to time T since the apriori bound holds uniformly for $t \le T$.

We need the following lemma whose proof is postponed.

Lemma 7.1. *For any $\gamma \in \mathcal{H}^{1,(k+1)}$, $k \ge 1$ and $\ell \le k$*

$$Tr_X \left|S^{(k)} \int \mathrm{d}y \; V(x_\ell - y)\gamma(X, y; X', y)S^{(k)}\right| \le C\|\gamma\|_{\mathcal{H}^{1,(k+1)}} \,.$$

where $X = (x_1, \dots x_k)$ for brevity and recall $S^{(k)} = S_{x_1}S_{x_2}\dots S_{x_k}$.

Armed with this estimate, we obtain from (2.6) and $\gamma_0^{(k)} = 0$ that

$$\begin{aligned} \|\gamma_t^{(k)}\|_{\mathcal{H}^{1,(k)}} &= \mathrm{Tr}\left|S^{(k)}\gamma_t^{(k)}S^{(k)}\right| && (7.1)\\ &\le 2\sum_{\ell=1}^k \int_0^t \mathrm{d}s \; \mathrm{Tr}\left|S^{(k)}\mathrm{Tr}_{x_{k+1}}\big[V(x_\ell - x_{k+1})\gamma_s^{(k+1)}\big]S^{(k)}\right| \\ &\le Ck\int_0^t \mathrm{d}s\|\gamma_s^{(k+1)}\|_{\mathcal{H}^{1,(k+1)}} \,. \end{aligned}$$

Let $A(k,t) := \|\gamma_t^{(k)}\|_{\mathcal{H}^{1,(k)}}$. After iteration, we have

$$A(k,t) \leq C^n k(k+1)\dots(k+n-1)\int_0^t \mathrm{d}s_1 \int_0^{s_1} \mathrm{d}s_2 \dots \int_0^{s_{n-1}} \mathrm{d}s_n\, A(n+k, s_n)\,. \tag{7.2}$$

Since Γ_t is bounded in $(\mathcal{D}, \|\cdot\|_\nu)$, there exists some K such that

$$A(n,s) \leq K\nu^n$$

for all $0 \leq s \leq T$. After using this estimate in (7.2), the multiple time integration is bounded by $t^n/n!$. This $n!$ compensates the combinatorial prefactor $k(k+1)\dots(k+n-1)$ and the result is

$$A(k,t) \leq K(2\nu)^{n+k}(Ct)^n \leq K(2\nu)^k(C\nu t)^n$$

for any n. If $t < T(\nu) := (C\nu)^{-1}$, then letting $n \to \infty$ we obtain that $A(k,t) = 0$. **Q.E.D.**

Proof of Lemma 7.1. Without loss of generality we can assume that $\ell = 1$ and we set $x = x_1$, i.e. $X = (x, x_2, \dots x_k)$. We define $\widehat{S} := S_{x_2}S_{x_3}\dots S_{x_k}$ and let $\widehat{\gamma} := \widehat{S}\gamma\widehat{S}$. Let $\widetilde{V}$ be the multiplication operator by the function $(x,y) \mapsto V(x-y)$ on $L^2(\mathbf{R}^{3k}_X \times \mathbf{R}^3_y)$ and let $P_x = (P_{x,1}, P_{x,2}, P_{x,3})$ be the components of momentum operator in the x variable, i.e., $P_{x,j} := -i\nabla_j$. P_y is similarly defined.

We denote $R := \mathrm{Tr}_y(\widetilde{V}\widehat{\gamma})$ the partial trace of $\widetilde{V}\widehat{\gamma}$, then we have to estimate

$$\mathrm{Tr}_X \left| S^{(k)} \int \mathrm{d}y\, V(x_1 - y)\gamma(X, y; X', y) S^{(k)} \right| = \mathrm{Tr}_X \left| S_x R S_x \right|.$$

Using the identity $S_x^2 = I + \sum_{j=1}^3 P_{x,j}^2$ and Lemma 9.2 from the Appendix, we estimate

$$\begin{aligned}
\mathrm{Tr}_X \left| S_x R S_x \right| &= \mathrm{Tr}_X \Big[S_x R S_x^2 R S_x \Big]^{1/2} &(7.3)\\
&\leq C\mathrm{Tr}_X \Big[S_x R^2 S_x \Big]^{1/2} + C\sum_{j=1}^3 \mathrm{Tr}_X \Big[S_x R P_{x,j}^2 R S_x \Big]^{1/2}\\
&= C\mathrm{Tr}_X |R S_x| + C\sum_{j=1}^3 \mathrm{Tr}_X \left| P_{x,j} R S_x \right|\\
&= C\mathrm{Tr}_X \left| \mathrm{Tr}_y(\widetilde{V}\widehat{\gamma}) S_x \right| + C\sum_{j=1}^3 \mathrm{Tr}_X \left| \mathrm{Tr}_y P_{x,j} \widetilde{V}\widehat{\gamma} S_x \right|.
\end{aligned}$$

Hardy's inequality (5.8) implies that

$$\widetilde{V}^2 \le S_x^2 , \qquad \widetilde{V}^2 \le S_y^2 . \tag{7.4}$$

In the first term of (7.3) we use (3.2), (7.4) and the fact that S_x and S_y commute to obtain

$$\begin{aligned}
\mathrm{Tr}_X\Big|\mathrm{Tr}_y(\widetilde{V}\widehat{\gamma})S_x\Big| &\le \mathrm{Tr}\Big|\widetilde{V}\widehat{\gamma}S_x\Big| = \mathrm{Tr}\Big(S_x\widehat{\gamma}\widetilde{V}^2\widehat{\gamma}S_x\Big)^{1/2}\\
&\le \mathrm{Tr}\Big(S_x\widehat{\gamma}S_x^2\widehat{\gamma}S_x\Big)^{1/2}\\
&\le \mathrm{Tr}\Big(S_x\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_x\Big)^{1/2} = \mathrm{Tr}\Big(S_xS_y\widehat{\gamma}S_x^2\widehat{\gamma}S_xS_y\Big)^{1/2}\\
&\le \mathrm{Tr}\Big(S_xS_y\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_xS_y\Big)^{1/2} = \mathrm{Tr}\big|S_xS_y\widehat{\gamma}S_yS_x\big|\\
&= \|\gamma\|_{\mathcal{H}^{1,(k+1)}}
\end{aligned}$$

where Tr denotes trace in all variables X, y.

For the second term in (7.3) we first notice the following identity

$$P_{x,j}\widetilde{V} = -P_{y,j}\widetilde{V} + \widetilde{V}P_{x,j} + \widetilde{V}P_{y,j}$$

Using $\mathrm{Tr}|A+B| \le \mathrm{Tr}|A| + \mathrm{Tr}|B|$, we obtain for each $j = 1, 2, 3$

$$\begin{aligned}
\mathrm{Tr}_X\Big|\mathrm{Tr}_y P_{x,j}\widetilde{V}\widehat{\gamma}S_x\Big| &\le \mathrm{Tr}_X\Big|\mathrm{Tr}_y P_{y,j}\widetilde{V}\widehat{\gamma}S_x\Big| \\
&\quad + \mathrm{Tr}_X\Big|\mathrm{Tr}_y\widetilde{V}P_{x,j}\widehat{\gamma}S_x\Big| + \mathrm{Tr}_X\Big|\mathrm{Tr}_y\widetilde{V}P_{y,j}\widehat{\gamma}S_x\Big|\\
&\le \mathrm{Tr}\Big|\widetilde{V}\widehat{\gamma}S_xP_{y,j}\Big| + \mathrm{Tr}\Big|\widetilde{V}P_{x,j}\widehat{\gamma}S_x\Big| + \mathrm{Tr}\Big|\widetilde{V}P_{y,j}\widehat{\gamma}S_x\Big| .
\end{aligned} \tag{7.5}$$

Notice that in the first term we used the cyclicity of the partial trace (9.4) from the Appendix before applying (3.2).

The estimate of the three terms in (7.5) are similar by using the inequalities (7.4) appropriately. The estimate of the first term is

$$\begin{aligned}
\mathrm{Tr}\Big|\widetilde{V}\widehat{\gamma}S_xP_{y,j}\Big| &= \mathrm{Tr}\Big(P_{y,j}S_x\widehat{\gamma}\widetilde{V}^2\widehat{\gamma}S_xP_{y,j}\Big)^{1/2}\\
&\le \mathrm{Tr}\Big(P_{y,j}S_x\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_xP_{y,j}\Big)^{1/2}\\
&= \mathrm{Tr}\Big(S_xS_y\widehat{\gamma}S_x^2P_{y,j}^2\widehat{\gamma}S_xS_y\Big)^{1/2}\\
&\le \mathrm{Tr}\Big(S_xS_y\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_xS_y\Big)^{1/2}\\
&= \|\gamma\|_{\mathcal{H}^{1,(k+1)}} ,
\end{aligned} \tag{7.6}$$

where we used the cyclicity (9.1) several times.

The second term in (7.5) is estimated as

$$\begin{aligned}
\mathrm{Tr}\left|\widetilde{V}P_{x,j}\widehat{\gamma}S_x\right| &= \mathrm{Tr}\left(S_x\widehat{\gamma}P_{x,j}\widetilde{V}^2P_{x,j}\widehat{\gamma}S_x\right)^{1/2} \\
&\le \mathrm{Tr}\left(S_x\widehat{\gamma}P_{x,j}S_y^2P_{x,j}\widehat{\gamma}S_x\right)^{1/2} \\
&\le \mathrm{Tr}\left(S_x\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_x\right)^{1/2} = \mathrm{Tr}\left(S_xS_y\widehat{\gamma}S_x^2\widehat{\gamma}S_xS_y\right)^{1/2} \\
&\le \mathrm{Tr}\left(S_xS_y\widehat{\gamma}S_x^2S_y^2\widehat{\gamma}S_xS_y\right)^{1/2} \\
&= \|\gamma\|_{\mathcal{H}^{1,(k+1)}} \, .
\end{aligned}$$

The estimate of the third term in (7.5) is identical just the second inequality (7.4) is used. This completes the proof of Lemma 7.1. **Q.E.D.**

8 Removing the Coulomb cutoffs

Proof of Lemma 4.7: We fix N, δ, ε and for simplicity let $\Psi_t := \Psi^{\delta}_{N,t}$, $\widetilde{\Psi}_t := \Psi^{\delta,\varepsilon}_{N,t}$. Clearly $\Psi_0 = \widetilde{\Psi}_0$. We compute

$$\begin{aligned}
\partial_t\|\Psi_t - \widetilde{\Psi}_t\|^2 &= 2\mathrm{Re}\,\langle\Psi_t - \widetilde{\Psi}_t, W\widetilde{\Psi}_t\rangle \\
&\le 2\|\Psi_t - \widetilde{\Psi}_t\|\;\|W\widetilde{\Psi}_t\| \, .
\end{aligned}$$

We need to estimate

$$\begin{aligned}
&\|W\widetilde{\Psi}_t\|^2 \qquad (8.1)\\
= \;& \frac{1}{N^2}\frac{N(N-1)}{2}\int |V_2(x_1-x_2)|^2|\widetilde{\Psi}(x_1,\dots x_N)|^2 \mathrm{d}x_1\dots\mathrm{d}x_N\\
&+\frac{1}{N^2}\frac{N(N-1)(N-2)}{2}\\
&\times\int |V_2(x_1-x_2)||V_2(x_2-x_3)||\widetilde{\Psi}(x_1,\dots x_N)|^2\mathrm{d}x_1\dots\mathrm{d}x_N\\
&+\frac{1}{N^2}\frac{N(N-1)}{2}\frac{(N-2)(N-3)}{2}\\
&\times\int |V_2(x_1-x_3)||V_2(x_2-x_4)||\widetilde{\Psi}(x_1,\dots x_N)|^2\mathrm{d}x_1\dots\mathrm{d}x_N\\
\le\;& \int |V_2(x_1-x_2)|^2|\widetilde{\Psi}(x_1,\dots x_N)|^2\mathrm{d}x_1\dots\mathrm{d}x_N\\
&+CN\int |V_2(x_2-x_3)|\Big(|\nabla_{x_1}\widetilde{\Psi}(x_1,\dots x_N)|^2+|\widetilde{\Psi}(x_1,\dots x_N)|^2\Big)\mathrm{d}x_1\dots\mathrm{d}x_N\\
&+N^2\int |V_2(x_1-x_3)||V_2(x_2-x_4)||\widetilde{\Psi}(x_1,\dots x_N)|^2\mathrm{d}x_1\dots\mathrm{d}x_N\,,
\end{aligned}$$

where we used that $|V_2(x_1-x_2)|\le|x_1-x_2|^{-1}\le C(I-\Delta_{x_1})$ for any x_2 fixed in the second term.

It follows from Lemma 5.3 that

$$\frac{\chi(|x-y|\le\lambda)}{|x-y|^\kappa}\le C(\kappa)\lambda^{3-\kappa}(I-\Delta_x)(I-\Delta_y) \qquad (8.2)$$

as operators on $L^2(\mathbf{R}^3\times\mathbf{R}^3)$ for $\lambda<1$, $0\le\kappa<3$. We use this estimate with $\kappa=1,2$, and $\lambda=\varepsilon/\sqrt{N}$ to continue the estimate as

$$\begin{aligned}
\|W\widetilde{\Psi}_t\|^2 \;\le\;& C\frac{\varepsilon}{\sqrt{N}}\Big(\|\nabla_{x_1}\nabla_{x_2}\widetilde{\Psi}\|^2+\|\widetilde{\Psi}\|^2\Big)\\
&+C\varepsilon^2\Big(\|\nabla_{x_1}\nabla_{x_2}\nabla_{x_3}\widetilde{\Psi}\|^2+\|\nabla_{x_1}\widetilde{\Psi}\|^2+\|\nabla_{x_2}\nabla_{x_3}\widetilde{\Psi}\|^2+\|\widetilde{\Psi}\|^2\Big)\\
&+C\varepsilon^4\Big(\|\nabla_{x_1}\nabla_{x_2}\nabla_{x_3}\nabla_{x_4}\widetilde{\Psi}\|^2+\|\nabla_{x_1}\nabla_{x_2}\widetilde{\Psi}\|^2+\|\widetilde{\Psi}\|^2\Big)\\
\le\;& C\varepsilon^2\|\gamma_{N,t}^{\delta,\varepsilon,(4)}\|_{\mathcal{H}^{1,(4)}}
\end{aligned}$$

for large enough N. Using (4.8) this completes the proof of Lemma 4.7.

Q.E.D.

9 Appendix on partial traces and density matrices

In this Appendix we collect a few facts about trace class operators. We start with the analogue of the monotone and dominated convergence theorems.

Lemma 9.1. *Let $0 \le A_n \le A$ be self-adjoint operators defined on a common separable Hilbert space.*

(i) If the sequence is monotone, $A_n \le A_{n+1}$, $A_n \to A$ strongly, A is compact and $\sup_n Tr\, A_n < \infty$, then $Tr\, A = \sup_n Tr\, A_n$, in particular A is trace class.

(ii) If $A_n \to 0$ strongly and $Tr\, A < \infty$, then $Tr\, A_n \to 0$.

Proof: Let $\{f_j\}$ be an eigenbasis for A. For part (i), we have for any M that

$$\sum_{i=1}^{M} \langle f_i, A f_i \rangle = \lim_{n\to\infty} \sum_{i=1}^{M} \langle f_i, A_n f_i \rangle \le \sup_n \mathrm{Tr}\, A_n \; .$$

Taking the limit $M \to \infty$ we see that $\mathrm{Tr}\, A \le \sup_n \mathrm{Tr}\, A_n$. The other direction is trivial.

For part (ii), we notice that A_n is trace class and

$$\mathrm{Tr}\, A_n = \sum_{i=1}^{\infty} \langle f_i, A_n f_i \rangle \le \sum_{i=1}^{N} \langle f_i, A_n f_i \rangle + \sum_{i=N+1}^{\infty} \langle f_i, A f_i \rangle \; .$$

The second term is smaller than any given $\eta > 0$ if N is big enough since $\mathrm{Tr}\, A < \infty$, while the first term goes to zero for any fixed N as $n \to \infty$.

Q.E.D.

Next we show the following properties of the trace of the square root.

Lemma 9.2. *Let A, B be positive self-adjoint operators on a common Hilbert space. Then*

$$Tr\, \sqrt{AB^2A} = Tr\, \sqrt{BA^2B} \tag{9.1}$$

and

$$Tr\, \sqrt{A+B} \le 2\Big(Tr\, \sqrt{A} + Tr\, \sqrt{B} \Big) \; . \tag{9.2}$$

Proof. We can assume that at least one side of (9.1), say $\mathrm{Tr}\sqrt{AB^2A}$ is finite. Then $X = AB^2A \geq 0$ is trace class and so is $Y := BA^2B$ by cyclicity of the trace. By functional calculus

$$\sqrt{X} = (const.)\int_0^\infty t^{-1/2} X e^{-tX} \mathrm{d}t \, .$$

Since X is bounded, the operator Xe^{-tX} can be expanded into convergent power series for any $t \geq 0$. Using the cyclicity of the trace

$$\mathrm{Tr}\; X^n = \mathrm{Tr}\; (AB^2A)^n = \mathrm{Tr}\; (BA^2B)^n = \mathrm{Tr}\; Y^n$$

for any n, therefore

$$\mathrm{Tr}\; Xe^{-tX} = \mathrm{Tr}\; Ye^{-tY}$$

for any $t \geq 0$ and (9.1) follows after integration. We used Lemma 9.1 (i) to interchange the trace and the improper $\mathrm{d}t$ integration.

For the proof of (9.2) we let $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq 0$ be the eigenvalues of A, similar notation is used for B and $A+B$. Fan's inequality [8] asserts that

$$\lambda_{n+m+1}(A+B) \leq \lambda_{n+1}(A) + \lambda_{m+1}(B)$$

for any $m, n \geq 0$. Therefore

$$\begin{aligned} \lambda_{2n+1}(A+B) &\leq \lambda_{n+1}(A) + \lambda_{n+1}(B) \\ \lambda_{2n}(A+B) &\leq \lambda_{2n-1}(A+B) \leq \lambda_n(A) + \lambda_n(B) \end{aligned}$$

and

$$\begin{aligned} \mathrm{Tr}\; \sqrt{A+B} &= \sum_{k=1}^\infty \sqrt{\lambda_k(A+B)} \\ &\leq 2\sum_{n=1}^\infty \sqrt{\lambda_n(A)+\lambda_n(B)} \\ &\leq 2\sum_{n=1}^\infty \Big(\sqrt{\lambda_n(A)} + \sqrt{\lambda_n(B)}\Big) \\ &= 2\Big(\mathrm{Tr}\; \sqrt{A} + \mathrm{Tr}\; \sqrt{B}\Big) \, . \end{aligned}$$

Q.E.D.

Now we give the definition of the partial trace. For more details, see e.g., [6].

Definition 9.3. *Let H_1, H_2 be separable Hilbert spaces. Let A be a trace class operator on the tensor product $H = H_1 \otimes H_2$. Then there exists a unique operator B in $\mathcal{L}^1(H_1)$ such that*

$$Tr_{H_1}\big[BK\big] = Tr_H\big[A(K \otimes I)\big] \tag{9.3}$$

for every $K \in \mathcal{K}(H_1)$. This operator B is called the partial trace of A with respect to H_2 and is denoted by $Tr_{H_2}A$ or just by $Tr_2\, A$.

The existence of the partial trace follows from duality: the linear functional $K \mapsto \mathrm{Tr}_H A(K \otimes I)$ defines a bounded linear map on $\mathcal{K}(H_1)$ and $[\mathcal{K}(H_1)]^* = \mathcal{L}^1(H_1)$.

Proposition 9.4. *The partial trace satisfies the following relations*

$$Tr_{H_1}\big|Tr_{H_2}(I \otimes A)B\big| = Tr_{H_1}\big|Tr_{H_2}B(I \otimes A)\big| \tag{9.4}$$

and

$$Tr_{H_1}\big|Tr_{H_2}A\big| \le Tr_H|A|\,. \tag{9.5}$$

Proof: For the proof of (9.4) we use the variational principle (2.12)

$$\begin{aligned}
\mathrm{Tr}_1\big|\mathrm{Tr}_2(I \otimes A)B\big| &= \sup \mathrm{Tr}_1\, K\big(\mathrm{Tr}_2(I \otimes A)B\big) \\
&= \sup \mathrm{Tr}_{1,2}\, (K \otimes I)(I \otimes A)B \\
&= \sup \mathrm{Tr}_{1,2}\, (K \otimes I)B(I \otimes A) \\
&= \sup \mathrm{Tr}_1\, K\big(\mathrm{Tr}_2 B(I \otimes A)\big) \\
&= \mathrm{Tr}_1\big|\mathrm{Tr}_2 B(I \otimes A)\big|
\end{aligned} \tag{9.6}$$

where the supremum is over all $K \in \mathcal{K}(H_1)$ with $\|K\| = 1$. We used the cyclicity of $\mathrm{Tr}_{1,2}$ and that $I \otimes A$ commutes with $K \otimes I$.

For the proof of (9.5) we first observe that the variational principle (2.12) extends to bounded operators as follows:

$$\|A\|_1 = \sup_{L \in \mathcal{L}(H)\,:\,\|L\|=1} \Big|\, \mathrm{Tr}\, AL\, \Big| \tag{9.7}$$

for any $A \in \mathcal{L}^1(H)$. The proof follows from $\mathcal{K}(H) \subset \mathcal{L}(H)$ on one hand, and from $|\mathrm{Tr}AL| \le \|A\|_1\|L\|$ on the other hand, using (3.1).

Therefore

$$\begin{aligned}
\mathrm{Tr}_1\, \big|\mathrm{Tr}_2\, A\big| &= \sup_{L \in \mathcal{L}(H_1)\,:\,\|L\|=1} \mathrm{Tr}_1\, L\, \big[\mathrm{Tr}_2\, A\big] \\
&= \sup_{L \in \mathcal{L}(H_1)\,:\,\|L\|=1} \mathrm{Tr}_{1,2}\, (L \otimes I)A \\
&\le \mathrm{Tr}_{1,2}\, |A|\,.
\end{aligned}$$

The estimate follows again from the variational principle (9.7) since $L \otimes I \in \mathcal{L}(H_1 \times H_2)$ and $\|L \otimes I\| = \|L\|$. **Q.E.D.**

Finally, we prove that the space $\mathcal{A}$ is indeed the dual space of the Sobolev space $\mathcal{H}^1$:

Proof of Lemma 3.1. (i) $\mathcal{H}^1 \subset \mathcal{A}^*$. Let $T \in \mathcal{H}^1$ and we write any element $A \in \mathcal{A}$ as $A = SKS$ with some $S \in \mathcal{K}$. Then

$$\Big|\text{Tr}\, TA\Big| = \Big|\text{Tr}\, TSKS\Big| \le \|K\|\ \text{Tr}\,|STS| = \|SKS\|_{\mathcal{A}}\ \|T\|_{\mathcal{H}^1} = \|A\|_{\mathcal{A}}\ \|T\|_{\mathcal{H}^1}$$

which shows that $\ell_T : A \to \text{Tr}\, TA$ is a continuous linear functional on $\mathcal{A}$ indexed by $T \in \mathcal{H}^1$ and $\|\ell_T\|_{\mathcal{A}^*} \le \|T\|_{\mathcal{H}^1}$.

(ii) $\mathcal{A}^* \subset \mathcal{H}^1$. Let $f \in \mathcal{A}^*$ be a continuous linear functional on $\mathcal{A}$. Consider $\psi, \phi \in H^1(\mathbf{R}^3) \subset H$, i.e., $S\psi, S\phi \in H$. Since $|\phi\rangle\langle\psi|$ is compact, $|S\phi\rangle\langle S\psi| = S|\phi\rangle\langle\psi|S$ is contained in $\mathcal{A}$. Consider the sesquilinear functional

$$\mathcal{F}_f : (\psi, \phi) \mapsto f\Big[\ S|\phi\rangle\langle\psi|S\ \Big]$$

from $H^1 \times H^1 \to \mathbf{C}$. This sesquilinear map is continuous since

$$\left|f\Big[\ S|\phi\rangle\langle\psi|S\ \Big]\right| \le \|f\|_{\mathcal{A}^*}\Big\|\ S|\phi\rangle\langle\psi|S\ \Big\|_{\mathcal{A}} = \|f\|_{\mathcal{A}^*}\Big\|\ |\phi\rangle\langle\psi|\ \Big\| \le \|f\|_{\mathcal{A}^*}\|\phi\|\ \|\psi\|\ ,$$

moreover it clearly extends from $H^1 \times H^1$ to $H \times H$.

Hence there exists a unique bounded operator $B = B_f$ such that

$$(\psi, B\phi) = f\Big[\ S|\phi\rangle\langle\psi|S\ \Big]$$

for every $\phi, \psi \in H$, and $\|B\| \le \|f\|_{\mathcal{A}^*}$ (uniqueness follows from the density of $H^1 \subset H$).

By polar decomposition we write $B = U|B|$ with some partial isometry U. Let $\{\psi_i\}$ be a finite orthonormal set, we can write

$$\begin{aligned} \sum_{i=1}^{N}(\psi_i, |B|\psi_i) &= f\Big[\ S\Big(\sum_i |\psi_i\rangle\langle U\psi_i|\Big)S\ \Big] \\ &\le \|f\|_{\mathcal{A}^*}\Big\|\sum_i |\psi_i\rangle\langle U\psi_i|\Big\| \\ &= \|f\|_{\mathcal{A}^*}\ . \end{aligned} \qquad (9.8)$$

Here we used that $U\psi_i$ is also orthonormal, and for any $h \in H$

$$\begin{aligned}\Big(h, \sum_i |\psi_i\rangle\langle U\psi_i|h\Big) &= \sum_i (h,\psi_i)(U\psi_i, h) \\ &\le \frac{1}{2}\sum_i \Big[|(h,\psi_i)|^2 + |(U\psi_i,h)|^2\Big] \\ &\le \|h\|^2 .\end{aligned}$$

Hence taking the supremum for all orthonormal sets in (9.8), we see that $\mathrm{Tr}\,|B| \le \|f\|_{\mathcal{A}^*}$.

Finally, we define $T_f := S^{-1}B_f S^{-1}$ for any $f \in \mathcal{A}^*$. Then clearly

$$\|T_f\|_{\mathcal{H}^1} = \mathrm{Tr}\,|B_f| \le \|f\|_{\mathcal{A}^*}$$

and for any $A \in \mathcal{A}$

$$\mathrm{Tr}\,(T_f A) = f(A)$$

since this is valid for all $A = S\sum_{i=1}^n |\phi_i\rangle\langle\psi_i|S$ finite range operators and these are dense in $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$. Hence for any $f \in \mathcal{A}^*$ we found a representative $T_f \in \mathcal{H}^1$ with smaller or equal norm. **Q.E.D.**

References

[1] C. Bardos, F. Golse and N. Mauser: *Weak coupling limit of the N-particle Schrödinger equation.* To appear in Math. Anal. and Appl. (2001)

[2] C. Bardos, L. Erdős, F. Golse, N. Mauser and H.T. Yau: *Derivation of the Schrödinger-Poisson equation from the quantum N-body problem.* Submitted to C. R. Acad. Sci. (2001)

[3] J. Ginibre and G. Velo: *The classical field limit of scattering theory for non-relativistic many-boson systems. I and II.* Commun. Math. Phys. **66**, 37-76 (1979) and **68**, 45-68 (1979).

[4] J. Ginibre and G. Velo: *On a class of nonlinear Schrödinger equations with nonlocal interactions.* Math. Z. **170**, 109-145 (1980)

[5] K. Hepp: *The classical limit for quantum mechanical correlation functions.* Commun. Math. Phys. **35**, 265-277 (1974).

[6] K. R. Parthasarathy. An Introduction to Quantum Stochastic Calculus. Birkhäuser, 1992.

[7] M. Reed and B. Simon. Methods of Modern Mathematical Physics. Vol.I., II. Academic Press, 1975.

[8] B. Simon: Trace ideals and their applications. Cambridge Univ. Press, 1979.

[9] H. Spohn: *Kinetic Equations from Hamiltonian Dynamics*. Rev. Mod. Phys. **52** no. 3 (1980), 569–615.