

# Functional Analysis II

Phan Thành Nam

Winter 2020-2021

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Weyl's law and semiclassical estimates . . . . .	5
1.2	Laplacian on bounded domains . . . . .	6
1.3	Applications to many-body fermionic systems . . . . .	7
<b>2</b>	<b>Basic spectral properties of Schrödinger operators</b>	<b>10</b>
2.1	Hilbert spaces . . . . .	10
2.2	Self-adjoint operators and Spectral theorem . . . . .	12
2.3	Min-max principle . . . . .	17
2.4	Sobolev inequalities . . . . .	19
2.5	Schrödinger operators . . . . .	25
<b>3</b>	<b>Semiclassical estimates</b>	<b>31</b>
3.1	Cwikel-Lieb-Rozenblum inequality . . . . .	31
3.2	Lieb-Thirring inequality . . . . .	36
3.3	Birman-Schwinger Principle . . . . .	42
3.4	Kato-Seiler-Simon and Cwikel's inequalities . . . . .	48

<b>4</b>	<b>Weyl's law</b>	<b>56</b>
4.1	Coherent States . . . . .	56
4.2	Weyl's law for sum of eigenvalues . . . . .	59
<b>5</b>	<b>Dirichlet Laplacian</b>	<b>70</b>
5.1	Berezin-Li-Yau inequality . . . . .	72
5.2	Sum of eigenvalues . . . . .	75
5.3	Distribution of eigenvalues . . . . .	77
5.4	Pólya conjecture . . . . .	80
5.5	Weyl's conjecture . . . . .	85
5.6	Can one hear the shape of a drum? . . . . .	91
<b>6</b>	<b>Neumann Laplacian</b>	<b>94</b>
6.1	Essential spectrum vs. Compact resolvent . . . . .	96
6.2	Extension domains . . . . .	97
6.3	Kröger's inequality . . . . .	103
6.4	Lieb-Thirring inequality for Neumann Laplacian . . . . .	105
6.5	Weyl's law . . . . .	110
6.6	Pólya conjecture . . . . .	113
<b>7</b>	<b>Many-body quantum systems</b>	<b>122</b>
7.1	Slater determinants . . . . .	123
7.2	Reduced density matrices . . . . .	126
7.3	Ideal Fermi gas . . . . .	133
<b>8</b>	<b>Thomas-Fermi theory</b>	<b>140</b>
8.1	Density functional theory . . . . .	140
8.2	Convergence of the kinetic density functional . . . . .	144
8.3	Convergence of the Levy-Lieb functional . . . . .	151
8.4	Convergence of ground state energy and ground states . . . . .	156
8.5	Atomic Thomas-Fermi minimizer . . . . .	160
<b>9</b>	<b>Hartree-Fock theory</b>	<b>167</b>
9.1	Lieb-Oxford inequality . . . . .	171
9.2	Error bound for atomic reduced Hartree-Fock energy . . . . .	177
9.3	Bach's correlation inequality . . . . .	183

9.4	Error bound for atomic Hartree–Fock energy . . . . .	191
9.5	Hartree–Fock energy of a homogeneous Fermi gas . . . . .	194
<b>10</b>	<b>Correlation energy</b>	<b>202</b>
10.1	Fock space formalism . . . . .	203
10.2	Particle–hole transformation . . . . .	210
10.3	Estimates for kinetic and number operators . . . . .	215
10.4	Removing the non-bosonizable term . . . . .	220
10.5	Diagonalization of bosonizable term . . . . .	231
<b>11</b>	<b>Stability of matter</b>	<b>235</b>
11.1	Teller’s no-binding theorem . . . . .	236
11.2	First proof of the stability of matter . . . . .	239
11.3	Baxter’s electrostatic inequality . . . . .	241
11.4	Second proof of the stability of matter . . . . .	246
11.5	Existence of thermodynamic limit . . . . .	248
11.6	Grand-canonical stability . . . . .	252
11.7	Instability for bosons . . . . .	254

# Chapter 1

## Introduction

In classical mechanics, a particle can be described by a pair of position-momentum

$$(x, p) \in \mathbb{R}^d \times \mathbb{R}^d.$$

In quantum mechanics, we cannot determine both position and momentum at the same time (by uncertainty principle). Therefore, a quantum particle has to be described by a normalized function  $\psi \in L^2(\mathbb{R}^d)$  with

$$|\psi(x)|^2 = \text{probability density of position}, \quad |\widehat{\psi}(k)|^2 = \text{probability density of momentum}.$$

The **semiclassical approximation** suggests to relate quantum quantities by classical ones, using the idea that

$$\text{a quantum state} \approx \text{a unit volum in the phase space } \mathbb{R}^d \times \mathbb{R}^d.$$

For example, for the sum of negative eigenvalues of a Schrödinger operator with a real-valued potential  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  we can expect

$$\text{Tr}[-\Delta + V(x)]_- \approx \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [|2\pi k|^2 + V(x)]_- dk dx = -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V(x)_-|^{1+d/2} dx$$

where  $a_- = \min(a, 0)$  and

$$L_{1,d}^{\text{cl}} = \int_{\mathbb{R}^d} |(|2\pi k|^2 - 1)_-| dk = \frac{2}{d+2} \cdot \frac{|B_1|}{(2\pi)^d}.$$

## 1.1 Weyl's law and semiclassical estimates

A cornerstone of the semiclassical analysis for Schrödinger operators is **Weyl's law**: for all  $d \geq 1$ , if  $V_- \in L^{1+d/2}(\mathbb{R}^d)$ , then

$$\mathrm{Tr}[-\Delta + \lambda V(x)]_- = -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |\lambda V(x)_-|^{1+d/2} dx + o(\lambda^{1+d/2})_{\lambda \rightarrow \infty}.$$

By introducing  $\hbar := \lambda^{-1/2}$ , Weyl's law is equivalent to

$$\mathrm{Tr}[-\hbar^2 \Delta + V(x)]_- = -\hbar^{-d} L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |V(x)_-|^{1+d/2} dx + o(\hbar^{-d})_{\hbar \rightarrow 0}.$$

This is consistent to Bohr's correspondence principle that the behavior of a system described by quantum mechanics reproduces classical mechanics in the limit  $\hbar \rightarrow 0$ .

In our applications, it is important to have quantitative estimates for finite parameters. In 1975, **Lieb and Thirring** proved that for all  $d \geq 1$ , if  $V_- \in L^{1+d/2}(\mathbb{R}^d)$ , then

$$\mathrm{Tr}[-\Delta + V(x)]_- \geq -L_{1,d} \int_{\mathbb{R}^d} |V(x)_-|^{1+d/2} dx.$$

Here  $L_{1,d}$  is independent of  $V$ .

Note that if we just look at the first eigenvalue, then using Sobolev's inequality it is not hard to prove that

$$\lambda_1(-\Delta + V(x)) \geq -L_{1,d}^{\mathrm{So}} \int_{\mathbb{R}^d} |V(x)_-|^{1+d/2} dx.$$

The Lieb-Thirring inequality is deeper! It is related to **Pauli's exclusion principle**, while Sobolev's inequality is a version of the **uncertainty principle**.

Obviously, the optimal constant in the Lieb-Thirring inequality satisfies  $L_{1,d} \geq \max(L_{1,d}^{\mathrm{cl}}, L_{1,d}^{\mathrm{So}})$ . Lieb and Thirring **conjectured** that

$$L_{1,d} = \max(L_{1,d}^{\mathrm{cl}}, L_{1,d}^{\mathrm{So}}) = \begin{cases} L_{1,d}^{\mathrm{cl}} & \text{if } d \geq 3, \\ L_{1,d}^{\mathrm{So}} & \text{if } d = 1, 2. \end{cases}$$

This is an important open problem in spectral theory and mathematical physics.

More generally, the semiclassical approximation also applies to the sum of moments of eigenvalues of Schrödinger operators. In particular, the number of negative eigenvalues of

$-\Delta + V(x)$  can be approximated by

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}(|2\pi k|^2 + V(x) < 0) dk dx = L_{0,d}^{\text{cl}} \int_{\mathbb{R}^d} |V(x)_-|^{d/2} dx, \quad L_{0,d}^{\text{cl}} = \frac{|B_1|}{(2\pi)^d}$$

Weyl's law also extends to this case, namely the approximation holds to the leading order with  $V \mapsto \lambda V$ ,  $\lambda \rightarrow +\infty$ . For finite parameters, the **Cwikel-Lieb-Rozenblum (CLR) inequality** states that if  $d \geq 3$ , then the number of negative eigenvalues of  $-\Delta + V(x)$  is bounded by

$$L_{0,d} \int_{\mathbb{R}^d} |V(x)_-|^{d/2} dx.$$

This bound fails in 1 and 2 dimensions. Moreover, in general  $L_{0,d} > L_{0,d}^{\text{cl}}$  for all dimensions.

## 1.2 Laplacian on bounded domains

Note that in the above discussion, we do not put conditions on  $V_+$ . In particular, we may consider the **hard sphere potential of a bounded domain**  $\Omega \subset \mathbb{R}^d$ :

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

The corresponding **Dirichlet Laplacian**  $-\Delta$  has eigenvalues  $0 < \lambda_1 < \lambda_2 \leq \dots$  with  $\lim_{n \rightarrow \infty} \lambda_n = +\infty$ . These eigenvalues can be computed using the **min-max principle** and the **quadratic form domain**  $H_0^1(\Omega)$ . They also satisfy the equations

$$-\Delta u_n = \lambda_n u_n, \quad u_n|_{\partial\Omega} = 0.$$

Obviously, the number of eigenvalues below  $\lambda$  of the Dirichlet Laplacian is the same to the number of negative eigenvalues of  $-\Delta - \lambda$ . Therefore, Weyl's law tells us

$$N(\lambda) = L_{0,d}^{\text{cl}} \lambda^{d/2} |\Omega| + o(\lambda^{d/2})_{\lambda \rightarrow \infty}, \quad L_{0,d}^{\text{cl}} = \frac{|B_1|}{(2\pi)^d},$$

which is equivalent to

$$\lambda_n = \frac{(2\pi)^2}{(|B_1| |\Omega|)^{2/d}} n^{2/d} + o(n^{2/d})_{n \rightarrow \infty}.$$

The asymptotic formula in this form was first proved by **Hermann Weyl** in 1911, solving a conjecture of **Sommerfeld-Lorentz** in 1910. There are also many works for the next order

correction (involving  $|\partial\Omega|$ ), including notable papers of **Victor Ivrii** in 1980.

The eigenvalue problem for the Laplacian on bounded domains was strongly motivated by music, going back to **Rayleigh** in 1877 with “The Theory of Sound”. This has become more popular since **Kac’s** 1966 paper “**Can one hear the shape of a drum?**”. This is an interesting **inverse problem**: assuming that we know all eigenvalues of the Dirichlet Laplacian on  $\Omega$ , can we determine the shape of  $\Omega$ ? Clearly, we can “hear” at least the volume of  $\Omega$  (or even  $|\partial\Omega|$ ). But can we hear more, e.g. can we determine  $\Omega$  uniquely (up to translations and rotations)?

An important open problem is **Pólya’s conjecture** for the sharp eigenvalue bound

$$\lambda_n \geq \frac{(2\pi)^2}{(|B_1||\Omega|)^{2/d}} n^{2/d}, \quad \forall n \geq 1.$$

It is straightforward to see that this bound holds when  $\Omega$  is a cube. Pólya proved this bound for **tiling domains**; however, the problem is still open even when  $\Omega$  is a ball!

The best known result in this direction is the **Berezin-Li-Yau inequality**

$$\sum_{n=1}^N \lambda_n \geq K^{\text{cl}} |\Omega|^{-2/d} N^{1+2/d}, \quad K^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{2/d}}.$$

This is a consequence of Pólya’s conjecture, and also a particular confirmation of the Lieb-Thirring conjecture for the hard sphere potential (up to a constant shift). It turns out that the Berezin-Li-Yau inequality can be used to prove a Lieb-Thirring inequality with the semiclassical constant plus some error which is small in applications.

### 1.3 Applications to many-body fermionic systems

From first principles of quantum mechanics, a system of  $N$  **identical fermions** in  $\mathbb{R}^d$  is described by a (normalized) wave function in  $L^2(\mathbb{R}^{dN})$  satisfying the **anti-symmetry**

$$\Psi(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -\Psi(x_1, \dots, x_j, \dots, x_i, \dots, x_N), \quad \forall i \neq j, \quad \forall x_i \in \mathbb{R}^d.$$

A typical **many-body Schrödinger operator** has the form

$$H_N = \sum_{i=1}^N (-\Delta_{x_i} + V(x_i)) + \sum_{1 \leq i < j \leq N} W(x_i - x_j)$$

acting on the anti-symmetric space  $L_a^2(\mathbb{R}^{dN})$ . The ground state energy is

$$E_N = \inf_{\|\Psi\|_{L^2}=1} \langle \Psi, H_N \Psi \rangle.$$

If a ground state exists, it satisfies the Schrödinger equation

$$H_N \Psi = E_N \Psi.$$

When  $N$  becomes large, this beautiful linear theory is very difficult to compute (even numerically) because there are so many variables. In practice, people often rely on effective theories, which are nonlinear but dependent on less variables. A popular method in computational physics and chemistry is **density functional theory**, where the complicated wave function  $\Psi : \mathbb{R}^{dN} \rightarrow \mathbb{C}$  is replaced by its one-body density  $\rho_\Psi : \mathbb{R}^d \rightarrow [0, \infty)$

$$\rho_\Psi(x) = N \int_{\mathbb{R}^{d(N-1)}} |\Psi(x, x_2, \dots, x_N)|^2 dx_2 \dots dx_N, \quad \int_{\mathbb{R}^d} \rho_\Psi = N.$$

The oldest density functional theory is Thomas-Fermi theory (1927), where the ground state energy is computed by

$$E_N^{\text{TF}} = \inf_{\int \rho = N} \left\{ K^{\text{cl}} \int_{\mathbb{R}^d} \rho^{1+2/d} + \int_{\mathbb{R}^d} V \rho + \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho(x) \rho(y) W(x-y) dx dy \right\}.$$

In particular, the Thomas-Fermi approximation for the kinetic energy

$$\left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i}) \Psi \right\rangle \approx K^{\text{cl}} \int_{\mathbb{R}^d} \rho_\Psi^{1+2/d}$$

can be interpreted as a **dual version** of Weyl's law for the eigenvalue sum of (one-body) Schrödinger operators. A key concept here is **Pauli's exclusion principle** which states that we cannot put two fermions in a common quantum state. Mathematically, if we define the **one-body density matrix**  $\gamma_\Psi^{(1)}$  as a trace class operator on  $L^2(\mathbb{R}^d)$  with kernel

$$\gamma_\Psi^{(1)}(x, y) = N \int_{\mathbb{R}^{d(N-1)}} \Psi(x, x_2, \dots, x_N) \overline{\Psi(y, x_2, \dots, x_N)} dx_2 \dots dx_N, \quad \text{Tr } \gamma_\Psi^{(1)} = N,$$

then Pauli's exclusion principle states that

$$0 \leq \gamma_\Psi^{(1)} \leq 1.$$



Consequently, for any one-body potential  $R : \mathbb{R}^d \rightarrow \mathbb{R}$  we have

$$\begin{aligned} \left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i}) \Psi \right\rangle + \int_{\mathbb{R}^d} R \rho_{\Psi} &= \left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i} + R(x_i)) \Psi \right\rangle = \text{Tr}((-\Delta + R)\gamma) \\ &\geq \sum_{i=1}^N \lambda_i(-\Delta + R(x)) \geq \text{Tr}[-\Delta + R(x)]_- \approx -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |R_-|^{1+d/2}. \end{aligned}$$

By optimizing over  $R$ , we obtain an appropriate justification for the Thomas-Fermi approximation for the kinetic energy. More generally, under suitable conditions on the interaction potential  $W$ , we will show that

$$\lim_{N \rightarrow \infty} \frac{E_N^{\text{TF}}}{E_N} = 1.$$

Normally, the Thomas-Fermi theory tells us the leading order behavior of weakly interacting Fermi gas. A better approximation is the **Hartree-Fock theory**, in which we restrict the wave functions to **Slater determinants**

$$(u_1 \wedge u_2 \wedge \dots \wedge u_N)(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det \left[ (u_i(x_j))_{1 \leq i, j \leq N} \right]$$

where  $\{u_i\}_{i=1}^N$  are orthonormal vectors in  $L^2(\mathbb{R}^d)$ . The corresponding Hartree-Fock energy is

$$E_N^{\text{HF}} = \inf \left\{ \langle \Psi, H_N \Psi \rangle : \Psi \text{ a Slater determinant} \right\}.$$

Obviously  $E_N \leq E_N^{\text{HF}}$  by the variational principle, but the lower bound of  $E_N - E_N^{\text{HF}}$  is nontrivial. We will discuss rigorous estimates allowing to justify the Hartree-Fock energy for a class of Fermi gas.

At the end of the course, we will turn to the **correlation energy**  $E_N - E_N^{\text{HF}}$ . This difficult question is an important topic of recent research. We will discuss the **random phase approximation**, first invented by Bohm and Pines in 1952 for Coulomb systems. In the case of **weak, short range interactions**, their argument can be justified rigorously. In this case, there are some collective pairs of fermions where each pair behaves as a **boson**, and the correlation energy is computed using **Bogoliubov's theory** for weakly interacting Bose gases.

# Chapter 2

## Basic spectral properties of Schrödinger operators

### 2.1 Hilbert spaces

**Definition** (Hilbert Spaces). *A space  $\mathcal{H}$  is a Hilbert space if*

- $\mathcal{H}$  is a complex vector space, equipped with an inner product  $\langle \cdot, \cdot \rangle$  which is linear in the second argument and anti-linear in the first

$$\langle x, \lambda y \rangle = \lambda \langle x, y \rangle, \quad \langle \lambda x, y \rangle = \bar{\lambda} \langle x, y \rangle;$$

- $(\mathcal{H}, \|\cdot\|)$  is a Banach (complete normed) space with norm  $\|x\| = \sqrt{\langle x, x \rangle}$ .

**Theorem** (Riesz's representation theorem). *For any Hilbert space  $\mathcal{H}$ , there exists an anti-linear unitary operator  $J : \mathcal{H} \rightarrow \mathcal{H}^*$  such that*

$$J(x)(y) = \langle x, y \rangle, \quad \forall x, y \in \mathcal{H}.$$

*In particular,  $\|J(x)\|_{\mathcal{H}^*} = \|x\|_{\mathcal{H}}$ .*

Recall that the **dual space**  $\mathcal{H}^*$  contains all linear bounded operators from  $\mathcal{H} \rightarrow \mathbb{C}$ . Sometimes, it is convenient to use the **bra-ket notation** where

$$|x\rangle \in \mathcal{H}, \quad \langle x| = J(x) \in \mathcal{H}^*.$$

Then  $\langle x|y\rangle = \langle x, y\rangle$  (the usual inner product) and  $|y\rangle\langle x|$  is a rank-one operator.

**Definition** (Weak convergence). Let  $\{x_n\}_{n=1}^{\infty} \subset \mathcal{H}$  and  $x \in \mathcal{H}$ . We write  $x_n \rightharpoonup x$ , namely  $x_n$  converges to  $x$  weakly in  $\mathcal{H}$ , if

$$\langle x_n, y\rangle \rightarrow \langle x, y\rangle, \quad \forall y \in \mathcal{H}.$$

Remark: In general, if  $\mathcal{H}$  is a normed space, people usually say that  $x_n \rightharpoonup x$  weakly in  $\mathcal{H}$  if

$$F(x_n) \rightarrow F(x), \quad \forall F \in \mathcal{H}^*.$$

When  $\mathcal{H}$  is a Hilbert space, the latter definition is equivalent to the former, thanks to Riesz's representation theorem. In this case, the weak convergence  $x_n \rightharpoonup x$  in  $\mathcal{H}$  is also equivalent to the **weak-\* convergence**  $J(x_n) \rightharpoonup^* J(x)$  in  $\mathcal{H}^*$ .

The concept of weak convergence is very helpful to gain the compactness.

**Theorem.** Let  $\mathcal{H}$  be a Hilbert space and let  $\{x_n\}_{n=1}^{\infty} \subset \mathcal{H}$ .

- (Banach–Steinhaus) If  $\{x_n\}$  converges weakly in  $\mathcal{H}$ , then  $\{x_n\}$  is bounded.
- (Banach–Alaoglu) If  $\{x_n\}$  is bounded, then there exists a subsequence  $\{x_{n_k}\}_{k=1}^{\infty}$  which converges weakly in  $\mathcal{H}$ .

**Definition.** The Hilbert space  $\mathcal{H}$  is **separable** if there exists an **orthonormal basis**  $\{u_n\}_{n \geq 1}$  (finite or countable).

We will always work on separable Hilbert spaces.

**Theorem.** If the Hilbert space  $\mathcal{H}$  is separable and  $\{u_n\}_{n \geq 1}$  is an orthonormal basis, then we have **Parseval's identity**

$$x = \sum_{n \geq 1} \langle u_n, x\rangle u_n, \quad \forall x \in \mathcal{H}.$$

Consequently,

$$\|x\|^2 = \sum_{n \geq 1} |\langle u_n, x \rangle|^2, \quad \forall x \in \mathcal{H}.$$

In particular, we have **Bessel's inequality**: for any orthonormal family  $\{u_n\}$ ,

$$\|x\|^2 \geq \sum_n |\langle u_n, x \rangle|^2, \quad \forall x \in \mathcal{H}.$$

**Exercise.** Let  $\{u_n\}_{n=1}^\infty$  be an orthonormal family in a Hilbert space  $\mathcal{H}$ . Prove that  $u_n \rightharpoonup 0$  weakly in  $\mathcal{H}$ .

## 2.2 Self-adjoint operators and Spectral theorem

**Definition** (Unbounded operators). An operator  $A$  on  $\mathcal{H}$  is a linear map  $A : D(A) \rightarrow \mathcal{H}$  with a dense, subspace  $D(A) \subset \mathcal{H}$  (domain of  $A$ ). The **adjoint operator**  $A^* : D(A^*) \rightarrow \mathcal{H}$  is defined by

$$D(A^*) = \left\{ x \in \mathcal{H} \mid \exists A^*x \in \mathcal{H} : \langle x, Ay \rangle = \langle A^*x, y \rangle, \quad \forall y \in D(A) \right\}.$$

The operator  $A$  is **self-adjoint** if  $A = A^*$ .

**Definition** (Spectrum). Let  $A : D(A) \rightarrow \mathcal{H}$  be an operator on a Hilbert space  $\mathcal{H}$ . Its **spectrum** is

$$\sigma(A) = \mathbb{C} \setminus \{ \lambda \in \mathbb{C} : (A - \lambda)^{-1} \text{ is a bounded operator} \}.$$

We can decompose

$$\sigma(A) = \sigma_{\text{dis}}(A) \cup \sigma_{\text{ess}}(A), \quad \sigma_{\text{dis}}(A) \cap \sigma_{\text{ess}}(A) = \emptyset,$$

where the **discrete spectrum**  $\sigma_{\text{dis}}(A)$  is the set of isolated eigenvalues with finite multiplicities and the **essential spectrum**  $\sigma_{\text{ess}}(A)$  is the complement.

Note that the spectrum and the essential spectrum are always **closed** sets.

**Exercise.** Let  $\Omega \subset \mathbb{R}^d$  be a Borel set,  $\mu$  a locally finite Borel measure on  $\Omega$ , and  $a \in L^\infty_{\text{loc}}(\Omega, \mu)$  a real-valued function. Consider the **multiplication operator**  $M_a$  on  $L^2(\Omega, \mu)$  defined by

$$(M_a f)(x) = a(x)f(x), \quad D(M_a) = \{f \in L^2(\Omega, \mu), af \in L^2(\Omega, \mu)\}.$$

Prove that

(i)  $M_a$  is a self-adjoint operator and  $\sigma(M_a) = \text{ess-range}(a) \subset \mathbb{R}$ , namely

$$\lambda \in \sigma(M_a) \quad \text{iff} \quad \mu(a^{-1}(\lambda - \varepsilon, \lambda + \varepsilon)) > 0, \quad \forall \varepsilon > 0.$$

(ii)  $\lambda$  is an eigenvalue of  $M_a$  iff  $\mu(a^{-1}(\lambda)) > 0$ .

(iii)  $\lambda \in \sigma_{\text{dis}}(M_a)$  iff  $\lambda$  is an isolated point of  $\sigma(M_a)$  and  $0 < \mu(a^{-1}(\lambda)) < \infty$ .

A cornerstone of spectral theory is the **Spectral theorem** which says that any self-adjoint operator is unitarily equivalent to a multiplication operator.

**Theorem** (Spectral theorem). Assume that  $A$  is a self-adjoint operator on a Hilbert space  $\mathcal{H}$ . Then there exists a Borel set  $\Omega \subset \mathbb{R}^d$  for some  $d \geq 1$ , a locally finite Borel measure  $\mu$  on  $\Omega$ , a real-valued function  $a \in L^\infty_{\text{loc}}(\Omega, \mu)$  and a unitary transformation  $U : \mathcal{H} \rightarrow L^2(\Omega, \mu)$  such that

$$UAU^* = M_a.$$

Here  $M_a$  is the multiplication operator on  $L^2(\Omega, \mu)$ , defined by

$$(M_a f)(x) = a(x)f(x), \quad D(M_a) = \{f \in L^2(\Omega, \mu), af \in L^2(\Omega, \mu)\}.$$

We can choose  $\Omega = \sigma(A) \times \mathbb{N}$ ,  $a(\lambda, n) = \lambda$  and  $\mu$  being a locally finite measure. In this case,  $\lambda$  is an eigenvalue of  $A$  if and only if there exists  $n \in \mathbb{N}$  such that  $\mu(\{(\lambda, n)\}) > 0$ ; moreover, the number of such  $n$ 's is equal to the multiplicity of  $\lambda$ .

Remark: As a consequence of the spectral theorem, if  $A$  is self-adjoint, then  $\sigma(A) \subset \mathbb{R}$ .

Given a self-adjoint operator  $A$  on a Hilbert space  $\mathcal{H}$  and a smooth function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , we

can define the self-adjoint operator  $f(A)$  on  $\mathcal{H}$  by

$$Uf(A)U^* = f(UAU^*) = f(M_a)$$

where  $U : \mathcal{H} \rightarrow L^2(\Omega)$  is a unitary transformation making  $UAU^* = M_a$ . This is called the **functional calculus**. Moreover, the spectral theorem can be used to prove several abstract results for self-adjoint operators.

**Exercise.** Let  $A : D(A) \rightarrow \mathcal{H}$  be a self-adjoint operator on a Hilbert space  $\mathcal{H}$ . Prove that  $A$  is a bounded operator if and only if  $D(A) = \mathcal{H}$ .

**Exercise.** Let  $A : D(A) \rightarrow \mathcal{H}$  be a self-adjoint operator on a Hilbert space  $\mathcal{H}$ . Prove that the following statements are equivalent:

- (i)  $A \geq 0$ , namely  $\langle u, Au \rangle \geq 0$  for all  $u \in D(A)$ ;
- (ii)  $\sigma(A) \subset [0, \infty)$ .

**Exercise (Weyl's Criterion).** For any self-adjoint operator  $A$  on a Hilbert space  $\mathcal{H}$ , prove that the following statements hold true:

- (i)  $\lambda \in \sigma(A)$  iff there exists a **Weyl sequence**  $\{u_n\} \subset D(A)$  such that

$$\|u_n\| = 1, \quad \|(A - \lambda)u_n\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

- (ii)  $\lambda \in \sigma_{\text{ess}}(A)$  iff there exists a Weyl sequence  $\{u_n\} \subset D(A)$  such that

$$\{u_n\} \text{ an orthonormal family, } \|(A - \lambda)u_n\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

**Definition.** Let  $A$  be a bounded operator on a Hilbert space  $\mathcal{H}$ . We say that  $A$  is a **compact operator** if  $A$  maps any bounded set to a pre-compact set.

**Exercise.** Let  $A$  be a bounded operator on a Hilbert space  $\mathcal{H}$ . Prove that  $A$  is a compact

operator iff  $A$  maps weak convergence to strong convergence, namely

$$(x_n \rightharpoonup x \text{ weakly in } \mathcal{H}) \implies (Ax_n \rightarrow Ax \text{ strongly in } \mathcal{H}).$$

**Theorem** (Spectral theorem for compact operators). *Let  $A$  be a self-adjoint compact operator on a separable Hilbert space  $\mathcal{H}$ . Then  $A$  has an orthonormal eigenbasis  $\{u_n\} \subset \mathcal{H}$  with eigenvalues  $\{\lambda_n\} \subset \mathbb{R}$  and  $\lambda_n \rightarrow 0$ . In short, we have the spectral decomposition*

$$A = \sum_{n \geq 1} \lambda_n |u_n\rangle \langle u_n|.$$

**Exercise.** *Let  $\{u_n\}_{n=1}^{\infty}$  be an orthonormal family in a Hilbert space  $\mathcal{H}$ . Let  $\{\lambda_n\}_{n=1}^{\infty} \subset \mathbb{R}$  be a bounded sequence. Consider the operator*

$$A := \sum_{n \geq 1} \lambda_n |u_n\rangle \langle u_n|.$$

- (i) *Prove that  $A$  is a bounded, self-adjoint operator.*
- (ii) *Prove that  $A$  is a compact operator if and only if  $\lim_{n \rightarrow \infty} \lambda_n = 0$ .*

In practice, the self-adjointness is not always easy to verify. A weaker concept is

**Definition.** *An operator  $A$  on a Hilbert space  $\mathcal{H}$  is **symmetric** if*

$$\langle x, Ay \rangle = \langle Ax, y \rangle, \quad \forall x, y \in D(A).$$

*This is also equivalent to*

$$\langle x, Ax \rangle \in \mathbb{R}, \quad x \in D(A).$$

Obviously, if  $A$  is self-adjoint, then  $A$  is symmetric. But the reverse is not true. Two useful methods to find **self-adjoint extensions** for symmetric operators are **Kato-Rellich theorem** and **Friedrichs' extension**.

**Theorem** (Kato-Rellich theorem). *Let  $A$  be a self-adjoint operator on a Hilbert space*

$\mathcal{H}$ . Let  $B$  be a symmetric operator on  $\mathcal{H}$  satisfying

$$\|Bx\| \leq (1 - \varepsilon)\|Ax\| + C_\varepsilon\|x\|, \quad \forall x \in D(A) \subset D(B),$$

for some constant  $\varepsilon > 0$  independent of  $x$  (we say that  $B$  is  $A$ -relatively bounded with the relative bound  $1 - \varepsilon$ ). Then  $A + B$  is a self-adjoint operator on  $\mathcal{H}$  with domain  $D(A + B) = D(A)$ .

**Exercise.** Let  $A$  be a self-adjoint operator on a Hilbert space. Let  $B$  be a symmetric operator which is  $A$ -relatively compact, namely  $D(A) \subset D(B)$  and  $B(A + i)^{-1}$  is a compact operator. Then  $A + B$  is self-adjoint on  $D(A)$  and

$$\sigma_{\text{ess}}(A + B) = \sigma_{\text{ess}}(A).$$

*Hint:* You can write  $B = B(A + i)^{-1}(A + i)$  and use Weyl's Criterion.

**Theorem** (Friedrichs' extension). Assume that  $A$  is bounded from below, namely

$$\langle x, Ax \rangle \geq -C\|x\|^2, \quad \forall x \in D(A)$$

with a finite constant  $C$  independent of  $x$ . Then  $A$  has a self-adjoint extension  $A_F$  which has the same **quadratic form domain**  $Q(A_F) = Q(A)$ . Recall that we define  $Q(A)$  as the closure of  $D(A)$  under the quadratic form norm  $\|x\|_{Q(A)} = \sqrt{\langle x, (A + C + 1)x \rangle}$ .

Note that in general the domain  $D(A_F)$  is often not known explicitly (unlike the extension given by Kato-Rellich theorem). Nevertheless, all eigenvalues below the essential spectrum of  $A_F$  can be computed without knowing the domain  $D(A_F)$ , thanks to the **min-max principle**. In particular, the Friedrichs extension preserves the **ground state energy**

$$\inf_{x \in D(A), \|x\|=1} \langle x, Ax \rangle = \inf_{x \in D(A_0), \|x\|=1} \langle x, A_0x \rangle.$$

In fact, the Friedrichs extension is the largest possible extension of an operator (in the sense Krein's characterization).



## 2.3 Min-max principle

**Theorem** (Min-Max Principle). *Let  $A$  be a self-adjoint operator on a Hilbert space  $\mathcal{H}$ . Assume that  $A$  is bounded from below and define the **min-max values***

$$\mu_n(A) = \inf_{\substack{M \subset D(A) \\ \dim M = n}} \sup_{\substack{u \in M \\ \|u\|=1}} \langle u, Au \rangle.$$

*Then  $\mu_n(A)$  is an increasing sequence and its the limit  $\mu_\infty(A) := \lim_{n \rightarrow \infty} \mu_n(A) \leq +\infty$  satisfies*

$$\mu_\infty(A) = \inf \sigma_{\text{ess}}(A).$$

*Moreover, if  $\mu_n(A) < \mu_\infty(A)$ , then  $\mu_1, \dots, \mu_n$  are the lowest eigenvalues of  $A$ .*

Remarks:

- Here we use the convention that if  $\sigma_{\text{ess}}(A) = \emptyset$ , then  $\inf \sigma_{\text{ess}}(A) = +\infty$ .
- In the above definition, the condition  $M \subset D(A)$  can be replaced by  $M \subset \mathfrak{D}$  for any subspace  $\mathfrak{D}$  which is dense in the quadratic form domain  $Q(A)$ . Thus if  $A$  is the Friedrichs' extension of a (densely defined) operator  $A_0$ , then the min-max values can be computed using the domain  $D(A_0)$ .
- The min-max values is **monotone increasing in operator**, namely if  $A \leq B$ , then

$$\mu_n(A) \leq \mu_n(B), \quad \forall n = 1, 2, \dots$$

*Proof. Step 1.* We prove that  $\mu_\infty(A) \leq \inf \sigma_{\text{ess}}(A)$ . We take  $\lambda \in \sigma_{\text{ess}}(A)$  and prove that

$$\mu_n(A) \leq \lambda, \quad \forall n \geq 1.$$

By Weyl's criterion, since  $\lambda \in \sigma_{\text{ess}}(A)$ , there exists an orthonormal family  $\{u_m\}_{m=1}^\infty \subset D(A)$  such that  $\|(A - \lambda)u_m\| \rightarrow 0$ . Consider the space

$$M_{m,n} = \text{Span}(u_{m+1}, \dots, u_{m+n}), \quad \dim M_{m,n} = n.$$

By the definition of the min-max values,

$$\mu_n(A) \leq \sup_{\substack{u \in M_{m,n} \\ \|u\|=1}} \langle u, Au \rangle.$$

On the other hand, since  $\lim_{m \rightarrow \infty} \|(A - \lambda)u_{m+k}\| = 0$  for all  $k = 1, 2, \dots, n$ , we have

$$\lim_{m \rightarrow \infty} \sup_{\substack{u \in M_{m,n} \\ \|u\|=1}} \langle u, Au \rangle = \lambda.$$

Therefore,  $\mu_n(A) \leq \lambda$ . This holds for all  $n \geq 1$ , implying that  $\mu_\infty(A) \leq \inf \sigma_{\text{ess}}(A)$ .

**Step 2.** We prove that if  $\mu_1(A) < \mu_\infty(A)$ , then  $\mu_1(A)$  is the lowest eigenvalue. Using  $\mu_\infty(A) \leq \inf \sigma_{\text{ess}}(A)$  from Step 1, we find that  $\mu_1(A) \notin \sigma_{\text{ess}}(A)$ . On the other hand,

$$\mu_1(A) = \inf_{\|u\|=1} \langle u, Au \rangle = \inf \sigma(A) \in \sigma(A).$$

Thus  $\mu_1(A) \in \sigma_{\text{dis}}(A)$ , namely it is an eigenvalue with finite multiplicity. Clearly it is the lowest eigenvalue.

**Step 3.** We prove that if  $\mu_2(A) < \mu_\infty(A)$ , then  $\mu_2(A)$  is the second lowest eigenvalue. By Step 2, we know that  $\mu_1(A)$  is an eigenvalue with an eigenvector  $u_1$ . Then  $A$  leaves invariant the space  $\mathcal{H}_1 = \{u_1\}^\perp$  and we can define  $A_1 = A|_{\mathcal{H}_1}$  as an operator on  $\mathcal{H}_1$ . Note that thanks to the decomposition

$$A = \mu_1(A)|u_1\rangle\langle u_1| \oplus A_1$$

we find that  $\sigma(A) = \{\mu_1(A)\} \cup \sigma(A_1)$  and

$$\mu_1(A_1) = \mu_2(A)$$

(**why?**). Consequently,  $\mu_2(A) \in \sigma(A_1) \subset \sigma(A)$ . Thus the condition  $\mu_2(A) < \mu_\infty(A)$  and the inequality  $\mu_\infty(A) \leq \inf \sigma_{\text{ess}}(A)$  from Step 1 imply that  $\mu_2(A) \in \sigma_{\text{dis}}(A)$ , namely it is an eigenvalue of  $A$ . Moreover,  $\mu_2(A) = \mu_1(A_1)$  the lowest eigenvalue of  $A_1$ , and hence  $\mu_2(A)$  is the second lowest eigenvalue of  $A$ .

**Step 4.** By the same argument, we can prove that if  $\mu_n(A) < \mu_\infty(A)$ , then  $\mu_n(A)$  is the  $n$ -th lowest eigenvalue of  $A$ . Moreover, if  $\mu_n(A) < \mu_{n+1}(A) = \mu_\infty(A)$ , then  $\mu_\infty(A) \in \sigma(A)$ . Thus in all cases, all min-max values  $\mu_n(A)$  belong to  $\sigma(A)$ . Therefore,  $\mu_\infty(A) = \lim_{n \rightarrow \infty} \mu_n(A)$

belongs to  $\sigma_{\text{ess}}(A)$ . Combining with the information  $\mu_{\infty}(A) \leq \inf \sigma_{\text{ess}}(A)$  in Step 1, we find that  $\mu_{\infty}(A) = \inf \sigma_{\text{ess}}(A)$ .  $\square$

**Exercise.** Let  $A$  be a self-adjoint operator on a Hilbert space. Assume that  $A$  is bounded from below and its min-max values satisfies

$$\lim_{n \rightarrow \infty} \mu_n(A) = +\infty.$$

- (i) Prove that  $A$  has eigenvalues  $\{\mu_n(A)\}_{n=1}^{\infty}$  and an orthonormal eigenbasis.
- (ii) Prove that  $(A + C)^{-1}$  is a compact operator for any constant  $C > -\mu_1(A)$ .

In this case we say that  $A$  has **compact resolvent**. As a consequence, the eigenfunctions of  $A$  form an orthonormal basis.

**Exercise.** Let  $A$  be a self-adjoint operator on a Hilbert space  $\mathcal{H}$ . Assume that  $A$  is bounded from below and let  $\mu_n(A)$  be its min-max values. Prove that for all  $N \in \mathbb{N}$ ,

$$\sum_{n=1}^N \mu_n(A) = \inf \left\{ \sum_{n=1}^N \langle u_n, Au_n \rangle : \{u_n\}_{n=1}^{\infty} \text{ an orthonormal family in } \mathcal{H} \right\}.$$

## 2.4 Sobolev inequalities

Next, we turn to the fact that the Schrödinger operators are defined on the real space  $\mathbb{R}^{dN}$ . Therefore, we recall some standard results from real analysis.

**Definition** (Sobolev Spaces). For any dimension  $d \in \mathbb{N}$  and  $s > 0$  (not necessarily an integer), define

$$H^s(\mathbb{R}^d) := \left\{ f \in L^2(\mathbb{R}^d) \mid |k|^s \widehat{f}(k) \in L^2(\mathbb{R}^d) \right\}$$

with  $\widehat{f}$  the Fourier transform of  $f$ . This is a Hilbert space with the inner product

$$\langle f, g \rangle_{H^s} = \int_{\mathbb{R}^d} \overline{\widehat{f}(k)} \widehat{g}(k) (1 + |2\pi k|^{2s}) dk$$

Remarks:

- We use the following convention of the Fourier transform

$$\widehat{f}(k) = \int_{\mathbb{R}^d} e^{-2\pi i k \cdot x} f(x) dx.$$

In this “engineering convention” we have the inverse formula

$$f(x) = \int_{\mathbb{R}^d} e^{2\pi i k \cdot x} \widehat{f}(k) dk$$

and the the Plancherel theorem

$$\|f\|_{L^2(\mathbb{R}^d)} = \|\widehat{f}\|_{L^2(\mathbb{R}^d)}.$$

- On the Sobolev space  $H^s(\mathbb{R}^d)$ , the weak derivative is defined via the Fourier transform

$$\widehat{D^\alpha f}(k) = (2\pi i k)^\alpha \widehat{f}(k) \in L^2(\mathbb{R}^d)$$

for any multiple index  $\alpha = (\alpha_1, \dots, \alpha_d)$  with  $|\alpha| = \alpha_1 + \dots + \alpha_d \leq s$ .

- In the course we will mostly think of  $s$  as an integer for simplicity. The non-integer case (the so-called **fractional Sobolev spaces**) is useful for studying relativistic quantum mechanics.

**Theorem** (Sobolev Inequalities/Continuous embedding). *Let  $d \geq 1$  and  $s > 0$ . Then*

$$\|f\|_{L^p(\mathbb{R}^d)} \leq C \|f\|_{H^s(\mathbb{R}^d)}, \quad \forall f \in H^s(\mathbb{R}^d)$$

where

$$\begin{cases} 2 \leq p \leq \frac{2d}{d-2s}, & \text{if } s < d/2, \\ 2 \leq p < \infty, & \text{if } s = d/2 \\ 2 \leq p \leq \infty, & \text{if } s > d/2. \end{cases}$$

We say that  $H^s(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$  with continuous embedding. When  $s > d/2$  we also have the continuous embedding  $H^s(\mathbb{R}^d) \subset \mathcal{C}(\mathbb{R}^d)$  (the space of continuous functions with sup-norm).

Remarks:

- In the case  $s < d/2$ , the power  $p^* := 2d/(d - 2s)$  is called the **Sobolev critical exponent**. In fact, this is the **only power** works for the following standard Sobolev inequality

$$\|f\|_{L^{p^*}(\mathbb{R}^d)} \leq C \|(-\Delta)^{s/2} f\|_{L^2(\mathbb{R}^d)}$$

(on the right side we do not put the full norm of  $H^s$ , but only the seminorm of  $\dot{H}^s$ ).

- In principle, for any given power  $s > 0$ , the Sobolev inequality becomes **weaker in higher dimensions**. For example,

$$H^1(\mathbb{R}) \subset L^2(\mathbb{R}) \cap \mathcal{C}(\mathbb{R}), \quad H^1(\mathbb{R}^2) \subset \bigcap_{2 \leq p < \infty} L^p(\mathbb{R}^2), \quad H^1(\mathbb{R}^3) \subset \bigcap_{2 \leq p \leq 6} L^p(\mathbb{R}^2).$$

Similarly,

$$H^2(\mathbb{R}^3) \subset L^2(\mathbb{R}^3) \cap \mathcal{C}(\mathbb{R}^3) \quad \text{but} \quad H^2(\mathbb{R}^4) \not\subset \mathcal{C}(\mathbb{R}^4).$$

*Proof.* Let us prove the standard Sobolev inequality in the case  $s < d/2$ :

$$\|f\|_{L^{p^*}(\mathbb{R}^d)} \leq C \|(-\Delta)^{s/2} f\|_{L^2(\mathbb{R}^d)}.$$

We use Rumin's method, which will be useful later. By Plancherel and Fubini theorems we can write

$$\begin{aligned} K &:= \|(-\Delta)^{s/2} f\|_{L^2(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^3} dk |2\pi k|^{2s} |\widehat{u}(k)|^2 \\ &= \int_{\mathbb{R}^3} dk \int_0^\infty dE \mathbf{1}(|2\pi k|^{2s} > E) |\widehat{u}(k)|^2 \\ &= \int_0^\infty dE \int_{\mathbb{R}^3} dk \mathbf{1}(|2\pi k|^{2s} > E) |\widehat{u}(k)|^2 \\ &= \int_0^\infty dE \int_{\mathbb{R}^3} dx |u^{E+}(x)|^2 = \int_{\mathbb{R}^3} dx \int_0^\infty dE |u^{E+}(x)|^2 \end{aligned}$$

where the function  $u^{E+}$  is defined via the Fourier transform

$$\widehat{u}^{E+}(k) = \mathbf{1}(|2\pi k|^{2s} > E) \widehat{u}(k).$$

When  $d > 2s$ , we have the uniform bound

$$|u(x) - u^{E+}(x)| = \left| \int_{\mathbb{R}^d} dk e^{i2\pi k \cdot x} \widehat{u}(k) \mathbf{1}(|2\pi k|^{2s} \leq E) \right|$$

$$\begin{aligned} &\leq \left( \int_{\mathbb{R}^d} |2\pi k|^{2s} |\widehat{u}(k)|^2 \right)^{1/2} \left( \int_{\mathbb{R}^d} dk \mathbf{1}(|2\pi k|^{2s} \leq E) |2\pi k|^{-2s} \right)^{1/2} \\ &= C_0 K^{\frac{1}{2}} E^{\frac{d-2s}{4s}} \end{aligned}$$

with a constant  $C_0$  depending only on  $d$  and  $s$ . By the triangle inequality,

$$|u^{E^+}(x)| \geq \left| |u(x)| - |u(x) - u^{E^+}(x)| \right| \geq \left[ |u(x)| - |u(x) - u^{E^+}(x)| \right]_+ \geq \left[ |u(x)| - C_0 K^{\frac{1}{2}} E^{\frac{d-2s}{4s}} \right]_+.$$

Of course this bound is nontrivial only when

$$E \leq \left( \frac{|u(x)|}{C_0 K^{\frac{1}{2}}} \right)^{\frac{4s}{d-2s}}.$$

Integrating over  $E \in (0, \infty)$  we get

$$\begin{aligned} \int_0^\infty dE |u^{E^+}(x)|^2 &\geq \int_0^\infty dE \left[ |u(x)| - C_0 K^{\frac{1}{2}} E^{\frac{d}{2s}-1} \right]_+^2 \\ &= C_1 |u(x)|^2 \left( \frac{|u(x)|}{K^{\frac{1}{2}}} \right)^{\frac{4s}{d-2s}} = C_1 |u(x)|^{\frac{2d}{d-2s}} K^{-\frac{2s}{d-2s}}. \end{aligned}$$

In conclusion,

$$K \geq C_1 K^{-\frac{2s}{d-2s}} \int_{\mathbb{R}^d} |u(x)|^{\frac{2d}{d-2s}}$$

which is equivalent to

$$K^{\frac{d}{d-2s}} \geq C_1 \int_{\mathbb{R}^d} |u(x)|^{\frac{2d}{d-2s}}.$$

Inserting the definition  $K = \|(-\Delta)^{s/2} f\|_{L^2(\mathbb{R}^d)}^2$  we arrive at the desired inequality.  $\square$

**Theorem** (Sobolev compact embedding). *Let  $d \geq 1$  and  $s > 0$ . Then for any bounded set  $\Omega \subset \mathbb{R}^d$ , the operator  $\mathbf{1}_\Omega : H^s(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)$  is a compact operator, where*

$$\begin{cases} 2 \leq p < \frac{2d}{d-2s}, & \text{if } s \leq d/2, \\ 2 \leq p \leq \infty, & \text{if } s > d/2. \end{cases}$$

*When  $s > d/2$ , we also have the compact embedding  $\mathbf{1}_\Omega : H^s(\mathbb{R}^d) \rightarrow \mathcal{C}(\mathbb{R}^d)$ .*

Remark: The Sobolev compact embedding means that if  $u_n \rightharpoonup u$  weakly in  $H^s(\mathbb{R}^d)$  with

$s > 0$ , then for any bounded set  $\Omega \subset \mathbb{R}^d$ ,

$$\mathbf{1}_\Omega u_n \rightarrow \mathbf{1}_\Omega u \quad \text{strongly in } L^p(\mathbb{R}^d).$$

The key point is the strong convergence in  $L^2$ ; and the convergence in  $L^p$  follows by a standard interpolation (for which we have to avoid the end-point). This kind of result can be interpreted as the operator  $\mathbf{1}_\Omega(1 - \Delta)^{-s/2}$  is compact on  $L^2(\mathbb{R}^d)$ .

We have the following more general result.

**Theorem.** *Let  $f, g \in L^\infty(\mathbb{R}^d)$  and  $f(x) \rightarrow 0$  and  $g(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . Then  $f(x)g(-i\nabla)$  is a compact operator on  $L^2(\mathbb{R}^d)$ .*

Here  $f(x)$  is the usual multiplication operator and  $g(-i\nabla)$  is defined by the spectral theorem, or equivalently via the Fourier transform:

$$(g(-i\nabla)u)(k) = g(2\pi k)\widehat{u}(k).$$

This theorem can be interpreted in the same spirit of the uncertainty principle: localizing both position and momentum gives us a compact operator. Further estimates for the operator  $f(x)g(-i\nabla)$  will be discussed in the next chapter.

*Proof.* We prove that if  $u_n \rightharpoonup 0$  weakly in  $L^2(\mathbb{R}^d)$ , then

$$v_n(x) = f(x)(g(-i\nabla)u_n)(x) \rightarrow 0 \quad \text{strongly in } L^2(\mathbb{R}^d).$$

**Step 1.** Let us consider the case when  $f$  and  $g$  are compactly supported. We write

$$(g(-i\nabla)u_n)(x) = \int_{\mathbb{R}^d} e^{i2\pi k \cdot x} g(2\pi k)\widehat{u}_n(k)dk.$$

Since  $u_n \rightharpoonup 0$  weakly in  $L^2(\mathbb{R}^d)$ ,  $\widehat{u}_n \rightharpoonup 0$  weakly in  $L^2(\mathbb{R}^d)$ . Since  $g$  is bounded and compactly supported,  $e^{i2\pi k \cdot x} g(2\pi k) \in L^2(\mathbb{R}^d, dk)$ . Thus

$$(g(-i\nabla)u_n)(x) \rightarrow 0, \quad \text{for a.e. } x \in \mathbb{R}^d.$$

Moreover, by Hölder inequality we also know that  $g(-i\nabla)u_n$  is bounded in  $L^\infty(\mathbb{R}^d)$ . Since  $f$  is bounded and compactly supported, we find that

$$v_n(x) = f(x)(g(-i\nabla)u_n)(x) \rightarrow 0, \quad \text{for a.e. } x \in \mathbb{R}^d$$

and

$$\|v_n\|_{L^\infty(\mathbb{R}^d)} \leq C, \quad \text{supp}(v_n) \subset \text{supp } f.$$

Thus  $v_n \rightarrow 0$  strongly in  $L^2(\mathbb{R}^d)$  by Lebesgue Dominated Convergence Theorem.

**Step 2.** Now we consider the case when  $g$  is compactly supported and  $f(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . Then for any  $\varepsilon > 0$ , we split

$$f = f_\varepsilon + \tilde{f}_\varepsilon$$

with  $f_\varepsilon$  being compactly supported and  $\|\tilde{f}_\varepsilon\|_{L^\infty} \leq \varepsilon$ . By the triangle inequality

$$\begin{aligned} \|f(x)g(-i\nabla)u_n\|_{L^2} &\leq \|f_\varepsilon(x)g(-i\nabla)u_n\|_{L^2} + \|\tilde{f}_\varepsilon g(-i\nabla)u_n\|_{L^2} \\ &\leq \|f_\varepsilon(x)g(-i\nabla)u_n\|_{L^2} + \|\tilde{f}_\varepsilon\|_{L^\infty} \|g\|_{L^\infty} \|u_n\|_{L^2} \\ &\leq \|f_\varepsilon(x)g(-i\nabla)u_n\|_{L^2} + C\varepsilon \end{aligned}$$

for a constant  $C$  independent of  $\varepsilon$  and  $n$ . Here we used the fact that  $u_n$  is bounded in  $L^2$  since it converges weakly in  $L^2$ . By Step 1,

$$\lim_{n \rightarrow \infty} \|f_\varepsilon(x)g(-i\nabla)u_n\|_{L^2} = 0.$$

Thus

$$\limsup_{n \rightarrow \infty} \|f(x)g(-i\nabla)u_n\|_{L^2} \leq \limsup_{n \rightarrow \infty} \|f_\varepsilon(x)g(-i\nabla)u_n\|_{L^2} + C\varepsilon = C\varepsilon$$

Since this holds for any  $\varepsilon > 0$ , we conclude that  $\|f(x)g(-i\nabla)u_n\|_{L^2} \rightarrow 0$ .

**Step 3.** Now we consider the case when  $f(x), g(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . Then for any  $\varepsilon > 0$ , we split

$$g = g_\varepsilon + \tilde{g}_\varepsilon$$

with  $g_\varepsilon$  being compactly supported and  $\|\tilde{g}_\varepsilon\|_{L^\infty} \leq \varepsilon$ . By triangle inequality

$$\begin{aligned} \|f(x)g(-i\nabla)u_n\|_{L^2} &\leq \|f(x)g_\varepsilon(-i\nabla)u_n\|_{L^2} + \|f(x)\tilde{g}_\varepsilon(-i\nabla)u_n\|_{L^2} \\ &\leq \|f(x)g_\varepsilon(-i\nabla)u_n\|_{L^2} + \|f\|_{L^\infty} \|\tilde{g}_\varepsilon\|_{L^\infty} \|u_n\|_{L^2} \\ &\leq \|f(x)g_\varepsilon(-i\nabla)u_n\|_{L^2} + C\varepsilon. \end{aligned}$$

Taking  $n \rightarrow \infty$ , and then  $\varepsilon \rightarrow 0$ , we find that  $\|f(x)g(-i\nabla)u_n\|_{L^2} \rightarrow 0$ .

□



**Exercise.** Let  $f \in L^\infty(\mathbb{R}^d)$  such that  $f(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ . Prove that if  $u_n \rightarrow 0$  weakly in the Sobolev space  $H^s(\mathbb{R}^d)$  for some constant  $s > 0$ , then  $fu_n \rightarrow 0$  strongly in  $L^2(\mathbb{R}^d)$ .

**Exercise.** Let  $F, G : \mathbb{R}^d \rightarrow \mathbb{R}$  be locally bounded functions satisfying  $F(x), G(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ . Prove that the operator  $F(x) + G(-i\nabla)$  on  $L^2(\mathbb{R}^d)$  has compact resolvent.

## 2.5 Schrödinger operators

Now we are ready to discuss some basic spectral properties of the Schrödinger operator  $-\Delta + V(x)$  on  $L^2(\mathbb{R}^d)$ . We will always assume that  $V$  is a real-valued potential.

First, consider the case when  $V$  is bounded or vanishing at infinity.

**Theorem.** Assume that  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty > p, q > \max(2, d/2)$ , then  $-\Delta + V$  is a self adjoint operator on  $L^2(\mathbb{R}^d)$  with domain  $H^2(\mathbb{R}^d)$  and

$$\sigma_{\text{ess}}(-\Delta + V) = [0, \infty).$$

More generally, the self-adjointness still holds if  $\infty \geq p, q \geq \max(2, d/2)$  when  $d \neq 4$  and  $\infty \geq p, q > 2$  when  $d = 4$ ; and the essential spectrum property still holds if  $\infty > p \geq 2$  when  $d \leq 3$  and  $\infty > p > d/2$  when  $d \geq 4$ .

Remark:  $L^s(\mathbb{R}^d) \subset L^{s_1}(\mathbb{R}^d) + L^{s_2}(\mathbb{R}^d)$  if  $s_1 < s < s_2$ .

*Proof. Step 1.* First we prove the self-adjointness. We use the Kato-Rellich theorem and show that  $V$  is  $\Delta$ -relatively bounded, more precisely

$$\|Vu\|_{L^2} \leq \epsilon \|\Delta u\|_{L^2} + C_\epsilon \|u\|_{L^2}, \quad \forall \epsilon > 0, \quad \forall u \in L^2(\mathbb{R}^d).$$

We can always write

$$V = V_1 + V_2, \quad \|V_1\|_{L^p} \leq \epsilon, \quad \|V_2\|_{L^\infty} \leq C_\epsilon$$

with  $p = \max(2, d/2)$  if  $d \neq 4$  and  $p \in (2, \infty)$  arbitrarily if  $d = 4$ .

**Case  $d \leq 3$ :** Using the Sobolev embedding  $H^2(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$  we can bound

$$\|V_1 u\|_{L^2} \leq \|V_1\|_{L^2} \|u\|_{L^\infty} \leq C\varepsilon \|u\|_{H^2} \leq C\varepsilon (\|\Delta u\|_{L^2} + \|u\|_{L^2}).$$

Moreover,

$$\|V_2 u\|_{L^2} \leq \|V_2\|_{L^\infty} \|u\|_{L^2} \leq C_\varepsilon \|u\|_{L^2}.$$

Therefore, by the triangle inequality

$$\|Vu\|_{L^2} \leq \|V_1 u\|_{L^2} + \|V_2 u\|_{L^2} \leq C\varepsilon \|\Delta u\|_{L^2} + C_\varepsilon \|u\|_{L^2}, \quad \forall \varepsilon > 0, \quad \forall u \in L^2(\mathbb{R}^d).$$

This is equivalent to the desired estimate (we can change  $C\varepsilon \mapsto \epsilon$ ).

**Case  $d > 4$ :** Using the Sobolev embedding  $H^2(\mathbb{R}^d) \subset L^{\frac{2d}{d-4}}(\mathbb{R}^d)$  and Hölder's inequality we can bound

$$\|V_1 u\|_{L^2}^2 = \int |V_1|^2 |u|^2 \leq \left( \int |V_1|^{2r} \right)^{\frac{1}{r}} \left( \int |u|^{2r'} \right)^{\frac{1}{r'}} \leq C \|V_1\|_{L^{2r}}^2 \|u\|_{H^2}^2$$

where

$$\frac{1}{r} + \frac{1}{r'} = 1, \quad 2r' = \frac{2d}{d-4}.$$

We find that  $r' = d/(d-4)$  and  $r = d/4$ . Thus

$$\|V_1 u\|_{L^2} \leq C \|V_1\|_{L^{d/2}} \|u\|_{H^2} \leq C\varepsilon (\|\Delta u\|_{L^2} + \|u\|_{L^2}).$$

The rest is similar to the case  $d \leq 3$ .

**Case  $d = 4$ :** Using the Sobolev embedding  $H^2(\mathbb{R}^d) \subset L^s(\mathbb{R}^d)$  for any  $s \in (2, \infty)$ , we have

$$\|V_1 u\|_{L^2} = \left( \int |V_1|^2 |u|^2 \right)^{\frac{1}{2}} \leq \left( \int |V_1|^{2r} \right)^{\frac{1}{2r}} \left( \int |u|^{2r'} \right)^{\frac{1}{2r'}} \leq C \|V_1\|_{L^{2r}} \|u\|_{H^2}$$

with  $2r = p$  and  $1/r + 1/r' = 1$ . The rest is similar to the case  $d \leq 3$ .

**Step 2.** Now we turn to the essential spectrum. By Weyl's criterion, we need to show that  $V$  is  $\Delta$ -relatively compact, namely  $V(x)(1 - \Delta)^{-1}$  is a compact operator on  $L^2(\mathbb{R}^d)$ . It suffices to consider the case  $V \in L^p(\mathbb{R}^d)$  with  $\infty > p \geq 2$  if  $d \leq 3$  and  $\infty > p > d/2$  if  $d \geq 4$ .

Take  $u_n \rightharpoonup 0$  weakly in  $L^2$ , then we have to prove that  $V(1 - \Delta)^{-1}u_n \rightarrow 0$  strongly in  $L^2$ .

Note that  $u_n \rightharpoonup 0$  weakly in  $L^2$  implies that (**why?**)

$$f_n := (1 - \Delta)^{-1}u_n \rightharpoonup 0 \quad \text{weakly in } H^2(\mathbb{R}^d).$$

By Sobolev compact embedding, we know that for any bounded set  $\Omega \subset \mathbb{R}^d$ , then

$$\mathbb{1}_\Omega f_n \rightarrow 0 \quad \text{strongly in } L^s(\mathbb{R}^d)$$

for any  $s \leq \infty$  if  $d \leq 3$  and  $s < 2d/(d-4)$  if  $d \geq 4$ . In particular, given the condition  $\infty > p \geq 2$  if  $d \leq 3$  and  $\infty > p > d/2$  if  $d \geq 4$ , we can choose  $s$  such that

$$\frac{2}{p} + \frac{2}{s} = 1.$$

Moreover, we know that  $f_n$  is bounded in  $L^s(\mathbb{R}^d)$  due to Sobolev continuous embedding.

Let us split

$$V = V_1 + V_2, \quad V_1 = V\mathbb{1}_{\{|x| \leq R\}}, \quad V_2 = V\mathbb{1}_{\{|x| > R\}}.$$

Then

$$\|V_1 f_n\|_2 = \left( \int_{|x| \leq R} |V|^2 |f_n|^2 \right)^{1/2} \leq \left( \int |V|^p \right)^{\frac{1}{p}} \left( \int_{|x| \leq R} |f_n|^s \right)^{\frac{1}{s}} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

and

$$\|V_2 f_n\|_2 \leq \left( \int_{|x| > R} |V|^p \right)^{\frac{1}{p}} \left( \int |f_n|^s \right)^{\frac{1}{s}} \leq C \left( \int_{|x| > R} |V|^p \right)^{\frac{1}{p}} \rightarrow 0 \quad \text{as } R \rightarrow \infty.$$

Thus  $V f_n \rightarrow 0$  strongly in  $L^2(\mathbb{R}^d)$  when  $n \rightarrow \infty$ . □

The condition on  $V$  can be relaxed slightly if we use Friedrich's extension.

**Exercise.** Assume that  $V_+ \in L^1_{\text{loc}}(\mathbb{R}^d)$  and  $V_- \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty \geq p, q \geq \max(1, d/2)$  when  $d \neq 2$  and  $\infty \geq p, q > 1$  when  $d = 2$ . Prove that:

- For all  $u \in H^1(\mathbb{R}^d)$  and  $\varepsilon > 0$ ,

$$\int_{\mathbb{R}^d} |V_-| |u|^2 \leq \varepsilon \int_{\mathbb{R}^d} |\nabla u|^2 + C_\varepsilon \int_{\mathbb{R}^d} |u|^2.$$

- The operator  $-\Delta + V$  with core domain  $C_c^\infty(\mathbb{R}^d)$  is bounded from below on  $L^2(\mathbb{R}^d)$ . Hence, it can be extended to be a self-adjoint operator by Friedrichs' method.
- If we assume further that  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty > p, q \geq \max(1, d/2)$  when  $d \neq 2$ , and  $\infty > p, q > 1$  when  $d = 2$ , then the quadratic form domain of  $-\Delta + V$  is  $H^1(\mathbb{R}^d)$  and

$$\sigma_{\text{ess}}(-\Delta + V) = [0, \infty).$$

*Hint: For the essential spectrum, you can use Weyl's criterion.*

In particular for  $d = 3$ , the Friedrich extension covers the case  $V(x) = -|x|^{-s}$  with  $0 < s < 2$ , while the Kato-Rellich theorem requires  $0 < s < 3/2$ . In the critical case  $s = 2$ , we have **Hardy's inequality**

$$\int_{\mathbb{R}^3} |\nabla u(x)|^2 dx \geq \frac{1}{4} \int_{\mathbb{R}^3} \frac{|u(x)|^2}{|x|^2} dx, \quad \forall u \in H^1(\mathbb{R}^3).$$

Next, we show that if the potential  $V(x)$  is negative and decays slowly, then  $-\Delta + V$  has infinitely many negative eigenvalues. The opposite regime, when  $-\Delta + V$  has finitely many eigenvalues, will be studied in the next chapter.

**Theorem.** Assume that  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty > p, q > \max(1, d/2)$  and

$$V(x) \leq -|x|^{-s}, \quad 0 < s < 2, \quad \text{for } |x| \text{ large.}$$

Then  $-\Delta + V$  has infinitely many negative eigenvalues. Here  $-\Delta + V$  is the self-adjoint operator obtained by Friedrichs' extension.

*Proof.* By the min-max principle, it suffices to show that all the min-max values are negative:

$$\mu_n := \inf_{\dim M=n} \max_{\substack{u \in M \\ \|u\|_{L^2}=1}} \langle u, (-\Delta + V)u \rangle < 0.$$

To choose  $M$ , we take normalized functions  $u_i \in \mathcal{C}_c^\infty(i+1 > |x| > i)$ ,  $i = 1, 2, \dots, n$ . Then take  $R > 0$  and define

$$M := \text{Span}(u_i^{(R)}, i = 1, 2, \dots, n), \quad u_i^{(R)}(x) := R^{-\frac{d}{2}} u_i(x/R).$$

Clearly  $\dim M = n$  since  $\{u_i^{(R)}\}_{i=1}^n$  is an orthonormal family. Indeed, they have disjoint supports and in their supports

$$(n+1)R \geq |x| \geq R.$$

Using  $V(x) \leq -|x|^{-s}$  with  $0 < s < 2$  for  $|x|$  large, we find that for  $R$  large,

$$\begin{aligned} \langle u_i^{(R)}, (-\Delta + V)u_i^{(R)} \rangle &= \int |\nabla u_i^{(R)}|^2 + \int V|u_i^{(R)}|^2 \leq \int |\nabla u_i^{(R)}|^2 - \int |x|^{-s}|u_i^{(R)}(x)|^2 dx \\ &= R^{-2} \int |\nabla u_i|^2 - R^{-s} \int |x|^{-s}|u_i(x)|^2 dx < 0. \end{aligned}$$

Next, since  $\{u_i^{(R)}\}_{i=1}^n$  have disjoint support, we find that

$$\max_{\substack{u \in M \\ \|u\|_{L^2} = 1}} \langle u, (-\Delta + V)u \rangle = \max_{i=1, \dots, n} \langle u_i^{(R)}, (-\Delta + V)u_i^{(R)} \rangle < 0.$$

This completes the proof. □

Next, we consider the case when the potential  $V(x)$  grows to  $\infty$  at infinity.

**Theorem.** *Assume that  $V \in L^p_{\text{loc}}(\mathbb{R}^d)$  with  $p \geq \max(1, d/2)$  when  $d \neq 2$  and  $p > 1$  when  $d = 1$ ; moreover  $V(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ . Then the operator  $-\Delta + V$  with core domain  $C_c^\infty(\mathbb{R}^d)$  is bounded from below and can be extended to be a self-adjoint operator by Friedrich's extension. This self-adjoint operator has compact resolvent.*

*Proof.* The condition  $V \in L^p_{\text{loc}}(\mathbb{R}^d)$  and  $V(x) \rightarrow \infty$  implies that for the negative part,  $V_- \in L^p(\mathbb{R}^d)$ . Therefore,

$$-\Delta + V \geq \frac{1}{2}(-\Delta + V_+) - C.$$

Thus  $-\Delta + V$  is bounded from below and hence it can be extended to be a self-adjoint operator by Friedrich's extension.

It remains to prove that  $-\Delta + V$  has compact resolvent. By the min-max principle, it suffices to show that  $-\Delta + V_+$  has compact resolvent. We prove by contradiction: assume that a finite element  $\lambda \in \sigma_{\text{ess}}(-\Delta + V_+)$ . Then by Weyl's criterion, there exists an orthonormal family  $\{u_n\}_{n=1}^\infty$  in  $L^2(\mathbb{R}^d)$  such that  $\|(-\Delta + V_+ - \lambda)u_n\|_{L^2} \rightarrow 0$ . Consequently,

$$\int |\nabla u_n|^2 + \int V_+|u_n|^2 \rightarrow \lambda.$$

Since  $u_n$  is bounded in  $H^1(\mathbb{R}^d)$  and  $u_n \rightharpoonup 0$  in  $L^2(\mathbb{R}^d)$ , we have  $u_n \rightharpoonup 0$  in  $H^1(\mathbb{R}^d)$  (**why?**). Hence, by Sobolev's embedding theorem, for any  $R > 0$  we have

$$\lim_{n \rightarrow \infty} \|\mathbf{1}_{\{|x| \leq R\}} u_n\|_{L^2} = 0.$$

Consequently,

$$\lim_{n \rightarrow \infty} \|\mathbf{1}_{\{|x| > R\}} u_n\|_{L^2} = 1.$$

Therefore,

$$\lambda \geq \lim_{n \rightarrow \infty} \int V_+ |u_n|^2 \geq \lim_{n \rightarrow \infty} \left[ \inf_{|y| \geq R} V(y) \right] \int_{|x| \geq R} |u_n(x)|^2 dx = \inf_{|y| \geq R} V(y).$$

Then sending  $R \rightarrow \infty$  we obtain  $\lambda = \infty$ , which is a contradiction. □

# Chapter 3

## Semiclassical estimates

### 3.1 Cwikel-Lieb-Rozenblum inequality

By semiclassical approximation, the number of negative eigenvalues of the Schrödinger operator  $-\Delta + V(x)$  on  $L^2(\mathbb{R}^d)$  can be related to its phase-space analogue

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}(|2\pi k|^2 + V(x) < 0) dk dx = \frac{|B_1|}{(2\pi)^d} \int_{\mathbb{R}^d} |V(x)_-|^{d/2} dx.$$

Recall  $t_- = \min(t, 0)$ . The following bound justifies this relation as a universal upper bound.

**Theorem** (Cwikel-Lieb-Rozenblum (CLR) inequality). *If  $d \geq 3$  and  $V_- \in L^{\frac{d}{2}}(\mathbb{R}^d)$ , then*

$$\mathcal{N}(-\Delta + V) \leq C_d \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}}.$$

*Here  $\mathcal{N}(-\Delta + V)$  is the number of negative eigenvalues of  $-\Delta + V(x)$  on  $L^2(\mathbb{R}^d)$ . The constant  $C_d$  is finite and independent of  $V$ .*

Remarks:

- Here the condition  $V_- \in L^{\frac{d}{2}}(\mathbb{R}^d)$  ensures that  $-\Delta + V$  is bounded from below, and hence it can be extended to be a self-adjoint operator on  $L^2(\mathbb{R}^d)$  by Friedrichs' method.
- It is not surprising that the positive part of  $V$  does not appear in the upper bound because  $\mathcal{N}(-\Delta + V) \leq \mathcal{N}(-\Delta + V_-)$  by the min-max principle. In general, we do not any serious condition on  $V_+$ , for example  $V_+ \in L^1_{\text{loc}}(\mathbb{R}^d)$  is sufficient for the operator to

be well-defined in the core domain  $C_c^\infty(\mathbb{R}^d)$ . In relation to the Laplacian on bounded domains (we will discuss later), it would be also useful to think of a hard sphere potential where  $V = +\infty$  outside  $\Omega \in \mathbb{R}^d$  (in this case the underlying Hilbert space will be  $L^2(\Omega)$ ).

The following proof is due to Frank, based on Rumin method. This is an extension of the previous proof of Sobolev's inequality to orthogonal functions.

*Proof.* Assume  $\mathcal{N}(-\Delta + V) \geq N$ . Let  $W$  be the space spanned by eigenfunctions of negative eigenvalues of  $-\Delta + V$ . Since  $\dim W \geq N$  and  $\sqrt{-\Delta}$  has a trivial kernel, we get

$$\dim(\sqrt{-\Delta}W) \geq N.$$

Thus we can choose  $N$  orthonormal functions in  $\sqrt{-\Delta}W$ , says  $\sqrt{-\Delta}u_i$ . Thus

$$\{u_i\}_{i=1}^N \subset W, \quad \langle u_i, -\Delta u_j \rangle = \langle \sqrt{-\Delta}u_i, \sqrt{-\Delta}u_j \rangle = \delta_{ij}.$$

Therefore,

$$0 \geq \sum_{i=1}^N \langle u_i, (-\Delta + V)u_i \rangle = N + \int V \rho \geq N - \int |V_-| \rho, \quad \rho(x) := \sum_{i=1}^N |u_i(x)|^2.$$

For any  $E > 0$ , we introduce the function  $u_i^{E+}$  via the Fourier transform

$$\widehat{u_i^{E+}}(k) = \mathbf{1}(|2\pi k|^2 > E) \widehat{u_i}(k).$$

Then similarly to the proof of Sobolev's inequality, we can write

$$N = \sum_{i=1}^N \int_{\mathbb{R}^d} |\nabla u_i|^2 dx = \sum_{i=1}^N \int_0^\infty dE \int_{\mathbb{R}^d} dx |u_i^{E+}(x)|^2 = \int_{\mathbb{R}^d} dx \int_0^\infty dE \sum_{i=1}^N |u_i^{E+}(x)|^2.$$

By the triangle inequality for the Euclidean norm in  $\mathbb{C}^d$ , we have

$$\sqrt{\sum_{i=1}^N |u_i^{E+}(x)|^2} \geq \left| \sqrt{\sum_{i=1}^N |u_i(x)|^2} - \sqrt{\sum_{i=1}^N |u_i(x) - u_i^{E+}(x)|^2} \right|$$



and hence

$$\sum_{i=1}^N |u_i^{E^+}(x)|^2 \geq \left[ \sqrt{\rho(x)} - \sqrt{\sum_{i=1}^N |u_i(x) - u_i^{E^+}(x)|^2} \right]_+^2.$$

On the other hand, for  $d \geq 3$  we have the uniform bound

$$\begin{aligned} \sum_{i=1}^N |u_i(x) - u_i^{E^+}(x)|^2 &= \sum_{i=1}^N \left| \int_{\mathbb{R}^d} dk \widehat{u}_i(k) \mathbf{1}_{\{|2\pi k|^2 \leq E\}} e^{2\pi i k x} \right|^2 = \\ &= \sum_{i=1}^N \left| \int_{\mathbb{R}^d} dk |2\pi k| \widehat{u}_i(k) \frac{\mathbf{1}_{\{|2\pi k|^2 \leq E\}}}{|2\pi k|} e^{2\pi i k x} \right|^2 \stackrel{\text{Bessel}}{\leq} \\ &\leq \int_{\mathbb{R}^d} \frac{\mathbf{1}_{\{|2\pi k|^2 \leq E\}}}{|2\pi k|^2} dk \leq C_d E^{\frac{d-2}{2}}. \end{aligned}$$

Here we have used Bessel's inequality and the fact that  $\{|2\pi k| \widehat{u}_i(k)\}_{i=1}^N$  is an orthonormal family in  $L^2(\mathbb{R}^d, dk)$  (as  $\{\sqrt{-\Delta} u_i\}_{i=1}^N$  is an orthonormal family). Thus

$$N = \sum_{i=1}^N \int |\nabla u_i|^2 \geq \int_{\mathbb{R}^d} \int_0^\infty \left[ \sqrt{\rho(x)} - C_d E^{\frac{d-2}{4}} \right]_+^2 dE dx \geq \frac{1}{C_d} \int \rho(x)^{\frac{d}{d-2}}.$$

Therefore we conclude that

$$N \leq \int_{\mathbb{R}^d} |V_-| \rho \leq \|V_-\|_{L^{d/2}} \|\rho\|_{L^{\frac{d}{d-2}}} \leq \|V_-\|_{L^{d/2}} (C_d N)^{\frac{d-2}{d}}$$

which implies  $N \leq C_d \|V_-\|_{L^{d/2}}^{d/2}$ . □

The following exercise shows that the CLR bound fails if  $d \leq 2$ .

**Exercise.** Let  $d = 1, 2$ . Let  $V \in L^1(\mathbb{R}^d)$  if  $d = 1$  and  $V \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$  for some  $p > 1$  if  $d = 2$ . Prove that if

$$\int_{\mathbb{R}^d} V < 0$$

Then  $-\Delta + V$  has at least one negative eigenvalue. Hint: You may consider  $u_\varepsilon(x) = e^{-\varepsilon|x|}$  when  $d = 1$ , and  $u_\varepsilon(x) = e^{-(1+|x|)^\varepsilon}$  when  $d = 2$ .

Nevertheless, we have a modified result for  $d \leq 2$ . Recall that for any function  $F : [0, \infty) \rightarrow [0, \infty]$  we define the **Legendre transform**  $F^* : [0, \infty) \rightarrow [0, \infty]$  by

$$F^*(x) = \sup_{y \geq 0} \{xy - F(y)\}, \quad \forall x \geq 0$$

Note that  $F \geq G$ , then  $F^* \leq G^*$ . Moreover, if  $F(x) = x^p/p$  then  $F^*(x) = x^q/q$  with  $1/p + 1/q = 1$ , thanks to Young's inequality.

**Theorem** (Bound states in one and two dimensions). *For  $d = 1, 2$  and any  $L > 0$ ,*

$$\mathcal{N}(-\Delta + V + L) \leq \int_{\mathbb{R}^d} dx F^* \left( \frac{2|V(x)_-|}{L} \right) L^{d/2}.$$

where  $\mathcal{N}(-\Delta + V + L)$  is the number of negative eigenvalues of  $-\Delta + V + L$  on  $L^2(\mathbb{R}^d)$  and

$$F(t) = \begin{cases} t(e^{4\pi t} - 1), & \text{if } d = 2, \\ t \tan^2(t) + \infty \mathbf{1}(t \geq \pi/2), & \text{if } d = 1. \end{cases}$$

*Proof.* We proceed similarly to the proof of the CLR bound, except  $-\Delta$  will be replaced by  $-\Delta + L$ . More precisely, assume  $\mathcal{N}(-\Delta + V + L) \geq N$ . Let  $W$  be the space spanned by eigenfunctions of negative eigenvalues of  $-\Delta + V + L$ . Since  $\dim W \geq N$ , we have  $\dim(\sqrt{-\Delta + L}W) \geq N$ . Thus we can choose  $N$  functions  $\{u_i\}_{i=1}^N$  such that

$$\{u_i\}_{i=1}^N \subset W, \quad \langle u_i, (-\Delta + L)u_j \rangle = \langle \sqrt{-\Delta + L}u_i, \sqrt{-\Delta + L}u_j \rangle = \delta_{ij}.$$

Therefore,

$$0 \geq \sum_{i=1}^N \langle u_i, (-\Delta + V + L)u_i \rangle = N + \int V\rho \geq N - \int |V_-|\rho, \quad \rho(x) := \sum_{i=1}^N |u_i(x)|^2.$$

For any  $E > 0$ , we introduce  $\widehat{u}_i^{E+}(k) = \mathbf{1}(|2\pi k|^2 + L > E)\widehat{u}_i(k)$ . Then

$$\begin{aligned} N &= \sum_{i=1}^N \|\sqrt{-\Delta + L}u_i\|_{L^2}^2 = \int_{\mathbb{R}^d} dx \int_0^\infty dE \sum_{i=1}^N |u_i^{E+}(x)|^2 \\ &\geq \int_{\mathbb{R}^d} dx \int_0^\infty dE \left[ \sqrt{\rho(x)} - \sqrt{\sum_{i=1}^N |u_i(x) - u_i^{E+}(x)|^2} \right]_+^2. \end{aligned}$$

Since  $\{\widehat{u}_i(k)\sqrt{|2\pi k|^2 + L}\}_{i=1}^N$  is an orthonormal family,

$$\sum_{i=1}^N |u_i(x) - u_i^{E+}(x)|^2 = \sum_{i=1}^N \left| \int_{\mathbb{R}^d} dk \widehat{u}_i(k) \mathbf{1}_{\{|2\pi k|^2 + L \leq E\}} e^{2\pi i k x} \right|^2 =$$

$$= \sum_{i=1}^N \left| \int_{\mathbb{R}^d} dk \widehat{u}_i(k) \sqrt{|2\pi k|^2 + L} \frac{\mathbb{1}_{\{|2\pi k|^2 + L \leq E\}}}{\sqrt{|2\pi k|^2 + L}} e^{2\pi i k x} \right|^2 \stackrel{\text{Bessel}}{\leq} \int_{\mathbb{R}^d} \frac{\mathbb{1}_{\{|2\pi k|^2 + L \leq E\}}}{|2\pi k|^2 + L} dk.$$

**When  $d = 2$ :** it is straightforward to see that when  $E \geq L$ ,

$$\int_{\mathbb{R}^2} \frac{\mathbb{1}_{\{|2\pi k|^2 + L \leq E\}}}{|2\pi k|^2 + L} dk = \frac{1}{2\pi} \int_0^{\sqrt{E-L}} \frac{r}{r^2 + L} dr = \frac{1}{4\pi} \log(E/L).$$

Thus

$$N \geq \int_{\mathbb{R}^2} dx \int_L^\infty dE \left[ \sqrt{\rho(x)} - \sqrt{\frac{1}{4\pi} \log(E/L)} \right]_+^2 = L \int_{\mathbb{R}^2} dx F(\rho(x))$$

with  $F(t) = t(e^{4\pi t} - 1)$ . Therefore

$$\begin{aligned} 0 &\geq 2N - 2 \int_{\mathbb{R}^2} |V_-| \rho \geq N + L \int_{\mathbb{R}^2} dx F(\rho(x)) - 2 \int_{\mathbb{R}^2} dx |V_-(x)| \rho(x) \\ &= N + L \int_{\mathbb{R}^2} dx \left( F(\rho) - \frac{2}{L} |V_-| \rho \right) \\ &\geq N - L \int_{\mathbb{R}^2} F^* \left( \frac{2|V_-|}{L} \right). \end{aligned}$$

Thus

$$N \leq L \int_{\mathbb{R}^2} F^* \left( \frac{2|V_-|}{L} \right).$$

**When  $d = 1$ :** if  $E \geq L$ ,

$$\int_{\mathbb{R}} \frac{\mathbb{1}_{\{|2\pi k|^2 + L \leq E\}}}{|2\pi k|^2 + L} dk = \frac{1}{\pi} \int_0^{\sqrt{E-L}} \frac{1}{r^2 + L} dr = \frac{1}{\pi\sqrt{L}} \arctan\left(\sqrt{\frac{E}{L} - 1}\right).$$

Thus

$$\begin{aligned} N &\geq \int_{\mathbb{R}} dx \int_L^\infty dE \left[ \sqrt{\rho(x)} - \sqrt{\frac{1}{\pi\sqrt{L}} \arctan\left(\sqrt{\frac{E}{L} - 1}\right)} \right]_+^2 \\ &\geq L \int_{\mathbb{R}} dx \rho(x) \tan^2(\sqrt{L}\rho(x)) = \sqrt{L} \int_{\mathbb{R}} dx F(\sqrt{L}\rho(x)) \end{aligned}$$

where  $F(t) = t \tan^2(t)$  if  $t < \pi/2$  and  $\infty$  if  $t \geq \pi/2$ . Therefore

$$\begin{aligned} 0 &\geq 2N - 2 \int_{\mathbb{R}} |V_-| \rho \geq N + \sqrt{L} \int_{\mathbb{R}} dx F(\sqrt{L}\rho(x)) - 2 \int_{\mathbb{R}^2} dx |V_-(x)| \rho(x) \\ &= N + \sqrt{L} \int_{\mathbb{R}} dx \left( F(\sqrt{L}\rho) - \frac{2|V_-|}{L} \sqrt{L}\rho \right) \\ &\geq N - \sqrt{L} \int_{\mathbb{R}} F^* \left( \frac{2|V_-|}{L} \right). \end{aligned}$$

Thus

$$N \leq \sqrt{L} \int_{\mathbb{R}} F^* \left( \frac{2|V_-|}{L} \right).$$

□

## 3.2 Lieb–Thirring inequality

Instead of focusing on the number of negative eigenvalues of  $-\Delta + V$ , we are also interested in the sum of negative eigenvalues (which is related to the ground state energy of the ideal Fermi gas). We have the following generalization of the CLR bound.

**Theorem** (Lieb–Thirring Inequality). *Let  $d \geq 1$  and*

$$\begin{cases} s \geq 1/2, & \text{if } d = 1 \\ s > 0, & \text{if } d = 2, \\ s \geq 0, & \text{if } d \geq 3. \end{cases}$$

*If the operator  $-\Delta + V(x)$  on  $L^2(\mathbb{R}^d)$  has negative eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$ , then*

$$\sum_{n \geq 1} |\mu_n|^s \leq L_{s,d} \int_{\mathbb{R}^d} |V_-|^{s + \frac{d}{2}}.$$

*Here the constant  $L_{s,d}$  is finite and independent of  $V$ . The range of  $s$  is optimal.*

Remark:

- This inequality was first derived by **Lieb–Thirring** in 1975 for  $s = 1$  in their proof of **the stability of matter**. Then they extended the inequality to the cases  $s > 0$  when  $d \geq 2$  and  $s > 1/2$  when  $d = 1$ . The case  $s = 0$  when  $d \geq 3$  was proved independently

by **Cwikel, Lieb, Rozenblum** (CLR bound) around 1977. The case  $s = 1/2$  when  $d = 1$  was solved by **Weidl** in 1996.

- Assuming that  $-\Delta + V$  can be defined as a self-adjoint operator, it is convenient to write the sum of moments of eigenvalues as

$$\sum_{n \geq 1} |\mu_n(-\Delta + V)|^s = \text{Tr} |(-\Delta + V)_-|^s$$

where the negative part  $(-\Delta + V)_-$  is defined by the spectral projection.

- Similarly to the CLR bound, in the Lieb-Thirring inequality only the negative part  $V_-$  is relevant. This follows from the fact that  $-\Delta + V \geq -\Delta + V_-$ , leading to  $\text{Tr} |(-\Delta + V)_-|^s \leq \text{Tr} |(-\Delta + V_-)_-|^s$  by the min-max principle.
- This bound agrees to the semiclassical approximation

$$\text{Tr} |(-\Delta + V)_-|^s \approx \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |(|2\pi k|^2 + V(x))_-|^s dk dx = L_{s,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{s+d/2} dx.$$

where

$$L_{s,d}^{\text{cl}} = \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(s+1)}{\Gamma(s+1 + \frac{d}{2})}.$$

When  $s \geq 3/2$ , the best constant in the Lieb-Thirring inequality coincides to the classical constant  $L_{s,d} = L_{s,d}^{\text{cl}}$ . This was proved by **Lieb-Thirring** (1976) for  $d = 1$  and extended by **Laptev-Weidl** (2000) to all  $d \geq 1$ .

- It is known that if  $s < 1$ , then  $L_{s,d} > L_{s,d}^{\text{cl}}$ . When  $d = 1$  and  $s = 1/2$ , **Hundertmark-Lieb-Thomas** (1998) proved that the sharp constant is  $L_{1/2,1} = 2L_{1/2,1}^{\text{cl}}$ .
- In the most interesting case  $s = 1$ , it is conjectured that  $L_{1,d} = L_{1,d}^{\text{cl}}$  for  $d \geq 3$  (we will come back to this case). When  $d \leq 2$ , the conjectured value of  $L_{s,d}$  is given by a one-bound-state/Sobolev inequality. Currently, the best known result is  $L_{1,d} \leq 1.456L_{1,d}^{\text{cl}}$  for all  $d \geq 1$ ; see FHJN (2018).

*Proof.* We will use the bound on the number of negative eigenvalues derived in the previous section. The proof covers all cases except the critical case  $d = 1$  and  $s = 1/2$ .

Our starting point is **the layer-cake representation**: for any  $s > 0$

$$\text{Tr} |(-\Delta + V)_-|^s = s \int_0^\infty \mathcal{N}(-\Delta + V + E) E^{s-1} dE$$

where  $\mathcal{N}(-\Delta + V + E)$  is the number of negative eigenvalue of  $-\Delta + V + E$  (which is the same to the number of eigenvalues  $\leq -E$  of  $-\Delta + V$ ). This formula is analogous to

$$\int_{\Omega} |f(x)|^s d\mu(x) = s \int_0^{\infty} \mu(\{x : |f(x)| > E\}) E^{s-1} dE.$$

**Case  $d \geq 3$ :** Inserting the CLR bound

$$\mathcal{N}(-\Delta + V + E) \leq C_d \int_{\mathbb{R}^d} |(V(x) + E)_-|^{\frac{d}{2}} dx$$

to the above layer-cake representation and using Fubini's theorem we obtain for all  $s > 0$

$$\begin{aligned} \text{Tr} |(-\Delta + V_-)_-|^s &= s \int_0^{\infty} \mathcal{N}(-\Delta + V + E) E^{s-1} dE \\ &\leq s C_d \int_0^{\infty} dE \int_{\mathbb{R}^d} dx |(V(x) + E)_-|^{\frac{d}{2}} E^{s-1} \\ &= s C_d \int_{\mathbb{R}^d} dx \int_0^{\infty} dE |(V(x) + E)_-|^{\frac{d}{2}} E^{s-1} = C_{s,d} \int_{\mathbb{R}^d} dx |V(x)_-|^{s+\frac{d}{2}}. \end{aligned}$$

**Case  $d \leq 2$ :** Recall that

$$\mathcal{N}(-\Delta + V + E) = \mathcal{N}(-\Delta + (V + E/2) + E/2) \leq \int_{\mathbb{R}^2} dx F^* \left( \frac{2|(V(x) + E/2)_-|}{(E/2)} \right) \left( \frac{E}{2} \right)^{\frac{d}{2}}$$

where  $F^*$  is the Legendre transform of the function

$$F(t) = \begin{cases} t(e^{4\pi t} - 1), & \text{if } d = 2, \\ t \tan^2(t) + \infty \mathbf{1}(t \geq \pi/2), & \text{if } d = 1. \end{cases}$$

Hence,

$$\begin{aligned} \text{Tr} |(-\Delta + V_-)_-|^s &= s \int_0^{\infty} \mathcal{N}(-\Delta + V + E) E^{s-1} dE \\ &\leq s \int_0^{\infty} dE \int_{\mathbb{R}^d} dx F^* \left( \frac{2|(V(x) + E/2)_-|}{(E/2)} \right) \left( \frac{E}{2} \right)^{\frac{d}{2}} E^{s-1} \\ &= s C_d \int_{\mathbb{R}^d} dx \int_0^{2|V(x)_-|} dE F^* \left( \frac{2|(V(x) + E/2)_-|}{(E/2)} \right) \left( \frac{E}{2} \right)^{\frac{d}{2}} E^{s-1} \\ &= C_{s,d} \int_{\mathbb{R}^d} dx |V(x)_-|^{s+\frac{d}{2}} \int_0^1 dy F^* \left( \frac{2}{y} - 2 \right) y^{s-1+\frac{d}{2}}. \end{aligned}$$

In the last equality we have changed the variable  $E = 2|V(x)_-|y$ . It remains to show that

$$\int_0^1 dy F^* \left( \frac{2}{y} - 2 \right) y^{s-1+\frac{d}{2}} < \infty.$$

**Consider the case  $d = 2$ ,  $s > 0$ .** Recall that if  $F \geq G$ , then  $F^* \leq G^*$ . Using

$$F(t) = t(e^{4\pi t} - 1) \geq C_p t^p, \quad \forall 2 \leq p < \infty$$

we find that

$$F^*(t) \leq C_q t^q, \quad \forall 1 < q \leq 2.$$

Since  $s > 0$ , we can take  $1 < q < 1 + s$  (such that  $s - q > -1$ ), and hence

$$\int_0^1 dy F^* \left( \frac{2}{y} - 2 \right) y^s \leq C_q \int_0^1 dy \left( \frac{2}{y} - 2 \right)^q y^s \leq C_q \int_0^1 dy y^{s-q} < \infty.$$

**Consider the case  $d = 1$ ,  $s > 1/2$ .** Using  $F(t) \geq 0$  for all  $t \geq 0$  and

$$F(t) = \infty, \quad \forall t \geq 2/\pi$$

we have

$$F^*(y) = \sup_{t \geq 0} (ty - F(t)) \leq \sup_{0 \leq t \leq 2/\pi} (ty - F(t)) \leq \frac{2}{\pi} y.$$

Therefore, using  $s > 1/2$  we have

$$\int_0^1 dy F^* \left( \frac{2}{y} - 2 \right) y^{s-\frac{1}{2}} \leq \int_0^1 dy \frac{2}{\pi} \left( \frac{2}{y} - 2 \right) y^{s-\frac{1}{2}} \leq \frac{4}{\pi} \int_0^1 dy y^{s-\frac{3}{2}} < \infty.$$

□

In the special case  $s = 1$  (sum of negative eigenvalues), the Lieb-Thirring inequality is equivalent to a kinetic inequality for **orthonormal functions** in  $L^2(\mathbb{R}^d)$  which will be useful to study large fermionic quantum systems.

**Theorem** (Lieb-Thirring Kinetic Inequality). *Let  $d \geq 1$ . Let  $\{u_n\}_n$  be an orthonormal family in  $L^2(\mathbb{R}^d)$  and define the **density**  $\rho(x) = \sum_n |u_n(x)|^2$ . Then*

$$\sum_n \int_{\mathbb{R}^d} |\nabla u_n(x)|^2 dx \geq K_d \int_{\mathbb{R}^d} \rho(x)^{1+\frac{2}{d}} dx.$$

Moreover, the best constant  $K_d > 0$  in the kinetic inequality is related to the best constant  $L_{1,d}$  for the sum of negative eigenvalues of  $-\Delta + V$  as

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1+\frac{2}{d}} = 1.$$

Remark: The **Lieb-Thirring conjecture** on the optimal value  $L_{1,d}$  is equivalent to

$$K_d = K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}} \quad \text{when } d \geq 3.$$

Here  $|B_1|$  is the volume of the unit ball in  $\mathbb{R}^d$ .

*Proof.* The kinetic inequality can be proved directly using Rumin's method (see an exercise). It remains to prove the relation between the best constants  $K_d$  and  $L_{1,d}$ . This follows a standard duality argument and Young's inequality (c.f. Legendre transform)

$$\frac{a^p}{p} = \sup_{b \geq 0} \left( ab - \frac{b^q}{q} \right), \quad \forall a \geq 0, \quad \forall p, q \geq 1, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Assume that the operator  $-\Delta + V$  has negative eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  with eigenfunctions  $u_1, u_2, \dots$ . By the LT kinetic inequality we have

$$\sum_n \int_{\mathbb{R}^d} |\nabla u_n|^2 \geq K_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}}, \quad \rho(x) = \sum_n |u_n(x)|^2.$$

Therefore,

$$\begin{aligned} \sum_n \mu_n &= \sum_n \langle u_n, (-\Delta + V)u_n \rangle = \sum_n \int_{\mathbb{R}^d} |\nabla u_n|^2 + \int_{\mathbb{R}^d} V \rho \\ &\geq K_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}} - \int_{\mathbb{R}^d} |V_-| \rho. \end{aligned}$$

Using Young's inequality we find that

$$K_d \rho(x)^{1+\frac{2}{d}} - |V(x)_-| \rho(x) \geq -\tilde{L}_{1,d} |V(x)_-|^{1+d/2}, \quad \forall x \in \mathbb{R}^d$$

where

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) \tilde{L}_{1,d} \right)^{1+\frac{2}{d}} = 1.$$



Thus

$$\sum_n \mu_n \geq \int_{\mathbb{R}^d} \left( K_d \rho^{1+\frac{2}{d}} - |V_-| \rho \right) \geq -\tilde{L}_{1,d} \int_{\mathbb{R}^d} |V_-|^{1+d/2}.$$

Consequently,

$$\sum_n |\mu_n| \leq \tilde{L}_{1,d} \int_{\mathbb{R}^d} |V_-|^{1+d/2}$$

and hence  $\tilde{L}_{1,d} \geq L_{1,d}$  (as  $L_{1,d}$  is the best constant). By the choice of  $\tilde{L}_{1,d}$ , we get

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1+\frac{2}{d}} \leq 1.$$

Reversely, consider any orthonormal family  $\{u_n\}_n$  in  $L^2(\mathbb{R}^d)$  and denote  $\rho(x) = \sum_n |u_n(x)|^2$ . Since Young's inequality is sharp, we can choose  $V(x) = -c_0 \rho(x)^{2/d}$  with an appropriate constant  $c_0 > 0$  such that

$$\tilde{K}_d \rho^{1+\frac{2}{d}} - |V_-| \rho = -L_{1,d} |V_-|^{1+d/2}$$

where

$$\left( \left(1 + \frac{2}{d}\right) \tilde{K}_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1+\frac{2}{d}} = 1.$$

On the other hand, by the LT inequality for the sum of negative eigenvalues of  $-\Delta + V$  and the min-max principle,

$$\sum_n \int_{\mathbb{R}^d} |\nabla u_n|^2 + \int_{\mathbb{R}^d} V \rho = \sum_n \langle u_n, (-\Delta + V) u_n \rangle \geq \text{Tr}(-\Delta + V)_- \geq -L_{1,d} \int_{\mathbb{R}^d} |V_-|^{1+d/2}.$$

Therefore, thanks to the choice of  $V$ ,

$$\sum_n \int_{\mathbb{R}^d} |\nabla u_n|^2 \geq - \int_{\mathbb{R}^d} V \rho - L_{1,d} \int_{\mathbb{R}^d} |V_-|^{1+d/2} = \tilde{K}_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}}.$$

Thus  $\tilde{K}_d \leq K_d$  since  $K_d$  is the best constant in the LT kinetic inequality. By the choice of  $\tilde{K}_d$ , we get

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1+\frac{2}{d}} \geq 1.$$

In conclusion, the best constants  $K_d$  and  $L_{1,d}$  satisfy

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1 + \frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1 + \frac{2}{d}} = 1.$$

□

**Exercise.** Let  $d \geq 1$ . Let  $\{u_n\}_{n=1}^N \subset H^1(\mathbb{R}^d)$  be an orthonormal family in  $L^2(\mathbb{R}^d)$  and define  $\rho(x) = \sum_{n=1}^N |u_n(x)|^2$ . Use Rumin's method to prove that

$$\sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla u_n(x)|^2 dx \geq K_d \int_{\mathbb{R}^d} \rho(x)^{1 + \frac{2}{d}} dx.$$

Here the constant  $K_d > 0$  depends only on  $d$ .

### 3.3 Birman-Schwinger Principle

In this section, we discuss an alternative approach to study the bound state problem for the Schrödinger operator  $-\Delta + V(x)$  on  $L^2(\mathbb{R}^d)$ . It is convenient to assume  $V \leq 0$  and denote  $U = -V \geq 0$ . Our starting point is the following reformulation of the eigenvalue problem.

**Theorem** (Birman-Schwinger principle). Let  $d \geq 1$ . Let  $0 \leq U \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty > p, q \geq \max(1, d/2)$  if  $d \neq 2$  and  $\infty > p, q > 1$  if  $d = 2$ . Recall that we can define  $-\Delta - U$  as a self-adjoint operator on  $L^2(\mathbb{R}^d)$  with the quadratic form domain  $H^1(\mathbb{R}^d)$  and  $\sigma_{\text{ess}}(-\Delta - U) = [0, \infty)$ . Then for all  $E > 0$ :

- (i)  $-E$  is an eigenvalue of  $-\Delta - U$  **if and only if** 1 is an eigenvalue of  $K_E = \sqrt{U(x)}(-\Delta + E)^{-1}\sqrt{U(x)}$  (with the same multiplicity).
- (ii) The number of eigenvalues  $\leq -E$  of  $-\Delta - U$  is equal to the number of eigenvalues  $\geq 1$  of  $K_E$ .

Moreover,  $K_E$  is a positive compact operator on  $L^2(\mathbb{R}^d)$ .

*Proof. Step 1.* We prove that if  $-E < 0$  is an eigenvalue of  $-\Delta - U(x)$ , then 1 is an

eigenvalue of  $K_E$ . Consider the eigenvalue equation

$$(-\Delta - U)f = -Ef, \quad \text{for some } 0 \neq f \in L^2(\mathbb{R}^d).$$

Then we can write

$$(-\Delta + E)f = Uf \implies f = (-\Delta + E)^{-1}Uf \implies \sqrt{U}f = \sqrt{U}(-\Delta + E)^{-1}\sqrt{U}(\sqrt{U}f).$$

Thus

$$\sqrt{U}f = K_E\sqrt{U}f.$$

To conclude that  $K_E$  has eigenvalue 1, we need to show that  $0 \neq \sqrt{U}f \in L^2(\mathbb{R}^d)$ . Since  $0 \neq f \in L^2(\mathbb{R}^d)$  and  $-\Delta + E \geq E > 0$ , we have

$$Uf = (-\Delta + E)f \neq 0.$$

Consequently,  $\sqrt{U}f \neq 0$ . Moreover, since  $f$  is an eigenfunction of  $-\Delta - U$ , it must belong to the quadratic form domain  $H^1(\mathbb{R}^d)$ . Hence,

$$\|\sqrt{U}f\|_{L^2}^2 = \int_{\mathbb{R}^d} U|f|^2 \leq C\|f\|_{H^1}^2 < \infty.$$

Thus if  $-E < 0$  is an eigenvalue of  $-\Delta - U(x)$ , then 1 is an eigenvalue of  $K_E$ .

**Step 2.** Reversely, we prove that if 1 is an eigenvalue of  $K_E$ , then  $-E < 0$  is an eigenvalue of  $-\Delta - U(x)$ . Consider the eigenvalue equation

$$g = K_Eg = \sqrt{U}(-\Delta + E)^{-1}\sqrt{U}g, \quad 0 \neq g \in L^2(\mathbb{R}^d).$$

Define

$$f = (-\Delta + E)^{-1}\sqrt{U}g.$$

Then

$$(-\Delta + E)f = \sqrt{U}g = U(-\Delta + E)^{-1}\sqrt{U}g = Uf$$

which is equivalent to

$$(-\Delta - U)f = -Ef.$$

To conclude that  $-E$  is an eigenvalue of  $-\Delta - U$ , it remains to prove that  $0 \neq f \in L^2(\mathbb{R}^d)$ .

Note that

$$\sqrt{U}f = \sqrt{U}(-\Delta + E)^{-1}\sqrt{U}g = g \neq 0,$$

and hence  $f \neq 0$ . Moreover, by Sobolev's inequality,  $(-\Delta + E)^{-1/2}\sqrt{U}$  is a bounded operator on  $L^2(\mathbb{R}^d)$  (it is indeed a compact operator; see an exercise below). Therefore,

$$(-\Delta + E)^{-1/2}\sqrt{U}g \in L^2(\mathbb{R}^d) \implies f = (-\Delta + E)^{-1}\sqrt{U}g \in H^1(\mathbb{R}^d).$$

From the above proof and the one-to-one correspondence between eigenfunctions  $f \leftrightarrow g$  (i.e.  $g = \sqrt{U}f$  and  $f = (-\Delta + E)^{-1}\sqrt{U}g$ ), we also obtain that the multiplicity of eigenvalue  $-E$  of  $-\Delta - U$  is the same with the multiplicity of the eigenvalue 1 of  $K_E$ .

**Step 3.** We can write  $K_E = BB^*$  with  $B = \sqrt{U(x)}(-\Delta + E)^{-1/2}$ . Since  $B$  is a compact operator on  $L^2(\mathbb{R}^d)$  (see an exercise below), we conclude that  $K_E$  is a positive compact operator on  $L^2(\mathbb{R}^d)$ . Therefore, by Spectral theorem, it has eigenvalues

$$\lambda_1(E) \geq \lambda_2(E) \geq \dots, \quad \lim_{n \rightarrow \infty} \lambda_n(E) = 0.$$

Note that for every  $n \in \mathbb{N}$ ,  $E \mapsto \lambda_n(E)$  is decreasing and continuous. To see the monotonicity, note that  $E \mapsto K_E$  is operator monotone: if  $E \geq E' > 0$ , then

$$(-\Delta + E)^{-1} \leq (-\Delta + E')^{-1} \implies K_E = \sqrt{U}(-\Delta + E)^{-1}\sqrt{U} \leq \sqrt{U}(-\Delta + E')^{-1}\sqrt{U} = K_{E'}.$$

Hence,  $\lambda_n(E) \leq \lambda_n(E')$  by the min-max principle (in fact,  $-\lambda_n(E)$  is the n-th min-max value of  $-K_E$ ). On the other hand, if  $E \geq E' > 0$ , then we also have

$$E(-\Delta + E)^{-1} \geq E'(-\Delta + E')^{-1} \implies EK_E \geq E'K_{E'}.$$

Thus  $E\lambda_n(E) \geq E'\lambda_n(E')$  by the min-max principle again. Combining with  $\lambda_n(E') \geq \lambda_n(E)$ , we conclude that  $E \mapsto \lambda_n(E)$  is also continuous.

By a standard counting argument combining the one-two-one correspondence in Steps 1&2 and the monotonicity/continuity of  $\lambda_n(E)$ , we find that the number of eigenvalues  $\leq E$  of  $-\Delta - U(x)$  is the same with the number of eigenvalues  $\geq 1$  of  $K_E$ . It is easiest to see via a Figure. □

**Exercise.** Let  $d \geq 1$ . Let  $0 \leq U \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $\infty > p, q \geq \max(1, d/2)$  if  $d \neq 2$  and  $\infty > p, q > 1$  if  $d = 2$ . Prove that  $\sqrt{U(x)}(-\Delta + E)^{-1/2}$  is a compact operator on  $L^2(\mathbb{R}^d)$ .

Historically, the Birman-Schwinger principle was used by Lieb and Thirring to prove their inequality

$$\mathrm{Tr} |(-\Delta + V)_-|^s \leq L_{s,d} \int_{\mathbb{R}^d} |V_-|^{s+\frac{d}{2}},$$

first for  $s = 1, d = 3$ , and then for  $s > 0$  when  $d \geq 2$  and  $s > 1/2$  when  $d = 1$  (it does not work for the critical cases  $s = 0, d \geq 3$  and  $s = 1/2, d = 1$ ). In the following, to illustration of the usefulness of this approach, we will represent

- The original proof of Lieb and Thirring in the physically interesting case  $s = 1, d = 3$ ;
- A proof of the existence of bound states with any negative potential in  $d \leq 2$ .

**The original proof of the Lieb-Thirring inequality for  $s = 1$  and  $d = 3$ :**

$$\mathrm{Tr} |(-\Delta + V)_-| \leq L_{1,3} \int_{\mathbb{R}^3} |V_-|^{\frac{5}{2}}.$$

It suffices to consider the case  $U = -V \geq 0$ . By the layer-cake representation:

$$\mathrm{Tr} |(-\Delta - U)_-| = \int_0^\infty \mathcal{N}(-\Delta - U + E) dE$$

where  $\mathcal{N}(-\Delta - U + E)$  is the number of negative eigenvalue of  $-\Delta - U + E$  (which is the same to the number of eigenvalues  $\leq -E$  of  $-\Delta + V$ ). By the Birman-Schwinger principle,  $\mathcal{N}(-\Delta + V + E)$  is equal to the number of eigenvalues  $\geq 1$  of

$$K_E = \sqrt{U(x)}(-\Delta + E)^{-1}\sqrt{U(y)}.$$

Consequently, this number is bounded by the Hilbert-Schmidt norm of  $K_E$ . Note that  $K_E$  has the kernel

$$K_E(x, y) = \sqrt{U(x)}G_E(x - y)\sqrt{U(y)}, \quad \widehat{G}_E(k) = \frac{1}{|2\pi k|^2 + E}.$$

Thus

$$\mathcal{N}(-\Delta - U + E) \leq \|K_E\|_{\mathrm{HS}}^2 = \iint_{\mathbb{R}^3 \times \mathbb{R}^3} U(x)|G_E(x - y)|^2 U(y) dx dy.$$

Using the Cauchy-Schwarz inequality and Plancherel's theorem we have

$$\begin{aligned}
\mathcal{N}(-\Delta - U + E) &\leq \iint_{\mathbb{R}^3 \times \mathbb{R}^3} |G_E(x-y)|^2 \frac{U^2(x) + U^2(y)}{2} dx dy \\
&= \left( \int_{\mathbb{R}^3} |U|^2 \right) \left( \int_{\mathbb{R}^3} |G_E|^2 \right) = \left( \int_{\mathbb{R}^3} |U|^2 \right) \left( \int_{\mathbb{R}^3} \frac{1}{(|2\pi k|^2 + E)^2} dk \right) \\
&= \left( \int_{\mathbb{R}^3} |U|^2 \right) \left( \int_{\mathbb{R}^3} \frac{1}{(|2\pi k|^2 + 1)^2} dk \right) E^{-1/2} \leq CE^{-1/2} \int_{\mathbb{R}^3} |U|^2.
\end{aligned}$$

This bound is not good enough for inserting to the layer cake representation. But we can adjust it by shifting  $U \mapsto (U - E/2)_+$ :

$$\begin{aligned}
\mathcal{N}(-\Delta - U + E) &= \mathcal{N}(-\Delta - (U - E/2) + E/2) \\
&\leq \mathcal{N}(-\Delta - (U - E/2)_+ + E/2) \leq CE^{-1/2} \int_{\mathbb{R}^3} dx (U(x) - E/2)_+^2.
\end{aligned}$$

Thus we conclude

$$\begin{aligned}
\text{Tr} |(-\Delta - U)_-| &= \int_0^\infty dE \mathcal{N}(-\Delta - U + E) \\
&\leq C \int_0^\infty dE \int_{\mathbb{R}^3} dx (U(x) - E/2)_+^2 E^{-1/2} \\
&= C \int_{\mathbb{R}^3} dx \int_0^\infty dE (U(x) - E/2)_+^2 E^{-1/2} = C \int_{\mathbb{R}^3} U^{5/2}(x) dx.
\end{aligned}$$

□

**An alternative proof of the existence of bound states for  $0 \not\equiv V \leq 0$  and  $d \leq 2$ .** (We assume that  $V$  is regular enough).

Denote  $U = -V \geq 0$ . By the Birman-Schwinger principle,  $-E < 0$  is an eigenvalue of  $-\Delta - U$  if and only if 1 is an eigenvalue of

$$K_E = \sqrt{U(x)}(-\Delta + E)^{-1} \sqrt{U(x)}.$$

Since  $K_E$  is a non-negative compact operator on  $L^2(\mathbb{R}^d)$ , the norm operator

$$\lambda_1(E) = \|K_E\|$$

is its largest eigenvalue. Let us prove that there exists  $E > 0$  such that  $\lambda_1(E) = 1$ .

Recall that  $E \mapsto \lambda_1(E)$  is decreasing and continuous. Moreover, clearly

$$\lim_{E \rightarrow \infty} \lambda_1(E) = 0.$$

(We can show that  $\lim_{E \rightarrow \infty} \|K_E\|_{\text{HS}} = 0$ ). On the other hand, let us show that if  $d \leq 2$  and  $0 \leq U \not\equiv 0$ , then

$$\lim_{E \rightarrow 0} \lambda_1(E) = \infty.$$

Indeed, when  $E \rightarrow 0^+$  we have, for any normalized function  $\varphi \in L^2(\mathbb{R}^d)$ ,

$$\lambda_1(E) \geq \langle \varphi, K_E \varphi \rangle = \left\langle \varphi \sqrt{U}, (-\Delta + E)^{-1} \varphi \sqrt{U} \right\rangle = \int_{\mathbb{R}^d} \frac{|\widehat{\varphi \sqrt{U}}(k)|^2}{|2\pi k|^2 + E} dk \rightarrow \int_{\mathbb{R}^d} \frac{|\widehat{\varphi \sqrt{U}}(k)|^2}{|2\pi k|^2} dk.$$

Here we have used Lebesgue Monotone Convergence theorem. Note that when  $d \leq 2$ , the function  $|k|^{-2}$  is not integrable at  $0 \in \mathbb{R}^d$ . On the other hand, since  $0 \leq U \not\equiv 0$ , we have

$$\widehat{\varphi \sqrt{U}}(0) = \int_{\mathbb{R}^d} \varphi(x) \sqrt{U(x)} dx \neq 0$$

for an appropriate choice of  $\varphi$ . Moreover, note that  $k \mapsto \widehat{\varphi \sqrt{U}}(k)$  is continuous (we have  $\varphi \sqrt{U} \in L^1(\mathbb{R}^d)$  when  $\varphi \in L^2$  and  $U \in L^1$ ). Thus in summary,  $E \mapsto \lambda_1(E)$  is continuous and

$$\lim_{E \rightarrow \infty} \lambda_1(E) = 0, \quad \lim_{E \rightarrow 0} \lambda_1(E) = \infty$$

Thus there exists  $E > 0$  such that  $\lambda_1(E) = 1$ . Then  $-E$  is an eigenvalue of  $-\Delta - U$ .  $\square$

Remark: As said, the Birman-Schwinger principle alone is not enough to derive the CLR bound. When  $d = 3$ , the Yukawa potential is given explicitly:

$$\widehat{G}_E(k) = \frac{1}{|2\pi k|^2 + E} \implies G_E(x) = \frac{e^{-\sqrt{E}|x|}}{4\pi|x|}.$$

Hence, from the above analysis we find that

$$\mathcal{N}(-\Delta - U + E) \leq \|K_E\|_{\text{HS}}^2 = \iint_{\mathbb{R}^3 \times \mathbb{R}^3} U(x) |G_E(x-y)|^2 U(y) dx dy \leq \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{U(x)U(y)}{(4\pi)^2 |x-y|^2} dx dy$$

for all  $E > 0$ . Thus we have the Birman-Schwinger inequality for the number of all negative

eigenvalues:

$$\mathcal{N}(-\Delta - U) \leq \frac{1}{(4\pi)^2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{U(x)U(y)}{|x-y|^2} dx dy.$$

Recall the **Hardy–Littlewood–Sobolev inequality**: for  $p, q > 1$ ,  $0 < r < d$  satisfying

$$\frac{1}{p} + \frac{1}{q} + \frac{r}{d} = 2$$

we have

$$\left| \iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{f(x)g(y)}{|x-y|^r} dx dy \right| \leq C_{d,s} \|f\|_{L^p} \|g\|_{L^q}.$$

Thus we can estimate further

$$\mathcal{N}(-\Delta - U) \leq C \|U\|_{L^{3/2}(\mathbb{R}^3)}^2.$$

In contrast, the CLR bound says that  $\mathcal{N}(-\Delta - U) \leq C \|U\|_{L^{3/2}(\mathbb{R}^3)}^{3/2}$ .

### 3.4 Kato–Seiler–Simon and Cwikel’s inequalities

In this section, we take a closer look to the operator  $f(x)g(-i\nabla)$  and its connection to spectral estimates for Schrödinger operators.

We have proved that if  $f, g : \mathbb{R}^d \rightarrow \mathbb{C}$  are uniformly bounded and vanishing at infinity, then  $f(x)g(-i\nabla)$  is a compact operator on  $L^2(\mathbb{R}^d)$ . This property can be extended to Schatten spaces.

**Definition** (Schatten spaces). *Let  $1 \leq p < \infty$  and let  $\mathcal{H}$  be a Hilbert space. The Schatten space  $\mathfrak{S}_p(\mathcal{H})$  contains all bounded operators  $A : \mathcal{H} \rightarrow \mathcal{H}$  such that*

$$\|A\|_{\mathfrak{S}_p} = (\text{Tr}(|A|^p))^{\frac{1}{p}} < \infty, \quad |A| = \sqrt{A^*A}.$$

*We denote by  $\mathfrak{S}_\infty(\mathcal{H})$  the space of compact operators, with the operator norm  $\|A\|$ . Thus  $\mathfrak{S}_p(\mathcal{H}) \subset \mathfrak{S}_q(\mathcal{H})$  if  $p \leq q$ .*

Remarks:

- $\mathfrak{S}_1(\mathcal{H})$  is the space of **trace class operators**. For any  $A \in \mathfrak{S}_1(\mathcal{H})$  and any orthonor-



mal basic  $\{u_n\}_{n \geq 1}$  of  $\mathcal{H}$ , we have

$$\mathrm{Tr}(A) = \sum_{n \geq 1} \langle u_n, Au_n \rangle < \infty.$$

The value of  $\mathrm{Tr}(A)$  is independent of the choice of the basis  $\{u_n\}$ . Moreover, we have **the cyclicity of the trace**: if  $AB$  and  $BA$  are trace class, then

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

- $\mathfrak{S}_2(\mathcal{H})$  is the space of **Hilbert-Schmidt operators**. This is a Hilbert space with the inner product

$$\langle A, B \rangle_{\mathfrak{S}_2} = \mathrm{Tr}(A^*B).$$

When  $\mathcal{H} = L^2(\Omega, \mu)$ , any operator  $A \in \mathfrak{S}_2(\mathcal{H})$  has a kernel  $K_A \in L^2(\Omega \times \Omega, \mu \times \mu)$  such that

$$(Af)(x) = \int_{\Omega} K_A(x, y) f(y) d\mu(y), \quad \forall f \in L^2(\Omega).$$

The mapping  $A \mapsto K_A$  is a unitary operator from  $L^2(\Omega)$  to  $L^2(\Omega^2)$ , in particular:

$$\|A\|_{\mathfrak{S}_2} = \|K_A\|_{L^2}.$$

- In general, when  $A$  is a compact operator on  $\mathcal{H} = L^2(\Omega, \mu)$ , we have the spectral decomposition

$$A = \sum_{n \geq 1} \lambda_n |u_n\rangle \langle v_n|$$

with  $\{u_n\}, \{v_n\}$  orthonormal families in  $\mathcal{H}$  and  $\lambda_n \in \mathbb{R}$ ,  $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ . The kernel of  $A$  is

$$K_A(x, y) = \sum_{n \geq 1} \lambda_n u_n(x) \overline{v_n(y)}.$$

With this convention, the trace of  $A$  can be computed by the diagonal part of its kernel

$$\mathrm{Tr}(A) = \int_{\Omega} K_A(x, x) d\mu(x) = \sum_{n \geq 1} \lambda_n \langle v_n, u_n \rangle_{L^2}$$

(which is well-defined if  $A$  is trace class). From the spectral decomposition, we also obtain the **polar decomposition**  $A = U|A|$  with  $U$  a unitary operator on  $\mathcal{H}$ .

- The Schatten space  $\mathfrak{S}_p$  satisfies properties similar to  $L^p$  spaces. They are Banach

spaces. Moreover, if  $p, q \geq 1$  and  $1/p + 1/q = 1$ , then we have **Hölder inequality**

$$\|AB\|_{\mathfrak{S}_1} \leq \|A\|_{\mathfrak{S}_p} \|B\|_{\mathfrak{S}_q}.$$

More precisely,

$$\|A\|_{\mathfrak{S}_p} = \sup_{\|B\|_{\mathfrak{S}_q}=1} |\mathrm{Tr}(AB)| = \sup_{\|B\|_{\mathfrak{S}_q}=1} |\langle A, B \rangle_{\mathfrak{S}_2}|.$$

**Exercise.** Prove Hölder inequality: if  $p, q \geq 1$  and  $1/p + 1/q = 1$  then

$$\|AB\|_{\mathfrak{S}_1} \leq \|A\|_{\mathfrak{S}_p} \|B\|_{\mathfrak{S}_q}.$$

*Hint: You can use the spectral decomposition.*

Now let us come back to the operator  $f(x)g(-i\nabla)$ . A basic and very useful property is

**Theorem** (Kato–Seiler–Simon inequality). Let  $f, g \in L^p(\mathbb{R}^d)$  with  $2 \leq p < \infty$ . Then

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_p} \leq \|f\|_{L^p} \|g\|_{L^p}.$$

*Proof.* It suffices to consider the case when  $0 \leq f, g \in L^\infty(\mathbb{R}^d)$  with compact supports. When  $p = \infty$ , it is obvious that

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_\infty} \leq \|f\|_{L^\infty} \|g\|_{L^\infty}.$$

When  $p = 2$ , we have the exact equality

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_2} = \|f\|_{L^2} \|g\|_{L^2}.$$

In fact, the integral kernel of  $f(x)g(-i\nabla)$  is

$$K(x, y) = f(x)\check{g}(x - y)$$

where  $\check{g}$  is the inverse Fourier transform of  $g$ . Therefore

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_2}^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x)|^2 |\check{g}(x - y)|^2 dx dy = \|f\|_{L^2}^2 \|\check{g}\|_{L^2}^2 = \|f\|_{L^2}^2 \|g\|_{L^2}^2.$$

Here we have used the identity between the Hilbert-Schmidt norm of an operator with its kernel and Plancherel theorem.

The case  $2 < p < \infty$  can be handled by **complex interpolation**. We skip the details and refer to Simon’s book “Trace Ideals and Their Applications” (Theorem 4.1).  $\square$

A deeper result concerning  $f(x)g(-i\nabla)$  is **Cwikel’s inequality**. This is related to the **weak  $L^p$  norm** and **weak Schatten norm**.

**Definition.** For  $1 \leq p < \infty$ , the space  $L_w^p(\mathbb{R}^d)$  contains functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  such that

$$\|f\|_{L_w^p} := \sup_{\tau > 0} \left( \tau |\{x : |f(x)| > \tau\}|^{\frac{1}{p}} \right) < \infty.$$

Remarks:

- Clearly the weak- $L^p$  is smaller than the usual  $L^p$  norm:

$$\|f\|_{L^p}^p = \int_{\mathbb{R}^d} |f|^p \geq \sup_{\tau > 0} \int_{\{|f| > \tau\}} |f|^p \geq \tau^p |\{x : |f(x)| > \tau\}| = \|f\|_{L_w^p}^p.$$

- We know that  $|x|^{-1} \notin L^p(\mathbb{R}^d)$  for all  $p$ , but  $|x|^{-1} \in L_w^d(\mathbb{R}^d)$  since

$$\tau |\{x : |x|^{-1} > \tau\}|^{\frac{1}{d}} = \tau |\{x : |x| < \tau^{-1}\}|^{\frac{1}{d}} = |B_1|^{\frac{1}{d}}.$$

- The expression  $\|f\|_{L_w^p}$  define a **quasi-norm** because instead of the triangle inequality we only have

$$\|f + g\|_{L_w^p} \leq C(\|f\|_{L_w^p} + \|g\|_{L_w^p}).$$

Alternatively we can also define

$$f \mapsto \sup_{\Omega} |\Omega|^{-\frac{1}{p'}} \int_{\Omega} |f(x)| dx, \quad \frac{1}{p} + \frac{1}{p'} = 1$$

which is indeed a norm (the supremum is taken over all set  $\Omega$  with  $0 < |\Omega| < \infty$ ).

- The Hardy-Littlewood-Sobolev inequality is equivalent to a **weak Young inequality**

$$\left| \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x)g(x-y)h(y) dx dy \right| \leq C \|f\|_{L^p} \|g\|_{L_w^q} \|h\|_{L^r}, \quad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2.$$

By optimizing over  $h$  we can also write equivalently,

$$\|f * g\|_{L^{r'}} \leq C \|f\|_{L^p} \|g\|_{L_w^q}, \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r'}.$$

This is stronger than the usual Young inequality

$$\|f * g\|_{L^{r'}} \leq C \|f\|_{L^p} \|g\|_{L^q}, \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r'}.$$

**Definition.** Let  $\mathcal{H}$  be a Hilbert space. For  $1 \leq p < \infty$ , the space  $\mathfrak{S}_{p,w}(\mathbb{R}^d)$  contains compact operators  $A$  on  $\mathcal{H}$  such that

$$\|A\|_{\mathfrak{S}_{p,w}} := \sup_{\tau > 0} \left( \tau \left( \mathcal{N}(\tau - |A|) \right)^{\frac{1}{p}} \right) < \infty.$$

Here  $\mathcal{N}(\tau - |A|)$  is the number of eigenvalues  $> \tau$  of  $|A| = \sqrt{A^*A}$ .

Obviously, we have  $\|A\|_{\mathfrak{S}_{p,w}} \leq \|A\|_{\mathfrak{S}_p}$ . The following deep result is interesting in its own and will imply the CLR bound.

**Theorem** (Cwikel's theorem). If  $f \in L^p(\mathbb{R}^d)$  and  $g \in L_w^p(\mathbb{R}^d)$  with  $2 \leq p < \infty$ , then

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_{p,w}} \leq C_p \|f\|_{L^p} \|g\|_{L_w^p}.$$

**Proof of the CLR bound using Cwikel's theorem.**

Let  $d \geq 3$  and  $0 \leq U \in L^{d/2}(\mathbb{R}^d)$ . By the Birman-Schwinger principle, for any  $E > 0$ , the number of eigenvalues  $\leq -E$  of  $-\Delta - U(x)$  is equivalent to the number of eigenvalues  $\geq 1$  of

$$K_E = \sqrt{U(x)}(-\Delta + E)^{-1} \sqrt{U(x)} = \sqrt{U(x)}g_E^2(-i\nabla)\sqrt{U(x)}$$

with  $g_E(p) = (|p|^2 + E)^{-1/2}$ . Consequently,

$$\mathcal{N}(-\Delta - U + E) \leq \|K_E\|_{\mathfrak{S}_{d/2,w}}^{d/2} = \|\sqrt{U(x)}g_E(-i\nabla)\|_{\mathfrak{S}_{d,w}}^d.$$

Using Cwikel's theorem and the uniform bound  $g_E(p) \leq |p|^{-1}$  we find that

$$\|\sqrt{U(x)}g_E(-i\nabla)\|_{\mathfrak{S}_{d,w}} \leq C \|\sqrt{U}\|_{L^d} \|g_E\|_{L_w^d} \leq C \|U\|_{L^{d/2}}^{1/2}.$$

Thus

$$\mathcal{N}(-\Delta - U + E) \leq C\|U\|_{L^{d/2}}^{d/2}.$$

Since it holds for all  $E > 0$ , we conclude that  $\mathcal{N}(-\Delta - U) \leq C\|U\|_{L^{d/2}}^{d/2}$ .  $\square$

**Proof of Cwikel’s theorem.** Assume that  $f, g \geq 0$  and  $\|f\|_{L^p} = \|g\|_{L^p_w} = 1$ .

**Step 1.** We decompose

$$\begin{aligned} f &= \sum_{n \in \mathbb{Z}} f_n, & f_n(x) &:= f(x) \mathbb{1}(2^{n-1} < f(x) \leq 2^n), \\ g &= \sum_{n \in \mathbb{Z}} g_n, & g_n(x) &:= g(x) \mathbb{1}(2^{n-1} < g(x) \leq 2^n). \end{aligned}$$

Then

$$X := f(x)g(-i\nabla) = \sum_{m, n \in \mathbb{Z}} f_n(x)g_m(-i\nabla) = A_k + B_k$$

where

$$A_k = \sum_{m+n \leq k} f_n(x)g_m(-i\nabla), \quad B_k = \sum_{m+n > k} f_n(x)g_m(-i\nabla), \quad \forall k \in \mathbb{Z}.$$

By the Cauchy-Schwarz inequality,

$$|X|^2 = (A_k^* + B_k^*)(A_k + B_k) \leq 2(A_k^*A_k + B_k^*B_k) = 2(|A_k|^2 + |B_k|^2).$$

**Step 2.** We prove that

$$\|A_k\| \leq 2^{k+1}, \quad \forall k \in \mathbb{Z}.$$

For any normalized functions  $u, v \in L^2(\mathbb{R}^d)$ , by the Cauchy-Schwarz inequality we have

$$\begin{aligned} |\langle u, A_k v \rangle| &\leq \sum_{m+n \leq k} \|f_n u\| \|g_m \hat{v}\| = \sum_{\ell \leq k} \sum_{n \in \mathbb{Z}} \|f_n u\| \|g_{\ell-n} \hat{v}\| \\ &= \sum_{\ell \leq k} 2^\ell \sum_{n \in \mathbb{Z}} \|2^{-n} f_n u\| \|2^{n-\ell} g_{\ell-n} \hat{v}\| \\ &\leq \sum_{\ell \leq k} 2^\ell \left( \sum_{n \in \mathbb{Z}} \|2^{-n} f_n u\|^2 \right)^{1/2} \left( \sum_{n \in \mathbb{Z}} \|2^{n-\ell} g_{\ell-n} v\|^2 \right)^{1/2} \\ &\leq \sum_{\ell \leq k} 2^\ell = 2^{k+1}. \end{aligned}$$

Here we have used

$$\sum_{n \in \mathbb{Z}} (2^{-n} f_n)^2 \leq 1, \quad \sum_{m \in \mathbb{Z}} (2^{-m} g_m)^2 \leq 1.$$

which follow from the facts that the functions  $\{2^{-n} f_n\}_n$  are  $\leq 1$  and have disjoint supports, and the same for  $\{2^{-m} g_m\}_m$ .

**Step 3.** We prove that

$$\mathrm{Tr}(|B_k|^2) \leq C 2^{(2-p)k}, \quad \forall k \in \mathbb{Z}.$$

Indeed, it is straightforward to see that

$$\begin{aligned} \mathrm{Tr}(|B_k|^2) &= \mathrm{Tr}(B_k^* B_k) = \sum_{m+n>k} \sum_{m'+n'>k} \mathrm{Tr} \left( g_m(-i\nabla) f_n(x) f_{n'}(x) g_{m'}(-i\nabla) \right) \\ &= \sum_{m+n>k} \sum_{m'+n'>k} \mathrm{Tr} \left( f_n(x) f_{n'}(x) g_{m'}(-i\nabla) g_m(-i\nabla) \right) \\ &= \sum_{m+n>k} \mathrm{Tr} \left( |f_n(x)|^2 |g_m(-i\nabla)|^2 \right) = \sum_{m+n>k} \|f_n(x) g_m(-i\nabla)\|_{\mathbb{S}^2}^2 \\ &= \sum_{m+n>k} \|f_n\|_{L^2}^2 \|g_m\|_{L^2}^2 = \sum_{n \in \mathbb{Z}} \|f_n\|_{L^2}^2 \sum_{m>k-n} \|g_m\|_{L^2}^2. \end{aligned}$$

Here we have used the cyclicity of the trace and the fact that  $\{f_n\}_n$  have disjoint supports, and that  $\{g_m\}_m$  have disjoint supports. The  $L^2$ -norm of  $g_m$  can be controlled by the weak  $L^p$  norm  $\|g\|_{L_w^p} = 1$  as follows:

$$\begin{aligned} \|g_m\|_{L^2}^2 &= \int_{\mathbb{R}^d} |g(x)|^2 \mathbf{1}(2^{m-1} < g(x) \leq 2^m) dx \\ &\leq 2^{2m} |\{2^{m-1} < g(x) \leq 2^m\}| \\ &\leq 2^{2m-p(m-1)} \left( 2^{p(m-1)} |\{2^{m-1} < g(x)\}| \right) \\ &\leq 2^{2m-p(m-1)} \sup_{\tau>0} \left( \tau^p |\{\tau < g(x)\}| \right) \\ &= 2^{2m-p(m-1)} \|g\|_{L_w^p}^p = 2^{(2-p)m} 2^p. \end{aligned}$$

Combining with  $f_n \geq 2^{n-1}$  on its support and  $\|f\|_{L^p} = 1$  we have

$$\begin{aligned} \mathrm{Tr}(|B_k|^2) &= \sum_{n \in \mathbb{Z}} \|f_n\|_{L^2}^2 \sum_{m>k-n} \|g_m\|_{L^2}^2 \leq \sum_{n \in \mathbb{Z}} \|f_n\|_{L^2}^2 \sum_{m>k-n} 2^{(2-p)m} 2^p \\ &\leq C_p \sum_{n \in \mathbb{Z}} \|f_n\|_{L^2}^2 2^{(2-p)(k-n)} = C_p 2^{(2-p)k} \int_{\mathbb{R}^d} dx \sum_{n \in \mathbb{Z}} |f_n(x)|^2 2^{n(p-2)} \end{aligned}$$

$$\leq C_p 2^{(2-p)k} \int_{\mathbb{R}^d} dx \sum_{n \in \mathbb{Z}} |f_n(x)|^p = C_p 2^{(2-p)k} \int_{\mathbb{R}^d} dx |f(x)|^p = C_p 2^{(2-p)k}.$$

**Step 4.** In summary, we have proved that for every  $k \in \mathbb{Z}$ , we can split

$$|X|^2 \leq 2(|A_k|^2 + |B_k|^2) \leq 2(2^{2(k+1)} + |B_k|^2), \quad \text{Tr}(|B_k|^2) \leq C_p 2^{k(2-p)}.$$

Note that  $\varepsilon_k := 2 \times 2^{2(k+1)}$  varies from 0 to  $\infty$  when  $k$  runs from  $-\infty$  to  $\infty$ . Therefore, for every  $\varepsilon > 0$ , we can choose  $k \in \mathbb{Z}$  such that

$$\varepsilon_k \leq \varepsilon \leq \varepsilon_{k+1} = 4\varepsilon_k.$$

Since  $\varepsilon \sim 2^{2k}$ , the above splitting argument tells us

$$|X|^2 \leq \varepsilon + Y_\varepsilon \quad \text{with an operator } Y_\varepsilon \geq 0, \quad \text{Tr}(Y_\varepsilon) \leq C_p \varepsilon^{1-p/2}.$$

By an exercise below (with  $|X|^2 = A$ ) we find that  $|X|$  is a compact operator and its eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  satisfies

$$\lambda_n \leq C_p n^{-1/p}, \quad \forall n \geq 1.$$

This implies the desired inequality

$$\|X\|_{\mathfrak{S}_{p,w}}^p = \sup_{\tau > 0} \tau^p |\{n \in \mathbb{N} : \lambda_n > \tau\}| \leq \sup_{\tau > 0} \tau^p |\{n \in \mathbb{N} : C_p n^{-1/p} > \tau\}| \leq C_p.$$

□

**Exercise.** Let  $A \geq 0$  be a self-adjoint operator on a Hilbert space. Let  $\infty > q > 1$ . Assume that for every  $\varepsilon > 0$ , we have the operator inequality

$$A \leq \varepsilon + B_\varepsilon \quad \text{with an operator } B_\varepsilon \geq 0, \quad \text{Tr}(B_\varepsilon) \leq \varepsilon^{1-q}.$$

Prove that  $A$  is a compact operator and its eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  satisfy

$$\lambda_n \leq C n^{-1/q}, \quad \forall n \geq 1.$$

# Chapter 4

## Weyl's law

Weyl's law states that the semiclassical approximation (recall  $t_- = \min(t, 0)$ )

$$\mathrm{Tr} |(-\Delta + V(x))_-|^s \approx \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |(|2\pi k|^2 + V(x))_-|^s dk dx = -L_{s,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |V(x)_-|^{s+d/2} dx$$

becomes correct in the **strong coupling regime**  $V \mapsto \lambda V$  with  $\lambda \gg 1$ . In principle this result holds for all  $d \geq 1$  and all  $s \geq 0$ . Moreover, the result for one  $s$  implies the result for all others  $s \geq 0$  via the layer cake representation (we will come to that).

### 4.1 Coherent States

In this section we discuss a very general method to connect the Schrödinger operator  $-\Delta + V(x)$  and its phase-space representation. The idea goes back to **Schrödinger** (1926).

**Definition** (Coherent States). Take  $G \in C_c^\infty(\mathbb{R}^d)$ ,  $G(x) = G(-x)$ ,  $\|G\|_{L^2} = 1$ . For every  $(k, y) \in \mathbb{R}^d \times \mathbb{R}^d$ , we defined the function  $F_{k,y} \in L^2(\mathbb{R}^d)$  by

$$F_{k,y}(x) := e^{2\pi i k \cdot x} G(x - y), \quad \forall x \in \mathbb{R}^d.$$

Note that  $\|F_{k,y}\|_{L^2(\mathbb{R}^d)} = 1$  for all  $(k, y) \in \mathbb{R}^d \times \mathbb{R}^d$ .

The key feature of the coherent states is that they provide a partition of the identity on  $L^2(\mathbb{R}^d)$  in terms of the phase space  $\mathbb{R}^d \times \mathbb{R}^d$ .



**Theorem** (Resolution of identity). *We have*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |F_{k,y}\rangle \langle F_{k,y}| dk dy = \mathbf{1}_{L^2(\mathbb{R}^d)},$$

namely for all  $u \in L^2(\mathbb{R}^d)$  we have

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dk dy = \|u\|_{L^2}^2.$$

Moreover,

$$\begin{aligned} \int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dk &= (|G|^2 * |u|^2)(y), \\ \int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dy &= (|\widehat{G}|^2 * |\widehat{u}|^2)(k). \end{aligned}$$

*Proof.* For any  $u \in L^2(\mathbb{R}^d)$ , by Plancherel theorem and  $G(x-y) = G(y-x)$  we have

$$\begin{aligned} \int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dk &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \overline{e^{2\pi i k \cdot x} G(y-x)} u(x) dx \right|^2 dk = \int_{\mathbb{R}^d} \left| \widehat{\overline{G(y-\cdot)} u(\cdot)}(k) \right|^2 dk = \\ &= \int_{\mathbb{R}^d} |G(y-x)|^2 |u(x)|^2 dx = (|G|^2 * |u|^2)(y). \end{aligned}$$

Consequently, integrating over  $k \in \mathbb{R}^d$  and using  $\|G\|_{L^2(\mathbb{R}^d)} = 1$  we obtain

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dk dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |G(y-x)|^2 |u(x)|^2 dx dy = \|G\|_{L^2}^2 \|u\|_{L^2}^2 = \|u\|_{L^2}^2.$$

The other identity is left as an exercise. □

**Exercise.** *Prove that for all  $u \in L^2(\mathbb{R}^d)$  and  $k \in \mathbb{R}^d$ ,*

$$\int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dy = (|\widehat{G}|^2 * |\widehat{u}|^2)(k).$$

Now we turn to the analysis of Schrodinger operators. We have the following exact phase-space representation for the kinetic and potential operators.

**Theorem.** *We have the quadratic form identities*

$$-\Delta_{\mathbb{R}^d} = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |2\pi k|^2 |F_{k,y}\rangle \langle F_{k,y}| dk dy - \|\nabla G\|_{L^2}^2$$

and

$$(V * |G|^2) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(y) |F_{k,y}\rangle \langle F_{k,y}| dk dy.$$

*Proof.* **Kinetic term:** we prove that if  $u \in H^1(\mathbb{R}^d)$ , then

$$\int_{\mathbb{R}^d} |\nabla u|^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |2\pi k|^2 |\langle F_{k,y}, u \rangle|^2 dk dy - \|\nabla G\|_{L^2}^2 \|u\|_{L^2}^2.$$

Using

$$\int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dy = (|\widehat{G}|^2 * |\widehat{u}|^2)(k) = \int_{\mathbb{R}^d} |\widehat{G}(k-q)|^2 |\widehat{u}(q)|^2 dq$$

we have

$$\begin{aligned} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |2\pi k|^2 |\langle F_{k,y}, u \rangle|^2 dk dy &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |2\pi k|^2 |\widehat{G}(k-q)|^2 |\widehat{u}(q)|^2 dq dk \\ &= (2\pi)^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |p+q|^2 |\widehat{G}(p)|^2 |\widehat{u}(q)|^2 dq dp = \\ &= (2\pi)^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|q|^2 + |p|^2 + 2p \cdot q) |\widehat{G}(p)|^2 |\widehat{u}(q)|^2 dq dp \end{aligned}$$

The first term is

$$(2\pi)^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |q|^2 |\widehat{G}(p)|^2 |\widehat{u}(q)|^2 dq dp = \|G\|_{L^2}^2 \int_{\mathbb{R}^d} |2\pi q|^2 |\widehat{u}(q)|^2 dq = \|\nabla u\|_{L^2}^2.$$

Similarly, the second term is

$$(2\pi)^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |p|^2 |\widehat{G}(p)|^2 |\widehat{u}(q)|^2 dq dp = \|\nabla G\|_{L^2}^2 \|u\|_{L^2}^2.$$

The cross term vanishes because of the symmetry  $|\widehat{G}(p)| = |\widehat{G}(-p)|$ :

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p \cdot q |\widehat{G}(p)|^2 |\widehat{u}(q)|^2 dq dk = \underbrace{\left( \int_{\mathbb{R}^d} p |\widehat{G}(p)|^2 dl \right)}_{=0} \cdot \left( \int_{\mathbb{R}^d} q |\widehat{u}(q)|^2 dq \right) = 0.$$

**Potential term:** we prove that if  $u \in L^2(\mathbb{R}^d)$ , then

$$\int_{\mathbb{R}^d} (V * |G|^2) |u|^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(y) |\langle F_{k,y}, u \rangle|^2 dk dy.$$

Recall

$$\int_{\mathbb{R}^d} |\langle F_{k,y}, u \rangle|^2 dk = (|G|^2 * |u|^2)(y) = \int_{\mathbb{R}^d} |G(x-y)|^2 |u(x)|^2 dx.$$

Here we have used  $G(x-y) = G(y-x)$ . Hence,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(y) |\langle F_{k,y}, u \rangle|^2 dk dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(y) |G(x-y)|^2 |u(x)|^2 dx dy = \int_{\mathbb{R}^d} (V * |G|^2) |u|^2.$$

□

## 4.2 Weyl's law for sum of eigenvalues

In this section we focus on the case  $s = 1$ , which is relevant to the ground state energy of Fermi gases.

**Theorem** (Weyl's law). *Let  $d \geq 1$ ,  $V \in L^p_{\text{loc}}(\mathbb{R}^d)$  with  $p > \max(1, d/2)$  and  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ . Then in the limit  $\lambda \rightarrow \infty$ :*

$$\text{Tr} |(-\Delta + \lambda V)_-| = L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} + o\left(\lambda^{1+\frac{d}{2}}\right).$$

*Proof. General strategy:* Take a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^d)$  such that  $\|G\|_{L^2} = 1$ . Denote  $\widetilde{V} = |G|^2 * V$ . For any normalized function  $u \in L^2(\mathbb{R}^d)$  we have the phase-space representation

$$\langle u, (-\Delta + \lambda \widetilde{V})u \rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|2\pi k|^2 + \lambda V(y)) |\langle F_{k,y}, u \rangle|^2 dk dy - \|\nabla G\|_{L^2}^2.$$

Therefore, for any orthonormal family  $\{u_n\}_{n=1}^N$ ,

$$\sum_{n=1}^N \langle u_n, (-\Delta + \lambda \tilde{V}) u_n \rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|2\pi k|^2 + \lambda V(y)) \sum_{n=1}^N |\langle F_{k,y}, u_n \rangle|^2 dk dy - N \|\nabla G\|_{L^2}^2.$$

The key observation is that for all  $(k, y) \in \mathbb{R}^d \times \mathbb{R}^d$ , by Bessel's inequality

$$0 \leq \sum_{n=1}^N |\langle F_{k,y}, u_n \rangle|^2 \leq \|F_{k,y}\|_{L^2}^2 = 1.$$

Therefore, by the **bathtub principle**

$$\begin{aligned} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|2\pi k|^2 + \lambda V(y)) \sum_{i=1}^N |\langle F_{k,y}, u_i \rangle|^2 dk dy &\geq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [ |2\pi k|^2 + \lambda V(y) ]_- dk dy \\ &= -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}}. \end{aligned}$$

Consequently, if  $\mathcal{N}(-\Delta + \lambda \tilde{V}) < \infty$ , then we obtain the lower bound

$$\text{Tr}(-\Delta + \lambda \tilde{V})_- \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} - \mathcal{N}(-\Delta + \lambda \tilde{V}) \|\nabla G\|_{L^2}^2.$$

This will lead to the desired lower bound

$$\text{Tr}(-\Delta + \lambda V)_- \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} + o(\lambda^{1+\frac{d}{2}})_{\lambda \rightarrow \infty}$$

provided that we can

- Replace  $\tilde{V} = |G|^2 * V$  by  $V$ , namely take  $|G|^2 \rightarrow \delta_0$ . The difference is controlled by the Lieb-Thirring inequality.
- Show that  $\mathcal{N}(-\Delta + \lambda \tilde{V}) \|\nabla G\|_{L^2}^2 \ll \lambda^{1+\frac{d}{2}}$ . If  $d \geq 3$  and  $V_- \in L^{d/2}$ , then it follows from the CLR bound

$$\mathcal{N}(-\Delta + \lambda \tilde{V}) \leq C_d \int_{\mathbb{R}^d} |\lambda \tilde{V}_-|^{d/2} \approx C_d \int_{\mathbb{R}^d} |\lambda V_-|^{d/2}.$$

In the case  $d \leq 2$  and/or  $V_- \notin L^{d/2}$  we need to refine the analysis slightly.

To achieve the matching upper bound, we use the min-max principle

$$\text{Tr}(-\Delta + V)_- \leq \text{Tr}((-\Delta + V)\gamma), \quad 0 \leq \gamma \leq 1$$

with the choice

$$\gamma = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 + \lambda V(y) < 0) dk dy.$$

Now let us go to the full proof.

**Step 1: Lower bound in the simplest case  $d \geq 3$  and  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d) \cap L^{\frac{d}{2}}(\mathbb{R}^d)$ .**

By the min-max principle, we can assume  $V = V_- \leq 0$ . Assume that  $-\Delta + \lambda V$  has  $N$  eigenvalues with orthonormal eigenfunctions  $\{u_i\}_{i=1}^N$ . By the CLR bound,

$$N \leq C_d \int_{\mathbb{R}^d} |\lambda V_-|^{\frac{d}{2}}.$$

We need to prove that

$$\mathrm{Tr}((-\Delta + \lambda V)_-) = \sum_{i=1}^N \langle u_i, (-\Delta + \lambda V) u_i \rangle \geq -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} + o(\lambda^{1+\frac{d}{2}})_{\lambda \rightarrow \infty}.$$

Let us decompose

$$\sum_{i=1}^N \langle u_i, (-\Delta + \lambda V) u_i \rangle = \sum_{i=1}^N \langle u_i, ((1-\varepsilon)(-\Delta) + \lambda \tilde{V}) u_i \rangle + \sum_{i=1}^N \langle u_i, (\varepsilon(-\Delta) + \lambda(V - \tilde{V})) u_i \rangle$$

with a parameter  $\varepsilon \in (0, 1)$  and  $\tilde{V} = G^2 * V$  with a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^d)$ ,  $\|G\|_{L^2} = 1$ . Recall that  $F_{k,y}(x) := e^{2\pi i k \cdot x} G(x - y)$ .

For the first term, we have the coherent state identity

$$\begin{aligned} \sum_{i=1}^N \langle u_i, ((1-\varepsilon)(-\Delta) + \tilde{V}) u_i \rangle &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} ((1-\varepsilon)|2\pi k|^2 + \lambda V(y)) \sum_{i=1}^N |\langle F_{k,y}, u_i \rangle|^2 dk dy \\ &\quad - N(1-\varepsilon) \|\nabla G\|_{L^2}^2. \end{aligned}$$

Using the uniform bound

$$0 \leq \sum_{i=1}^N |\langle F_{k,y}, u_i \rangle|^2 \leq \|F_{k,y}\|_2^2 = 1, \quad \forall (k, y) \in \mathbb{R}^d \times \mathbb{R}^d$$

we obtain

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} ((1-\varepsilon)|2\pi k|^2 + \lambda V(y)) \sum_{i=1}^N |\langle F_{k,y}, u_i \rangle| dk dy$$

$$\geq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [(1-\varepsilon)|2\pi k|^2 + \lambda V(y)]_- dk dy = -\frac{L_{1,d}^{\text{cl}}}{(1-\varepsilon)^{d/2}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}}.$$

For the second term, we use the Lieb-Thirring inequality to get the lower bound

$$\sum_{i=1}^N \left\langle u_i, \left( \varepsilon(-\Delta) + \lambda(V - \tilde{V}) \right) u_i \right\rangle \geq -\frac{C_d}{\varepsilon^{d/2}} \int_{\mathbb{R}^d} |\lambda(V - \tilde{V})_-|^{1+\frac{d}{2}}.$$

Combining with the above upper bound on  $N$ , we conclude that

$$\text{Tr}((-\Delta + \lambda V)_-) \geq -\frac{L_{1,d}^{\text{cl}}}{(1-\varepsilon)^{d/2}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} - \frac{C_d}{\varepsilon^{d/2}} \int_{\mathbb{R}^d} |\lambda(V - \tilde{V})_-|^{1+\frac{d}{2}} - C_d \|\nabla G\|_{L^2}^2 \int_{\mathbb{R}^d} |\lambda V_-|^{\frac{d}{2}}.$$

Therefore,

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr}((-\Delta + \lambda V)_-) \geq -\frac{L_{1,d}^{\text{cl}}}{(1-\varepsilon)^{d/2}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}} - \frac{C_d}{\varepsilon^{d/2}} \int_{\mathbb{R}^d} |(V - \tilde{V})_-|^{1+\frac{d}{2}}.$$

This holds for every  $\varepsilon \in (0, 1)$  and  $V = G^2 * V$ . Replacing  $G$  by

$$G_n(x) = n^{d/2} G_1(nx), \quad n \geq 1$$

for a fixed function  $G_1$  we find that  $|G_n|^2 * V \rightarrow V$  strongly in  $L^{1+d/2}(\mathbb{R}^d)$  as  $n \rightarrow \infty$  (recall that we are assuming  $V = V_- \in L^{1+d/2}$ ). Thus

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr}((-\Delta + \lambda V)_-) \geq -\frac{L_{1,d}^{\text{cl}}}{(1-\varepsilon)^{d/2}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}$$

for all  $\varepsilon \in (0, 1)$ . Finally, sending  $\varepsilon \rightarrow 0$  we obtain the desired lower bound

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr}((-\Delta + \lambda V)_-) \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}.$$

**Step 2: Lower bound in the general case  $d \geq 1$  and  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ .**

Let us explain how to remove the restriction  $d \geq 3$  and  $V_- \in L^{\frac{d}{2}}(\mathbb{R}^d)$ .

**Removal of the restriction on regularity.** First, the additional regularity condition on  $V$  can be removed easily using the Lieb-Thirring inequality. To be precise, let us assume that we have proved the desired lower bound

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr}((-\Delta + \lambda V)_-) \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}$$

for all  $V_- \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . Now consider a general  $V$  satisfying only  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ . Then for the lower bound, we can focus on the case  $V = V_- \leq 0$ . Take a sequence  $\{V_n\}_{n=1}^\infty$  such that  $V_n \rightarrow V$  strongly in  $L^{1+\frac{d}{2}}(\mathbb{R}^d)$ . Taking  $\varepsilon \in (0, 1)$ , we can split

$$-\Delta + \lambda V = \left[ (1 - \varepsilon)(-\Delta) + \lambda V_n \right] + \left[ \varepsilon(-\Delta) + \lambda(V - V_n) \right].$$

By the min-max principle (see an exercise below), we have the lower bound

$$\mathrm{Tr}((-\Delta + \lambda V)_-) \geq \mathrm{Tr} \left( \left[ (1 - \varepsilon)(-\Delta) + \lambda V_n \right]_- \right) + \mathrm{Tr} \left( \left[ \varepsilon(-\Delta) + \lambda(V - V_n) \right]_- \right).$$

By the assumed lower bound for functions in  $L^1 \cap L^\infty$ , we have

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr} \left( \left[ (1 - \varepsilon)(-\Delta) + \lambda V_n \right]_- \right) \geq -\frac{L_{1,d}^{\mathrm{cl}}}{(1 - \varepsilon)^{d/2}} \int_{\mathbb{R}^d} |V_n|^{1+\frac{d}{2}}.$$

By the Lieb-Thirring inequality

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr} \left( \left[ \varepsilon(-\Delta) + \lambda(V - V_n) \right]_- \right) \geq -\frac{C_d}{\varepsilon^{d/2}} \int_{\mathbb{R}^d} |V - V_n|^{1+\frac{d}{2}}.$$

Thus

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr}((-\Delta + \lambda V)_-) \geq -\frac{L_{1,d}^{\mathrm{cl}}}{(1 - \varepsilon)^{d/2}} \int_{\mathbb{R}^d} |V_n|^{1+\frac{d}{2}} - \frac{C_d}{\varepsilon^{d/2}} \int_{\mathbb{R}^d} |V - V_n|^{1+\frac{d}{2}}.$$

Sending  $n \rightarrow \infty$ , and then  $\varepsilon \rightarrow 0$  we obtain the desired lower bound for  $\mathrm{Tr}((-\Delta + \lambda V)_-)$ .

**Exercise.** Let  $A$  be a self-adjoint operator on a Hilbert space such that  $A_- = A\mathbf{1}(A < 0)$  is a trace class operator. Prove that

$$\mathrm{Tr}(A_-) = \inf_{0 \leq \gamma \leq 1} \mathrm{Tr}(A\gamma).$$

Here we use the convention  $\mathrm{Tr}(A\gamma) = \mathrm{Tr}(\sqrt{\gamma}A\sqrt{\gamma}) = \mathrm{Tr}(\sqrt{\gamma}A_-\sqrt{\gamma}) + \mathrm{Tr}(\sqrt{\gamma}A_+\sqrt{\gamma})$ .

**Removal of the restriction on the dimension.** Let us consider  $d \leq 2$  and prove

$$\liminf_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr}((-\Delta + \lambda V)_-) \geq -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}$$

for  $V = V_- \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . Introducing the energy cut-off  $E > 0$ , we can write

$$\mathrm{Tr}((-\Delta + \lambda V)_-) = \mathrm{Tr}((-\Delta + \lambda V)\mathbf{1}(-\Delta + \lambda V \leq -E)) + \mathrm{Tr}((-\Delta + \lambda V)\mathbf{1}(0 > -\Delta + \lambda V > -E)).$$

Note that in low dimensions, the number of eigenvalues  $\leq -E$  of  $-\Delta + \lambda V$  is bounded (see an exercise below). Therefore, we can prove

$$\mathrm{Tr}((-\Delta + \lambda V)\mathbf{1}(-\Delta + \lambda V \leq -E)) \geq -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+\frac{d}{2}} + o(\lambda^{1+\frac{d}{2}})$$

by repeating the proof in Step 1 and replacing the CLR bound by

$$\mathcal{N}(-\Delta + \lambda V + E) \leq C_d E^{\frac{d-4}{2}} \int_{\mathbb{R}^d} |\lambda V|^2.$$

This contributes to the error  $o(\lambda^{1+\frac{d}{2}})$  if we choose  $E$  such that

$$E^{\frac{d-4}{2}} \lambda^2 \ll \lambda^{1+d/2} \iff E \gg \lambda^{\frac{2-d}{4-d}}.$$

On the other hand, using the obvious operator inequality

$$A\mathbf{1}(0 > A > -E) \geq -E^{1-s}|A_-|^s, \quad \forall s \in (0, 1)$$

with  $A = -\Delta + \lambda V$  and the Lieb-Thirring inequality we have

$$\mathrm{Tr}((-\Delta + \lambda V)\mathbf{1}(0 > -\Delta + \lambda V > -E)) \geq -C_d E^{1-s} \int_{\mathbb{R}^d} |\lambda V|^{s+\frac{d}{2}}, \quad \forall 1/2 < s < 1.$$

This contributes to the error  $o(\lambda^{1+\frac{d}{2}})$  if we choose  $E$  such that

$$E^{1-s} \lambda^{s+\frac{d}{2}} \ll \lambda^{1+\frac{d}{2}} \iff E \ll \lambda.$$

Thus eventually we choose  $\lambda \gg E \gg \lambda^{\frac{2-d}{4-d}}$  and  $s \in (1/2, 1)$  arbitrary. This concludes the proof of the lower bound for  $d \leq 2$ .

**Exercise.** Let  $3 \geq d \geq 1$  and  $V \in L^2(\mathbb{R}^d)$ . Prove that for every  $E > 0$

$$\mathcal{N}(-\Delta + V + E) \leq C_d E^{\frac{d-4}{2}} \int_{\mathbb{R}^d} |V|^2.$$



Here  $\mathcal{N}(-\Delta + V + E)$  is the number of negative eigenvalue of  $-\Delta + V + E$ .

**Step 3: Upper bound in the case  $V \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$  and  $V_- \in L^{\frac{d}{2}}(\mathbb{R}^d)$ .**

We use the variational principle (see an exercise above)

$$\mathrm{Tr}((-\Delta + \lambda V)_-) \leq \mathrm{Tr}((-\Delta + \lambda V)\gamma)$$

with the choice

$$\gamma := \iint_{\mathcal{M}} |F_{k,y}\rangle \langle F_{k,y}| dk dy, \quad \mathcal{M} = \{(k, y) \in \mathbb{R}^d \times \mathbb{R}^d : |2\pi k|^2 + \lambda V(y) < 0\}.$$

Recall the coherent states

$$F_{k,y}(x) := e^{2\pi i k \cdot x} G(x - y)$$

with a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^d)$ ,  $\|G\|_{L^2} = 1$ . Clearly  $0 \leq \gamma \leq 1$  (by the resolution of identity) and

$$\mathrm{Tr} \gamma = |\mathcal{M}| = L_{0,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} < \infty.$$

Thus

$$\mathrm{Tr} [ -\Delta + \lambda V ]_- \leq \mathrm{Tr} [ (-\Delta + \lambda V)\gamma ] = \iint_{\mathcal{M}} \langle F_{k,y}, (-\Delta + \lambda V) F_{k,y} \rangle dk dy$$

Now we calculate

$$\langle F_{k,y}, (-\Delta + \lambda V) F_{k,y} \rangle = \int_{\mathbb{R}^d} (|\nabla_x F_{k,y}(x)|^2 + \lambda V(x) |F_{k,y}(x)|^2) dx.$$

The potential part is easy as  $|F_{k,y}(x)|^2 = |G(x - y)|^2$ . For the kinetic part we use the fact that  $G$  is real-valued:

$$\begin{aligned} |\nabla_x F_{k,y}(x)|^2 &= |2\pi i k e^{2\pi i k \cdot x} G(x - y) + e^{2\pi i k \cdot x} \nabla_x G(x - y)|^2 \\ &= |2\pi k|^2 |G(x - y)|^2 + |\nabla_x G(x - y)|^2. \end{aligned}$$

Hence,

$$\begin{aligned} \langle F_{k,y}, (-\Delta + \lambda V)F_{k,y} \rangle &= \int_{\mathbb{R}^d} (|2\pi k|^2 |G(x-y)|^2 + |\nabla_x G(x-y)|^2 + \lambda V(x)|G(x-y)|^2) dx = \\ &= |2\pi k|^2 + \|\nabla G\|_{L^2}^2 + (\lambda V * G^2)(y). \end{aligned}$$

Therefore we conclude that

$$\begin{aligned} \text{Tr} [-\Delta + \lambda V]_- &\leq \iint_{\mathcal{M}} \langle F_{k,y}, (-\Delta + \lambda V)F_{k,y} \rangle dk dy \\ &= \iint_{\mathcal{M}} (|2\pi k|^2 + \|\nabla G\|_{L^2}^2 + (\lambda V * G^2)(y)) dk dy \\ &= \iint_{\mathcal{M}} (|2\pi k|^2 + \lambda V(y)) dk dy + \lambda \iint_{\mathcal{M}} ((V * G^2)(y) - V(y)) dk dy + \|\nabla G\|_{L^2}^2 |\mathcal{M}| \\ &= -L_{1,d}^{\text{cl}} \lambda^{1+\frac{d}{2}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}} + L_{0,d}^{\text{cl}} \lambda^{1+\frac{d}{2}} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V * G^2 - V) dy + \|\nabla G\|_{L^2}^2 L_{0,d}^{\text{cl}} \lambda^{d/2} \int_{\mathbb{R}^d} |V_-|^{d/2}. \end{aligned}$$

Thus

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr} [-\Delta + \lambda V]_- \leq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}} + L_{0,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V * G^2 - V) dy.$$

We can replace  $G$  by  $G_n(x) = n^{d/2} G_1(nx)$  for a fixed function  $0 \leq G_1 \in C_c^\infty$  with  $\|G_1\|_{L^2(\mathbb{R}^d)} = 1$ . Since  $V \in L^{1+\frac{d}{2}}$ , we have  $V * G_n^2 \rightarrow V$  strongly in  $L^{1+\frac{d}{2}}(\mathbb{R}^d)$  when  $n \rightarrow \infty$ , and hence by Hölder inequality

$$\begin{aligned} \left| \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V * G_n^2 - V) dy \right| &\leq \| |V_-|^{\frac{d}{2}} \|_{L^{1+\frac{2}{d}}} \|V * G_n^2 - V\|_{L^{1+\frac{d}{2}}} \\ &= \|V_-\|_{L^{1+\frac{d}{2}}}^{\frac{d}{2}} \|V * G_n^2 - V\|_{L^{1+\frac{d}{2}}} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Thus we obtain the desired upper bound

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr} [-\Delta + \lambda V]_- \leq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}.$$

**Step 4: Upper bound for  $V_+ \in L_{\text{loc}}^p(\mathbb{R}^d)$  with some  $p > \max(1, d/2)$ ,  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ .**

Let us explain how to remove the restrictions  $V_+ \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$  and  $V_- \in L^{\frac{d}{2}}(\mathbb{R}^d)$ .

**Removal the restriction**  $V_+ \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ . Take  $V_+ \in L_{\text{loc}}^p(\mathbb{R}^d)$  with  $p > \max(1, d/2)$ . For a technical reason, we assume also that  $V_- \in L^\infty(\mathbb{R}^d)$  and  $V_-$  has compact support. Then the proof in Step 3 gives us

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr} [-\Delta + \lambda V]_- \leq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} V_-^{1+\frac{d}{2}} + L_{0,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V * G^2 - V) \, dy.$$

We can decompose

$$\begin{aligned} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V * G^2 - V) &= \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V_- * G^2 - V_-) \, dy + \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V_+ * G^2) \, dy \\ &= \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V_- * G^2 - V_-) \, dy + \int_{\mathbb{R}^d} (|V_-|^{\frac{d}{2}} * G^2) V_+ \, dy \end{aligned}$$

Again, we replace  $G$  by  $G_n(x) = n^{d/2} G_1(nx)$  for a fixed function  $0 \leq G_1 \in C_c^\infty$  with  $\|G_1\|_{L^2(\mathbb{R}^d)} = 1$ . Then similarly to Step 3, using  $V_- \in L^{1+\frac{d}{2}}$  we have

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} |V_-|^{\frac{d}{2}} (V_- * G_n^2 - V_-) \, dy = 0.$$

Since  $V_-$  and  $G_1$  have compact supports, the supports of  $V_-$  and  $|V_-|^{\frac{d}{2}} * G_n^2$  are contained in a ball  $B_R$  independent of  $n$ . Moreover, since  $V_- \in L^\infty$  we have  $|V_-|^{\frac{d}{2}} * G_n^2 \rightarrow |V_-|^{\frac{d}{2}}$  strongly in any  $L^q(B_R)$  with  $1 \leq q < \infty$ . Since  $V_+ \in L_{\text{loc}}^p$ , we have  $V_+ \in L^p(B_R)$ , and hence by Holder's inequality

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} (|V_-|^{\frac{d}{2}} * G_n^2) V_+ = \lim_{n \rightarrow \infty} \int_{B_R} (|V_-|^{\frac{d}{2}} * G_n^2) V_+ = \int_{B_R} |V_-|^{\frac{d}{2}} V_+ = 0.$$

Thus we get the desired upper bound

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \text{Tr} [-\Delta + \lambda V]_- \leq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} V_-^{1+\frac{d}{2}}.$$

**Removal the restriction in  $V_-$ .** Now we assume only  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$ . For every  $R > 0$ , we consider

$$V_R = V_+ + V_- \mathbf{1}(|x| \leq R) \mathbf{1}(|V_-| \leq R).$$

Since  $V \leq V_R$ , by the min-max principle and the

$$\mathrm{Tr}((-\Delta + \lambda V)_-) \leq \mathrm{Tr}((-\Delta + \lambda V_R)_-).$$

Moreover,

$$(V_R)_- = V_- \mathbf{1}(|x| \leq R) \mathbf{1}(|V_-| \leq R)$$

is compactly supported and uniform bounded. Therefore, by using the above upper bound for  $V_R$  we have

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr} [ -\Delta + \lambda V ]_- \leq \limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr} [ -\Delta + \lambda V_R ]_- \leq -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |(V_R)_-|^{1+\frac{d}{2}}.$$

Then by Monotone Convergence Theorem,

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \mathrm{Tr} [ -\Delta + \lambda V ]_- \leq -L_{1,d}^{\mathrm{cl}} \lim_{R \rightarrow \infty} \int_{\mathbb{R}^d} |(V_R)_-|^{1+\frac{d}{2}} = -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}.$$

This completes the proof of Weyl's law for the sum of negative eigenvalues.  $\square$

In the above proof of the upper bound, we have used the standard choice of the trial operator

$$\gamma := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 + \lambda V(y) < 0) dk dy.$$

We can also use a modified trial operator which simplifies the computation a bit.

**Exercise.** Let  $d \geq 1$ . Assume that  $V_- \in L^{1+\frac{d}{2}}(\mathbb{R}^d)$  and  $V_+ \in L_{\mathrm{loc}}^p(\mathbb{R}^d)$  with  $p \geq \max(1, d/2)$  if  $d \neq 2$  and  $p > 1$  if  $d = 2$ . Let  $F_{k,y}(x) = e^{2\pi i k \cdot x} G(x - y)$  with a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^d)$  satisfying  $\|G\|_{L^2(\mathbb{R}^d)} = 1$  and define the operator on  $L^2(\mathbb{R}^d)$

$$\tilde{\gamma} := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 + \lambda(G^2 * V)(y) + \|\nabla G\|_{L^2}^2 < 0) dk dy.$$

(i) Prove that

$$\mathrm{Tr}((-\Delta + \lambda V)\tilde{\gamma}) = -L_{1,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} \left| (\lambda G^2 * V + \|\nabla G\|_{L^2}^2)_- \right|^{1+\frac{d}{2}}.$$

(ii) Using an appropriate choice of  $G$  to deduce that

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-(1+d/2)} \operatorname{Tr}((-\Delta + \lambda V)_-) \leq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}.$$

# Chapter 5

## Dirichlet Laplacian

**Definition** (Dirichlet Laplacian). Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . Consider  $-\Delta$  on  $L^2(\Omega)$  with the core domain  $C_c^\infty(\Omega)$ . Since  $-\Delta \geq 0$ , namely

$$\langle u, -\Delta u \rangle = \int_{\Omega} |\nabla u|^2 \geq 0, \quad \forall u \in C_c^\infty(\Omega),$$

it defines the **Dirichlet Laplacian**  $-\Delta_D$  by Friedrich's extension. The quadratic form domain of  $-\Delta_D$  is denoted by  $H_0^1(\Omega)$ .

Remarks:

- If we consider  $-\Delta$  on  $L^2(\Omega)$  with the core domain  $C^\infty(\bar{\Omega})$ , then we also have  $-\Delta \geq 0$  and it defines the **Neumann Laplacian**  $-\Delta_N$  by Friedrich's extension. The corresponding quadratic form is  $H^1(\Omega)$  which is the Hilbert space with the natural norm

$$\|u\|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2.$$

Here the derivatives  $\nabla u = (\partial_{x_1} u, \dots, \partial_{x_d} u)$  should be interpreted in the **distributional sense**, namely

$$\int_{\Omega} (\partial_{x_d} u) \varphi dx = - \int_{\Omega} u \partial_{x_d} \varphi dx, \quad \forall \varphi \in C_c^\infty(\Omega).$$

- In general,  $H_0^1(\Omega) \neq H^1(\Omega)$  and  $-\Delta_D \neq -\Delta_N$ . **The boundary matters here!** In fact, by the definition of the quadratic forms:

$$H_0^1(\Omega) = \overline{C_c^\infty(\Omega)}^{H^1(\Omega)}, \quad H^1(\Omega) = \overline{C^\infty(\bar{\Omega})}^{H^1(\Omega)}.$$

In practice, any function in  $H_0^1(\Omega)$  can be approximated by a function in  $C_c^\infty(\Omega)$ , and the latter can be think of as a function  $\mathbb{R}^d \rightarrow \mathbb{C}$  (extended by 0 outside  $\Omega$ ). This density argument makes the computations on the Dirichlet Laplacian rather similar to the usual Laplacian on  $L^2(\mathbb{R}^d)$ .

- Obviously for any  $u \in H^1(\mathbb{R}^d)$ , the restriction  $u|_\Omega$  always belong to  $H^1(\Omega)$ . The reverse direction requires the smoothness of the boundary  $\partial\Omega$ : if the boundary is  $C^1$ , then

$$H^1(\Omega) = \{u|_\Omega : u \in H^1(\mathbb{R}^d)\}.$$

In this case, we also have

$$H_0^1(\Omega) = \{u \in H^1(\Omega) : u|_{\partial\Omega} = 0\}.$$

Here the **trace operator**  $u \mapsto u|_{\partial\Omega}$ , first defined for smooth functions  $u$ , can be extended to be a continuous linear operator  $H^1(\Omega) \rightarrow L^2(\partial\Omega)$ .

We will need only the following simple fact on  $H_0^1(\Omega)$ .

**Lemma.** *For any open set  $\Omega \subset \mathbb{R}^d$ , if  $u \in H^1(\Omega)$  and  $\text{supp } u \subset\subset \Omega$ , then  $u \in H_0^1(\Omega)$ . Consequently, if  $u \in H^1(\mathbb{R}^d)$  and  $\text{supp } u \subset\subset \Omega$ , then  $u|_\Omega \in H_0^1(\Omega)$ .*

*Proof.* Because  $\text{supp } u \subset\subset \Omega$  there exists a  $\epsilon > 0$  such that

$$\text{supp } u + B_\epsilon(0) \subset \Omega.$$

Choose

$$g \in C_c^\infty(\mathbb{R}^d), \quad \text{supp } g \subset B_1(0), \quad \int_{\mathbb{R}^d} g = 1.$$

Define  $g_n(x) = n^d g(nx)$ . Then

$$g_n \in C_c^\infty(\mathbb{R}^d), \quad \text{supp } g_n \subset B_{n^{-1}}(0), \quad \int_{\mathbb{R}^d} g_n = 1.$$

Denote the function  $\tilde{u} : \mathbb{R}^d \rightarrow \mathbb{C}$  by

$$\tilde{u} = \begin{cases} u(x), & \text{if } x \in \Omega \\ 0, & \text{if } x \notin \Omega \end{cases}$$

Then  $\varphi_n := \tilde{u} * g_n \in C_c^\infty(\mathbb{R}^d)$ ,

$$\text{supp } \varphi_n \subset \text{supp}(\tilde{u}) + \text{supp}(g_n) \subset \text{supp}(\tilde{u}) + B_{n^{-1}}(0) \subset\subset \Omega$$

and  $\varphi_n \rightarrow u$  strongly in  $H^1(\Omega)$ . Thus  $u \in H_0^1(\Omega)$ .  $\square$

## 5.1 Berezin-Li-Yau inequality

Recall the Lieb-Thirring kinetic inequality: for any orthonormal family  $\{u_n\}_{n=1}^N$  in  $L^2(\mathbb{R}^d)$ ,

$$\sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla u_n|^2 \geq K_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}}, \quad \rho(x) = \sum_{n=1}^N |u_n(x)|^2.$$

The **Lieb-Thirring conjecture** states that we actually have

$$K_d = K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}} \quad \text{when } d \geq 3.$$

Here  $|B_1|$  is the volume of the unit ball in  $\mathbb{R}^d$ . If we assume that all functions  $\{u_n\}$  are supported on a bounded set  $\Omega$ , then the LT conjecture implies that

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq K_d^{\text{cl}} \int_{\Omega} \rho^{1+\frac{2}{d}} \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( \int_{\Omega} \rho \right)^{1+\frac{2}{d}} = \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}}.$$

This weaker inequality has been proved rigorously by **Berezin** (1972) and **Li-Yau** (1983), and it holds in all dimensions  $d \geq 1$ .

**Theorem** (Berezin-Li-Yau inequality). *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. For  $N \geq 1$  and any orthonormal family  $\{u_n\}_{n=1}^N$  in  $L^2(\Omega)$  with  $u_n \in H_0^1(\Omega)$ ,*

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}}.$$

*Proof.* By a density argument, we can take  $u_n \in C_c^\infty(\Omega)$  for all  $n$  and think of  $\{u_n\}$  as functions  $\mathbb{R}^d \rightarrow \mathbb{C}$  (extended by 0 outside  $\Omega$ ). Using Fourier transform, we can write

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 = \sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla u_n|^2 = \sum_{n=1}^N \int_{\mathbb{R}^d} |2\pi k|^2 |\hat{u}_n|^2 = \int_{\mathbb{R}^d} |2\pi k|^2 F(k) dk$$



where

$$0 \leq F(k) := \sum_{i=1}^N |\widehat{u}_n(k)|^2 = \sum_{n=1}^N \left| \int_{\Omega} u_n(x) e^{-2\pi i k \cdot x} dx \right|^2 \leq \int_{\Omega} |e^{-2\pi i k \cdot x}|^2 dx = |\Omega|.$$

Here we have used the fact that  $\{u_n\}$  is an orthonormal family in  $L^2(\Omega)$  and Bessel's inequality. By the **bathtub principle**, it is easy to see that the minimum

$$\inf \left\{ \int_{\mathbb{R}^d} |2\pi k|^2 F(k) dk \mid 0 \leq F \leq |\Omega|, \int_{\mathbb{R}^d} F = N \right\}$$

is attained by

$$F_0(k) = |\Omega| \mathbf{1}_{B_R}(k)$$

where ball  $B_R = B_R(0)$  is determined by

$$N = \int_{\mathbb{R}^d} F_0 = |\Omega| |B_R| = |\Omega| |B_1| R^d \iff R = \left( \frac{N}{|\Omega| |B_1|} \right)^{\frac{1}{d}}.$$

Thus

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq \int_{B_R} |2\pi k|^2 |\Omega| dk = |\Omega| R^{d+2} \int_{|k| \leq 1} |2\pi k|^2 dk = \underbrace{\frac{d}{d+2} \frac{4\pi^2}{|B_1|^{\frac{2}{d}}}}_{=K_{cl}} \frac{1}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}}.$$

□

As a direct consequence of the Berezin-Li-Yau inequality, we have the lower bound for the sum of eigenvalues of the Dirichlet Laplacian on  $\Omega$ .

**Exercise.** Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Let  $\mu_1 \leq \mu_2 \leq \dots$  be the min-max values of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$ .

(i) Prove that

$$\sum_{i=1}^N \mu_i \geq \frac{K_d^{cl}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}}, \quad \forall N \geq 1.$$

- (ii) Deduce that  $-\Delta_{\mathbb{D}}$  has compact resolvent (hence all  $\{\mu_n\}$  are eigenvalues).
- (iii) Prove that  $0 < \mu_1 < \mu_2$ .

The Berezin-Li-Yau inequality can be rewritten in the following dual form.

**Theorem** (Berezin-Li-Yau inequality). *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Let  $\mu_1 \leq \mu_2 \leq \dots$  be the eigenvalues of the Dirichlet Laplacian  $-\Delta_{\mathbb{D}}$  on  $L^2(\Omega)$ . Then for all  $\lambda > 0$ , we have*

$$\sum_{n=1}^{\infty} [\mu_n - \lambda]_- \geq -L_{1,d}^{\text{cl}} |\Omega| \lambda^{1+\frac{d}{2}}.$$

*Proof.* For every  $N \geq 1$ , we have

$$\sum_{n=1}^N (\mu_n - \lambda) = \sum_{n=1}^N \mu_n - N\lambda \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}} - N\lambda \geq -L_{1,d}^{\text{cl}} |\Omega| \lambda^{1+\frac{d}{2}}.$$

Here we have used Young's inequality

$$\frac{a^p}{p} + \frac{b^q}{q} \geq ab, \quad a, b \geq 0, \quad p, q \geq 1, \quad \frac{1}{p} + \frac{1}{q} = 1$$

and the relation

$$\left( \left(1 + \frac{2}{d}\right) K_d \right)^{1+\frac{d}{2}} \left( \left(1 + \frac{d}{2}\right) L_{1,d} \right)^{1+\frac{2}{d}} = 1.$$

Thus we can take  $N = N(\lambda)$  the largest index such that  $\mu_{N(\lambda)} < \lambda$  and obtain

$$\sum_{n=1}^{\infty} [\mu_n - \lambda]_- = \sum_{n=1}^{N(\lambda)} (\mu_n - \lambda) \geq -L_{1,d}^{\text{cl}} |\Omega| \lambda^{1+\frac{d}{2}}.$$

□

Remark: Heuristically, the above inequality justifies **the Lieb-Thirring conjecture**

$$\text{Tr}_{L^2(\mathbb{R}^d)} (-\Delta + V)_- \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |V_-|^{1+\frac{d}{2}}$$

for the hard core potential

$$V(x) = \begin{cases} -\lambda & \text{if } x \in \Omega, \\ +\infty & \text{if } x \notin \Omega. \end{cases}$$

## 5.2 Sum of eigenvalues

The lower bound in the Berezin-Li-Yau inequality is sharp in the limit  $N \rightarrow \infty$ , namely the semiclassical constant  $K_d^{\text{cl}}$  is optimal.

**Theorem** (Weyl's law for the sum of eigenvalues). *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Then the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  satisfy*

$$\sum_{i=1}^N \mu_i = \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}} + o(N^{1+\frac{2}{d}})_{N \rightarrow \infty}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}}.$$

*Proof.* The Berezin-Li-Yau inequality gives the lower bound (even without error). It remains to prove the upper bound

$$\sum_{i=1}^N \mu_i \leq \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}} + o(N^{1+\frac{2}{d}})_{N \rightarrow \infty}.$$

Recall that by the min-max principle,

$$\begin{aligned} \sum_{i=1}^N \mu_i &= \inf \left\{ \sum_{i=1}^N \langle u_i, -\Delta_D u_i \rangle \left| \{u_i\}_{i=1}^N \subset H_0^1(\Omega) \text{ orthonormal family in } L^2(\Omega) \right. \right\} \\ &= \inf \left\{ \sum_{i \geq 1} \lambda_i \langle u_i, -\Delta_D u_i \rangle \left| \{u_i\}_{i \geq 1} \subset H_0^1(\Omega) \text{ an ONF in } L^2(\Omega), 0 < \lambda_i \leq 1, \sum_i \lambda_i = N \right. \right\}. \end{aligned}$$

We choose the trial operator  $\gamma = \sum_{i \geq 1} \lambda_i |u_i\rangle \langle u_i|$  using the coherent states:

$$\gamma := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{B_R}(k) \mathbf{1}_{\tilde{\Omega}}(y) |F_{k,y}\rangle \langle F_{k,y}| dk dy$$

with a ball  $B_R = B_R(0)$  and a set  $\tilde{\Omega} \subset\subset \Omega$ . Recall that  $F_{k,y}(x) = e^{2\pi i k \cdot x} G(x - y)$  with a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^d)$  is a radial function. Then  $\gamma$  is a trace class operator on

$L^2(\mathbb{R}^d)$  satisfying  $0 \leq \gamma \leq 1$ . Moreover, we can choose  $R > 0$  such that

$$\mathrm{Tr} \gamma = |\tilde{\Omega}| |B_R| = N \iff |\tilde{\Omega}| |B_1| R^d = N \iff R = \left( \frac{N}{|\tilde{\Omega}| |B_1|} \right)^{\frac{1}{d}}.$$

In particular, we have the spectral decomposition

$$\gamma = \sum_{i \geq 1} \lambda_i |u_i\rangle \langle u_i|, \quad 0 < \lambda_i \leq 1, \quad \sum_{i \geq 1} \lambda_i = N, \quad \{u_i\} \text{ ONF in } L^2(\mathbb{R}^d).$$

We can also require

$$\mathrm{supp} G \subset B_\delta(0) \quad \text{with } \delta := \frac{1}{2} \mathrm{dist}(\tilde{\Omega}, \Omega^c) > 0.$$

Then all  $\mathrm{supp} u_i \subset \tilde{\Omega} + \mathrm{supp} G \subset \subset \Omega$  because for any test function  $\varphi \geq 0$  supported outside  $\tilde{\Omega} + \mathrm{supp} G$  we have

$$\sum_{i \geq 1} \lambda_i \left| \int_{\mathbb{R}^d} \varphi u_i \right|^2 = \mathrm{Tr}(\varphi^2 \gamma) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{B_R}(k) \mathbf{1}_{\tilde{\Omega}}(y) \underbrace{|\langle F_{k,y}, \varphi \rangle|^2}_{=0} dk dy = 0$$

(as  $F_{k,y} = e^{2\pi k \cot x} G(x-y)$  and  $\varphi(x)$  have disjoint supports). Next, we compute

$$\begin{aligned} \sum_{i \geq 1} \lambda_i \int_{\mathbb{R}^d} |\nabla u_i|^2 &= \mathrm{Tr}_{L^2(\mathbb{R}^d)}(-\Delta \gamma) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{B_R(0)}(k) \mathbf{1}_{\tilde{\Omega}}(y) \|\nabla F_{k,y}\|_{L^2(\mathbb{R}^d)}^2 dk dy \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{B_R(0)}(k) \mathbf{1}_{\tilde{\Omega}}(y) \left( |2\pi k|^2 + \|\nabla G\|_{L^2}^2 \right) dk dy \\ &= |\tilde{\Omega}| R^{d+2} \int_{|k| \leq 1} |2\pi k|^2 dk + |\tilde{\Omega}| |B_R| \|\nabla G\|_{L^2}^2 \\ &= \frac{K_d^{\mathrm{cl}}}{|\tilde{\Omega}|^{\frac{2}{d}}} N^{1+\frac{2}{d}} + N \|\nabla G\|_{L^2}^2. \end{aligned}$$

This calculation also shows that  $u_i \in H^1(\mathbb{R}^d)$ , and hence  $(u_i)|_\Omega \in H_0^1(\Omega)$ . Thus in summary, by the min-max principle we conclude that

$$\sum_{i=1}^N \mu_i(-\Delta_{\mathrm{D}}) \leq \frac{K_d^{\mathrm{cl}}}{|\tilde{\Omega}|^{\frac{2}{d}}} N^{1+\frac{2}{d}} + N \|\nabla G\|_{L^2}^2.$$

Consequently,

$$\limsup_{N \rightarrow \infty} N^{-(1+2/d)} \sum_{i=1}^N \mu_i(-\Delta_{|D}) \leq \frac{K_d^{\text{cl}}}{|\widetilde{\Omega}|^{\frac{2}{d}}}.$$

The latter inequality holds true for any set  $\widetilde{\Omega} \subset \subset \Omega$ . We can optimize over  $\widetilde{\Omega}$  and use **the inner regularity of Lebesgue measure** to conclude that

$$\limsup_{N \rightarrow \infty} N^{-(1+2/d)} \sum_{i=1}^N \mu_i(-\Delta_{|D}) \leq \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}}.$$

This completes the proof of Weyl's law for the sum of eigenvalues of  $-\Delta_{|D}$ .  $\square$

The above formula for the sum of eigenvalues can be rewritten in the following dual form.

**Exercise.** Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Prove that the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  satisfy

$$\sum_{i=1}^{\infty} [\mu_i - \lambda]_- = -L_{1,d}^{\text{cl}} |\Omega| \lambda^{1+\frac{d}{2}} + o\left(\lambda^{1+\frac{d}{2}}\right)_{\lambda \rightarrow \infty}$$

Remark: Heuristically, the formula in the above exercise is consistent with the semiclassical formula

$$\text{Tr}_{L^2(\mathbb{R}^3)}(-\Delta + \lambda V)_- \approx \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|2\pi k|^2 + \lambda V(x))_- dk dx = -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{1+d/2}$$

in the case  $V \equiv -1$  on  $\Omega$  and  $\geq 0$  (even  $+\infty$ ) elsewhere.

## 5.3 Distribution of eigenvalues

Now we come to the asymptotic behavior of a single eigenvalue, which goes back to the original result of **Weyl** in 1911.

**Theorem** (Weyl's law for distribution of eigenvalues). Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Then the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  satisfy

$$\mu_N = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}} |\Omega|^{\frac{2}{d}}} N^{\frac{2}{d}} + o(N^{\frac{2}{d}})_{N \rightarrow \infty}.$$

Equivalently, if we denote by  $N(\lambda)$  the number of eigenvalues  $< \lambda$ , then

$$N(\lambda) = \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}} + o(\lambda^{\frac{d}{2}})_{\lambda \rightarrow \infty}.$$

Remark: The above formula of  $N(\lambda)$  is consistent with the semiclassical formula

$$\text{Tr } \mathbf{1}(-\Delta + \lambda V < 0) \approx \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}(|2\pi k|^2 + \lambda V(x) < 0) dk dx = L_{0,d}^{\text{cl}} \int_{\mathbb{R}^d} |\lambda V_-|^{d/2}, \quad L_{0,d}^{\text{cl}} = \frac{|B_1|}{(2\pi)^d}$$

with  $V \equiv -1$  on  $\Omega$  and  $\geq 0$  (even  $+\infty$ ) elsewhere.

We will derive the above theorem using Weyl's law for the sum of eigenvalues and a simple Tauberian lemma.

**Lemma** (Tauberian). *Given any increasing sequence  $0 \leq \mu_1 \leq \mu_2 \leq \dots$  and two constants  $A > 0$ ,  $a > 0$ . Then*

$$\lim_{N \rightarrow \infty} N^{-1-a} \sum_{n=1}^N \mu_n = A \iff \lim_{N \rightarrow \infty} N^{-a} \mu_N = A(1+a).$$

*Proof.* Assume that

$$S_N := \sum_{n=1}^N \mu_n = AN^{1+a} + o(N^{1+a})_{N \rightarrow \infty}.$$

Then for every constant  $\varepsilon \in (0, 1)$ , with  $N$  sufficiently large and  $m \in [\varepsilon N, \varepsilon N + 1)$  we have

$$|S_N - AN^{1+a}| \leq \varepsilon^2 N^{1+a}, \quad |S_{N+m} - A(N+m)^{1+a}| \leq \varepsilon^2 (N+m)^{1+a}.$$

Therefore,

$$\begin{aligned} \mu_N &\leq \frac{\mu_{N+1} + \mu_{N+2} + \dots + \mu_{N+m}}{m} = \frac{S_{N+m} - S_N}{m} \\ &\leq \frac{A(N+m)^{1+a} - AN^{1+a} + \varepsilon^2(N+m)^{1+a} + \varepsilon^2 N^{1+a}}{m} \\ &\leq \frac{A(N+\varepsilon N+1)^{1+a} - AN^{1+a} + 2\varepsilon^2(N+\varepsilon N+1)^{1+a}}{\varepsilon N} \\ &\leq N^a \left[ A \frac{(1+\varepsilon+N^{-1})^{1+a} - 1}{\varepsilon} + 2\varepsilon(1+\varepsilon+N^{-1})^{1+a} \right]. \end{aligned}$$

Thus

$$\limsup_{N \rightarrow \infty} N^{-a} \mu_N \leq A \frac{(1+\varepsilon)^{1+a} - 1}{\varepsilon} + 2\varepsilon(1+\varepsilon)^{1+a}.$$

Sending  $\varepsilon \rightarrow 0^+$  we obtain

$$\limsup_{N \rightarrow \infty} N^{-a} \mu_N \leq A(1+a).$$

Similarly, using

$$\mu_N \geq \frac{\mu_N + \mu_{N-1} + \dots + \mu_{N-m+1}}{m} \frac{S_N - S_{N-m}}{m}$$

we find that

$$\liminf_{N \rightarrow \infty} N^{-a} \mu_N \geq A(1+a).$$

Thus

$$\lim_{N \rightarrow \infty} N^{-a} \mu_N = A(1+a).$$

The reverse direction is left as an exercise. □

**Exercise.** Given an increasing sequence  $0 \leq \mu_1 \leq \mu_2 \leq \dots$  satisfying

$$\lim_{N \rightarrow \infty} N^{-a} \mu_N = A(1+a)$$

for two constants  $A > 0$ ,  $a > 0$ . Prove that

$$\lim_{N \rightarrow \infty} N^{-1-a} \sum_{n=1}^N \mu_n = A.$$

**Proof of Weyl's law for the distribution of eigenvalues.** In the previous section we have proved that

$$\lim_{N \rightarrow \infty} N^{-1-\frac{2}{d}} \sum_{i=1}^N \mu_i = \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}}.$$

Therefore, the Tauberian lemma implies that

$$\lim_{N \rightarrow \infty} N^{-\frac{2}{d}} \mu_N = \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} \left(1 + \frac{2}{d}\right) = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}} |\Omega|^{\frac{2}{d}}}.$$

Now consider  $N(\lambda)$  the number of eigenvalues  $< \lambda$ . By definition of  $N(\lambda)$ , we have

$$\mu_{N(\lambda)} < \lambda \leq \mu_{N(\lambda)+1}.$$

Of course, when  $\lambda \rightarrow \infty$  then  $N(\lambda) \rightarrow \infty$ . Hence by the asymptotic formula of  $\mu_N$  for  $N$

large,

$$\lambda = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{\frac{2}{d}}} N(\lambda)^{\frac{2}{d}} + o(N(\lambda)^{\frac{2}{d}})_{\lambda \rightarrow \infty}$$

which is equivalent to

$$N(\lambda) = \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}} + o(\lambda^{\frac{d}{2}}).$$

□

## 5.4 Pólya conjecture

Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Recall Weyl's law for the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$ :

$$\mu_N = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{\frac{2}{d}}} N^{\frac{2}{d}} + o(N^{\frac{2}{d}})_{N \rightarrow \infty}.$$

An important open problem is **Pólya's conjecture**:

$$\mu_N \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} N^{\frac{2}{d}}, \quad \forall N \geq 1$$

which is equivalent to (**why?**)

$$N(\lambda) \leq \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}}, \quad \forall \lambda > 0.$$

**The Berezin-Li-Yau inequality follows from Pólya's conjecture.** Clearly

$$\sum_{n=1}^N \mu_n \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} \sum_{n=1}^N n^{\frac{2}{d}} \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} \int_0^N t^{\frac{2}{d}} dt = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} \left(1 + \frac{2}{d}\right) N^{1+\frac{2}{d}} = \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}}.$$

However, obtaining the sharp lower bound for **every eigenvalue** is much more difficult. Nevertheless, using the Berezin-Li-Yau inequality we get the non-optimal bound

$$\mu_N \geq \frac{1}{N} \sum_{n=1}^N \mu_n \geq \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{\frac{2}{d}} = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} \frac{d}{d+2} N^{\frac{2}{d}}, \quad \forall N \geq 1.$$

□

**Proof of Pólya's conjecture for cubes.** Pólya's conjecture can be verified easily for cubes.



By a simple scaling argument, it suffices to consider the case  $\Omega = [0, \pi]^d$  where the eigenvalues are given explicitly by

$$\{|x|^2 = (x_1^2 + \dots + x_d^2) \mid x = (x_1, \dots, x_d) \in \mathbb{N}^d\}.$$

A key observation is that the number of integer points inside a ball can be controlled by the volume of the ball. More precisely, any point  $x = (x_1, \dots, x_d) \in \mathbb{N}^d$  can be associated with the unit cube

$$Q_x = (x_1, x_1 - 1) \times (x_2, x_2 - 1) \times \dots \times (x_d, x_d - 1).$$

Since  $Q_x \cap Q_y = \emptyset$  if  $x \neq y$  and  $Q_x \subset B_R(0) \cap \mathbb{R}_+^d$  if  $x \in \mathbb{N}^d \cap B_R(0)$ , we find that

$$|\mathbb{N}^d \cap B_R(0)| = \left| \bigcup_{x \in \mathbb{N}^d \cap B_R(0)} Q_x \right| \leq 2^{-d} |B_R| = 2^{-d} R^d |B_1|.$$

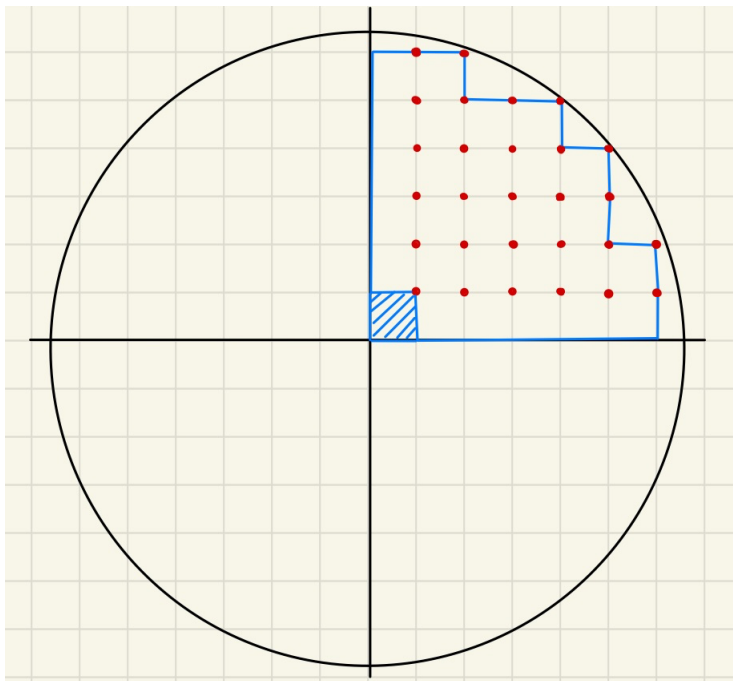


Figure: Positive integer points inside a circle

On the other hand, since  $\{\mu_n\}$  is an increasing sequence, there must be at least  $N$  points inside  $|\mathbb{N}^d \cap B_R(0)|$  with  $R = \sqrt{\mu_N}$ . Thus

$$N \leq 2^{-d} \mu_N^{\frac{d}{2}} |B_1| \iff \mu_N \geq \frac{2^2}{|B_1|^{\frac{2}{d}}} N^{\frac{2}{d}} = \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}} |\Omega|^{\frac{2}{d}}} N^{\frac{2}{d}}, \quad \forall N \geq 1.$$

□

In 1961, Pólya proved

**Theorem (Pólya).** *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Assume that  $\Omega$  is a **tiling domain**, namely we can cover  $\mathbb{R}^d$  (up to a set of 0 measure) by a union of disjoint copies of  $\Omega$  (each copy is obtained from by  $\Omega$  up to translation, rotation and reflection). Then the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  satisfy*

$$\mu_N \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} N^{\frac{2}{d}}, \quad \forall N \geq 1.$$

Remark: A cube is a tiling domain, but a ball is not (this case remains open).

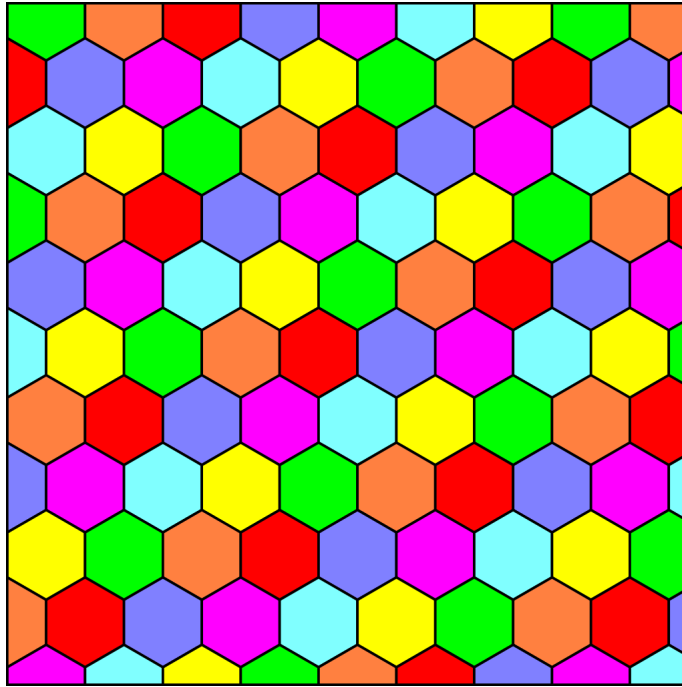


Figure: Tiling by hexagonal

*Proof.* Let us denote by  $\mu_k(\Omega)$  the  $k$ -th eigenvalue of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$ .

**Step 1.** Assuming that we can put  $N$  disjoint copies  $\{\Omega_n\}_{n=1}^N$  of  $\Omega$  inside a large cube  $Q \subset \mathbb{R}^d$ . We will prove

$$\mu_k(\Omega) \geq \mu_{kN}(Q), \quad \forall k = 1, 2, \dots$$

Take  $k \geq 1$ . By the min-max principle, for every  $\varepsilon > 0$  we can find a subspace  $M_k(\Omega) \subset C_c^\infty(\Omega)$  such that  $\dim M_k(\Omega) = k$  and

$$\mu_k(\Omega) \geq \sup_{\substack{u \in M_k(\Omega) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 - \varepsilon.$$

Since each  $\Omega_n$  is a copy of  $\Omega$ , we find that  $\mu_k(\Omega_n) = \mu_k(\Omega)$  and we can also find a subspace  $M_k(\Omega_n) \subset C_c^\infty(\Omega)$  such that  $\dim M_k(\Omega_n) = k$  and

$$\mu_k(\Omega) = \mu_k(\Omega_n) \geq \sup_{\substack{u \in M_k(\Omega_n) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 - \varepsilon.$$

Note that the functions in  $M_k(\Omega_n)$  have disjoint supports to the functions in  $M_k(\Omega_m)$  if  $n \neq m$ . Therefore, the space

$$M = \bigoplus_{n=1}^N M_k(\Omega_n) = \text{Span} \left\{ u \in \bigcup_{n=1}^N M_k(\Omega_n) \right\} \subset C_c^\infty(Q)$$

has  $\dim M = kN$ . Then by the min-max principle,

$$\mu_{kN}(\Omega) \leq \sup_{\substack{u \in M \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 \leq \sup_{1 \leq n \leq N} \sup_{\substack{u \in M_k(\Omega_n) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 \leq \mu_k(\Omega) + \varepsilon.$$

To see the second inequality, we can write any vector  $\varphi \in M$  as

$$\varphi = \sum_{n=1}^n \varphi_n, \quad \varphi_n \in M_k(\Omega_n).$$

Since  $\{\varphi_n\}_{n=1}^N$  has disjoint supports, we find that

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla \varphi|^2 &= \sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla \varphi_n|^2 \leq \sum_{n=1}^N \left( \int_{\mathbb{R}^d} |\varphi_n|^2 \right) \sup_{\substack{u \in M_k(\Omega_n) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 \\ &\leq \sup_{1 \leq n \leq N} \sup_{\substack{u \in M_k(\Omega_n) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 \sum_{n=1}^N \left( \int_{\mathbb{R}^d} |\varphi_n|^2 \right) = \sup_{1 \leq n \leq N} \sup_{\substack{u \in M_k(\Omega_n) \\ \|u\|_{L^2} = 1}} \|\nabla u\|_{L^2}^2 \int_{\mathbb{R}^d} |\varphi|^2 \end{aligned}$$

Thus we have proved that

$$\mu_{kN}(\Omega) \leq \mu_k(\Omega) + \varepsilon.$$

Sending  $\varepsilon \rightarrow 0$  we find the desired inequality  $\mu_{kN}(Q) \leq \mu_k(\Omega)$ .

**Step 2.** Since Pólya's conjecture holds for cube, we have

$$\mu_k(\Omega) \geq \mu_{kN}(Q) \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}} |Q|^{2/d}} (kN)^{\frac{2}{d}}, \quad \forall k \geq 1.$$

This inequality holds for all  $N \geq 1$  such that we can put at least  $N$  disjoint copies of  $\Omega$  inside the cube  $Q \subset \mathbb{R}^d$ . Since  $\Omega$  is a tiling domain, we can choose a very big cube  $Q_N \subset \mathbb{R}^d$  such that we can put  $N$  disjoint copies of  $\Omega$  inside  $Q_N$  and at the same time

$$\lim_{N \rightarrow \infty} \frac{|Q_N|}{N|\Omega|} = 1.$$

Thus for every  $k \geq 1$ , we have

$$\mu_k(\Omega) \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|Q_N|^{2/d}}(kN)^{\frac{2}{d}} \rightarrow_{N \rightarrow \infty} \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}}k^{\frac{2}{d}}, \quad \forall k \geq 1$$

This completes the proof. □

Remark: For a general domain  $\Omega$ , by using the above proof, we find that

$$\mu_k(\Omega) \geq R(\Omega)^{2/d} \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} k^{\frac{2}{d}}, \quad \forall k \geq 1.$$

where  $R(\Omega) \in (0, 1]$  is the **packing density** of  $\Omega$ , namely the largest fraction of the space  $\mathbb{R}^d$  that we can cover by disjoint copies of  $\Omega$ . Determination the packing density of a ball is the standard **packing problem**.

**2D Packing problem.** For a disc  $\Omega = \{x \in \mathbb{R}^2 : |x| < 1\}$ , the packing density is

$$R(\Omega) = \frac{\pi}{\sqrt{12}} \approx 0.9069\dots$$

and it is achieved by the “hexagonal packing arrangement”. This optimality was proved by **Lagrange** in 1773 for “lattice packings”, by **Gauss** in 1831 for “periodic packings”, and finally by **Tóth** in 1940 for the general case.

**3D Packing problem.** For a ball  $\Omega = \{x \in \mathbb{R}^3 : |x| < 1\}$ , the “sphere packing problem” is more difficult. In 1611, **Kepler** conjectured that the optimal parking arrangement is obtained by a family of “close-packed structures”, leading to

$$R(\Omega) = \frac{\pi}{3\sqrt{2}} \approx 0.74048\dots$$

This optimality was proved by **Gauss** in 1831 for “lattice packings” and by **Hales** in 1998 for the general case but his proof is involving a heavy computer checking of many individual

cases. Finally, a formal proof was published in 2017 by Hales and collaborators (Forum of Mathematics, Pi).

**8D & 24D Packing problems.** In 2016, **Viazovska** published a surprisingly short solution for the parking problem in 8 dimensions. Shortly later, she and collaborators solved the problem in 24 dimensions.



Figure: Circle packing and “Sphere packing”

## 5.5 Weyl’s conjecture

Recall that Weyl’s law states that the number of eigenvalues  $< \lambda$  of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  satisfies

$$N(\lambda) = L_{0,d}^{\text{cl}} |\Omega| \lambda^{\frac{d}{2}} + o(\lambda^{\frac{d}{2}})_{\lambda \rightarrow \infty}, \quad L_{0,d}^{\text{cl}} = \frac{|B_{\mathbb{R}^d}(0, 1)|}{(2\pi)^d}$$

**Weyl’s conjecture** (1911) states that the second order term is involving  $|\partial\Omega|$

$$N(\lambda) = L_{0,d}^{\text{cl}} |\Omega| \lambda^{\frac{d}{2}} - \frac{1}{4} L_{0,d-1}^{\text{cl}} |\partial\Omega| \lambda^{\frac{d-1}{2}} + o(\lambda^{\frac{d-1}{2}})_{\lambda \rightarrow \infty}.$$

(The second order term is negative, so it is consistent with Pólya’s conjecture). This is a hard problem. A proof of Weyl’s conjecture for a class of smooth domains was given by **Ivrii** in 1980 (if you really love semiclassical approximation, check Ivrii’s Monsterbook (2007-2019)).

**Proof of Weyl’s conjecture for squares.**

Recall that for  $\Omega = [0, \pi]^2$ , the eigenvalues of the Dirichlet Laplacian  $-\Delta_D$  on  $L^2(\Omega)$  are given explicitly by

$$\{|x|^2 = (x_1^2 + x_2^2) \mid x = (x_1, x_2) \in \mathbb{N}^2\}.$$

Thus

$$N(\lambda) = \frac{1}{4}(S(\sqrt{\lambda}) - 4[\sqrt{\lambda}] - 1)$$

where the **floor function**  $[t]$  is the integer part of  $t$  (i.e.  $[t] \leq t < [t] + 1 \in \mathbb{N}$ ) and

$$S(R) := \text{number of integer points inside } B(0, R).$$

The Weyl's conjecture tells us that for  $\Omega = [0, \pi]^2$

$$\begin{aligned} N(\lambda) &= L_{0,d}^{\text{cl}}|\Omega|\lambda^{\frac{d}{2}} - \frac{1}{4}L_{0,d-1}^{\text{cl}}|\partial\Omega|\lambda^{\frac{d-1}{2}} + o(\lambda^{\frac{d-1}{2}})_{\lambda \rightarrow \infty} \\ &= \frac{1}{4}\pi\lambda - \lambda^{1/2} + o(\lambda^{1/2})_{\lambda \rightarrow \infty} \end{aligned}$$

which is equivalent to

$$S(R) = \pi R^2 + o(R)_{R \rightarrow \infty}.$$

The asymptotic estimate for  $|S(R) - \pi R^2|$  when  $R \rightarrow \infty$  is called the **Gauss circle problem** (1801). **Hardy conjectured** that

$$|S(R) - \pi R^2| \leq O(R^{1/2+\varepsilon})_{R \rightarrow \infty}, \quad \forall \varepsilon > 0.$$

The lower bound is sharp since **Hardy** and **Landau** independently showed that the error cannot be better than  $O(R^{1/2} \ln(R))$ . On the other hand, the upper bound remains open. **Gauss** himself managed to prove that

$$|S(R) - \pi R^2| \leq 2\sqrt{2}\pi R.$$

**Sierpinski** (1905) proved that

$$|S(R) - \pi R^2| \leq O(R^\theta)_{R \rightarrow \infty}, \quad \forall \theta > \frac{2}{3} \approx 0.66666\dots$$

**Huxley** (2003) proved the currently best bound

$$|S(R) - \pi R^2| \leq O(R^\theta)_{R \rightarrow \infty}, \quad \theta = \frac{131}{208} \approx 0.62981\dots$$

Now let us prove  $|S(R) - \pi R^2| \leq o(R)$  which is necessary to justify Weyl's conjecture for squares. This means we need to improve Gauss' bound, which is a nontrivial task.

We start with

**Theorem** (Poisson summation formula). *If  $f \in C_c^\infty(\mathbb{R}^d)$ , then*

$$\sum_{x \in \mathbb{Z}^d} f(x) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k).$$

In general, this formula holds if  $f$  is sufficiently smooth and decays sufficiently fast (e.g.  $f \in C^\infty$  and it decays faster than any polynomial).

*Proof.* Consider the function

$$F(x) = \sum_{y \in \mathbb{Z}^d} f(x + y).$$

Then  $F : \mathbb{R}^d \rightarrow \mathbb{C}$  is a periodic function (of period 1) and it can be written by the Fourier series

$$F(x) = \sum_{k \in \mathbb{Z}^d} a_k e^{2\pi i k \cdot x}$$

where

$$\begin{aligned} a_k &= \int_{[0,1]^d} F(x) e^{-2\pi i k \cdot x} dx = \sum_{y \in \mathbb{Z}^d} \int_{[0,1]^d} f(x + y) e^{-2\pi i k \cdot x} dx \\ &= \sum_{y \in \mathbb{Z}^d} \int_{y+[0,1]^d} f(x) e^{-2\pi i k \cdot x} dx = \int_{\mathbb{R}^d} f(x) e^{-2\pi i k \cdot x} dx = \widehat{f}(k). \end{aligned}$$

Thus

$$\sum_{y \in \mathbb{Z}^d} f(x + y) = F(x) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e^{2\pi i k \cdot x}.$$

In particular, taking  $x = 0$  we obtain

$$\sum_{y \in \mathbb{Z}^d} f(y) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k).$$

□

Remark: Strictly speaking, the formula

$$\mathbb{1}_{[0,1]^d} F(x) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e^{2\pi i k \cdot x}$$

follows from Parseval's identity on  $L^2([0,1]^d)$ , so in principle it holds pointwise *almost everywhere*. However, it holds pointwise *everywhere* because of the continuity of the functions in both sides, and extends globally because of the periodicity. Thus all we need is that  $F$  and  $\sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e^{2\pi i k \cdot x}$  are continuous. The condition  $f \in C_c^\infty(\mathbb{R}^d)$  implies that  $\widehat{f}(k)$  decays faster than any polynomial when  $|k| \rightarrow \infty$  (which can be seen by integration by part).

Now we prove

**Theorem.** *Let  $S(R)$  be the number of integer points inside  $B(0, R) \subset \mathbb{R}^2$ . Then*

$$|S(R) - \pi R^2| \leq O(R^{2/3})_{R \rightarrow \infty}.$$

*Proof.* The proof is due to Hugh Montgomery.

**Step 1.** Note that

$$S(R) = \sum_{x \in \mathbb{Z}^2} \mathbb{1}_{B_R(0)}(x).$$

We want to use the Poisson summation formula, and hence we will replace  $\mathbb{1}_{B_R(0)}$  by smooth functions. Fix a radial function

$$0 \leq \varphi_1 \in C_c^\infty(B_1(0)), \quad \int_{\mathbb{R}^d} \varphi_1 = 1$$

and define for some parameter  $1 \gg h > 0$

$$\varphi_h(x) = h^{-2} \varphi_1(x/h), \quad 0 \leq \varphi_h \in C_c^\infty(\mathbb{R}^2), \quad \int_{\mathbb{R}^d} \varphi_h = 1.$$

For every  $r > 0$  we denote

$$f_r(x) = \varphi_h * \mathbb{1}_{B_r(0)}(x) = \int_{B_r(0)} \varphi_h(x-y) dy.$$



Since  $f_r \in C_c^\infty$ , we have Poisson summation formula

$$\tilde{S}(r) := \sum_{x \in \mathbb{Z}^2} f_r(x) = \sum_{k \in \mathbb{Z}^2} \widehat{f}_r(k).$$

Moreover, clearly

$$f_{R-h} \leq \mathbb{1}_{B_R(0)} \leq f_{R+h} \implies \tilde{S}(R-h) \leq S(R) \leq \tilde{S}(R+h)$$

**Step 2.** It remains to estimate  $|\tilde{S}(r) - \pi r^2|$ . In the sum  $\tilde{S}(r) = \sum_{k \in \mathbb{Z}^2} \widehat{f}_r(k)$  the main contribution comes from  $k = 0$ , namely

$$\widehat{f}_r(0) = \int_{\mathbb{R}^2} f_r = \int_{\mathbb{R}^2} \varphi_h \int_{\mathbb{R}^2} \mathbb{1}_{B_r(0)} = \pi r^2.$$

Thus it remains to bound  $\sum_{k \neq 0} \widehat{f}_r(k)$ . For  $k \neq 0$ , we have

$$\widehat{f}_r(k) = \widehat{\varphi}_h(k) \widehat{\mathbb{1}_{B_r(0)}}(k) = \widehat{\varphi}_1(hk) r^2 \widehat{\mathbb{1}_{B_1(0)}}(rk) = \widehat{\varphi}_1(hk) r^2 \frac{J_1(2\pi r|k|)}{r|k|}$$

with the **Bessel function**

$$J_1(t) = \frac{1}{i\pi} \int_0^\pi e^{it \cos \theta} \cos \theta d\theta.$$

Here we used the fact that

$$\begin{aligned} \widehat{\varphi}_h(k) &= \int_{\mathbb{R}^2} e^{-2\pi i k \cdot x} \varphi_h(x) dx = \int_{\mathbb{R}^2} e^{-2\pi i (kh) \cdot (x/h)} h^{-2} \varphi_1(x/h) dx \\ &= \int_{\mathbb{R}^2} e^{-2\pi i (kh) \cdot y} \varphi_1(y) dy = \widehat{\varphi}_1(hk) \end{aligned}$$

(with the variable  $y = x/h$ ) and similarly

$$r^{-2} \widehat{\mathbb{1}_{B_r(0)}}(k) = \widehat{\varphi}_1(rk).$$

Moreover, since  $\mathbb{1}_{B_1(0)}$  is radial, its Fourier transform is also radial and can be computed as

$$\begin{aligned} \widehat{\mathbb{1}_{B_1(0)}}(k) &= \widehat{\mathbb{1}_{B_1(0)}}(|k|, 0) = \iint_{|x_1|^2 + |x_2|^2 < 1} e^{-2\pi i |k| x_1} dx_1 dx_2 = \int_{-1}^1 e^{-2\pi i |k| x_1} 2\sqrt{1-x_1^2} dx_1 \\ &= \int_0^\pi e^{2\pi i |k| \cos \theta} 2(\sin \theta)^2 d\theta \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{\pi i |k|} \int_0^\pi \frac{d}{d\theta} \left( e^{2\pi i |k| \cos \theta} \right) \sin \theta d\theta \\
&= \frac{1}{\pi i |k|} \int_0^\pi e^{2\pi i |k| \cos \theta} \cos \theta d\theta = \frac{J_1(2\pi |k|)}{|k|}.
\end{aligned}$$

Here we have changed the variable  $x_1 = -\cos \theta$  and used the integration by part. We need the following bound.

**Exercise.** Prove that  $J_1(t) \leq Ct^{-1/2}$  for all  $t > 0$ .

Moreover, since  $\varphi_1 \in C_c^\infty$ ,  $\widehat{\varphi}_1 \in L^\infty$  and it decays faster than any polynomial. Thus

$$|\widehat{f}_r(k)| = |\widehat{\varphi}_1(hk)| \frac{r}{|k|} J_1(r|k|) \leq \frac{Cr^{1/2}}{(h|k|)^\ell |k|^{3/2}}, \quad \forall \ell \geq 0.$$

Take a cut-off  $K > 0$ . We have

$$\sum_{0 < |k| \leq K} |\widehat{f}_r(k)| \leq \sum_{0 < |k| \leq K} \frac{Cr^{1/2}}{|k|^{3/2}} \leq Cr^{1/2} K^{1/2}.$$

and

$$\sum_{|k| > K} |\widehat{f}_r(k)| \leq \sum_{|k| > K} \frac{Cr^{1/2}}{h|k|^{5/2}} \leq C \frac{r^{1/2}}{hK^{1/2}}.$$

Thus

$$\sum_{k \neq 0} |\widehat{f}_r(k)| \leq Cr^{1/2} K^{1/2} + C \frac{r^{1/2}}{hK^{1/2}}.$$

Optimizing over  $K > 0$  we find that

$$|\widetilde{S}(r) - \pi r^2| = \sum_{k \neq 0} |\widehat{f}_r(k)| \leq C \frac{r^{1/2}}{h^{1/2}}.$$

**Step 3.** We have

$$\pi(R-h)^2 - C \frac{(R-h)^{1/2}}{h^{1/2}} \leq \widetilde{S}(R-h) \leq S(R) \leq \widetilde{S}(R+h) \leq \pi(R+h)^2 + C \frac{(R+h)^{1/2}}{h^{1/2}}$$

and hence

$$|S(R) - \pi R^2| \leq CRh + C \frac{R^{1/2}}{h^{1/2}}$$

Optimizing over  $h$  (i.e. taking  $h \sim R^{-1/3}$ ) we obtain

$$|S(R) - \pi R^2| \leq CR^{2/3}.$$

□

## 5.6 Can one hear the shape of a drum?

So far, we have seen that from the eigenvalues of the Dirichlet Laplacian on  $L^2(\Omega)$ , we can reconstruct some geometric properties of  $\Omega$ , e.g.  $|\Omega|$  or  $|\partial\Omega|$ . Thus it is a natural question that: can one completely determine  $\Omega$  (up to usual symmetries) from all of its Dirichlet eigenvalues? This was made popular by **Kac** (1966) in his paper “Can one hear the shape of a drum?”

A **negative answer** was found immediately by **Milnor** who proved the existence of two tori in 16-dimensions which have the same eigenvalues but with different shapes. However, counter examples in 2 dimensions was found only in 1992 by **Gordon-Webb–Wolpert**, using a sufficient condition for isospectrality by **Sunada**. In the following, we will represent a 2D example by **Buser-Conway-Doyle-Semmler** (1994) via their **transplantation method** (see **Okada-Shudo** (2001) for the relation between two methods).

**2D example by transplantation method.** Clearly the two domains  $\Omega_1$  and  $\Omega_2$  in the figure below have different shapes. Let us prove that they have the same Dirichlet eigenvalues. Let  $u$  be an eigenfunction on  $\Omega_1$  with eigenvalue  $\mu$ , namely

$$-\Delta u = \mu u \quad \text{on } \Omega_1, \quad u|_{\partial\Omega_1} = 0.$$

Let us construct an eigenfunction  $v$  on  $\Omega_2$  with the same eigenvalue  $\mu$  as follows.

- We divide each domain into 7 congruent triangles.
- We decompose  $u = \sum_{i=1}^7 u_i$  where  $u_i$  is the restriction on the  $i$ -th triangle of  $\Omega_1$ . Similarly, we decompose  $v = \sum_{i=1}^7 v_i$  on  $\Omega_2$ .
- Each  $v_i$  is the linear combination of three  $u_j$ 's, for example  $v_1 = u_5 - u_7 - u_4$ .

The full information of the definition of  $v_i$ 's can be found in the figure below. The rules are:

- On  $\Omega_1$ , the triangles with red numbers (1, 3, 5) are oriented clockwise, and the triangles with blue numbers (2, 4, 6, 7) are oriented counter-clockwise.
- On  $\Omega_2$ , if a number comes with + sign, then it keeps its orientation as in  $\Omega_1$ ; and if a number comes with - sign, then it changes its orientation from  $\Omega_1$ . For example, +1 and -2 are clockwise, -1 and +2 are counter-clockwise.
- We transform the functions  $u_i$  from triangles of  $\Omega_1$  to put in triangles in  $\Omega_2$ . In each triangle in  $\Omega_2$ , if the number  $i$  comes with + (i.e. it has the same orientation as in  $\Omega_1$ ), we simply transform  $u_i$  by translation and rotation; however if the number  $i$  comes with - (i.e. it changes the orientation), then we also transform  $u_i$  by reflection.

For example, in the first triangle of  $\Omega_2$ , we see  $5 - 7 - 4$ . This means in that triangle we set

$$v_1 = u_5 - u_7 - u_4$$

where  $u_5$  is transform directly from  $\Omega_1$ , while  $u_7$  and  $u_4$  are also reflected.

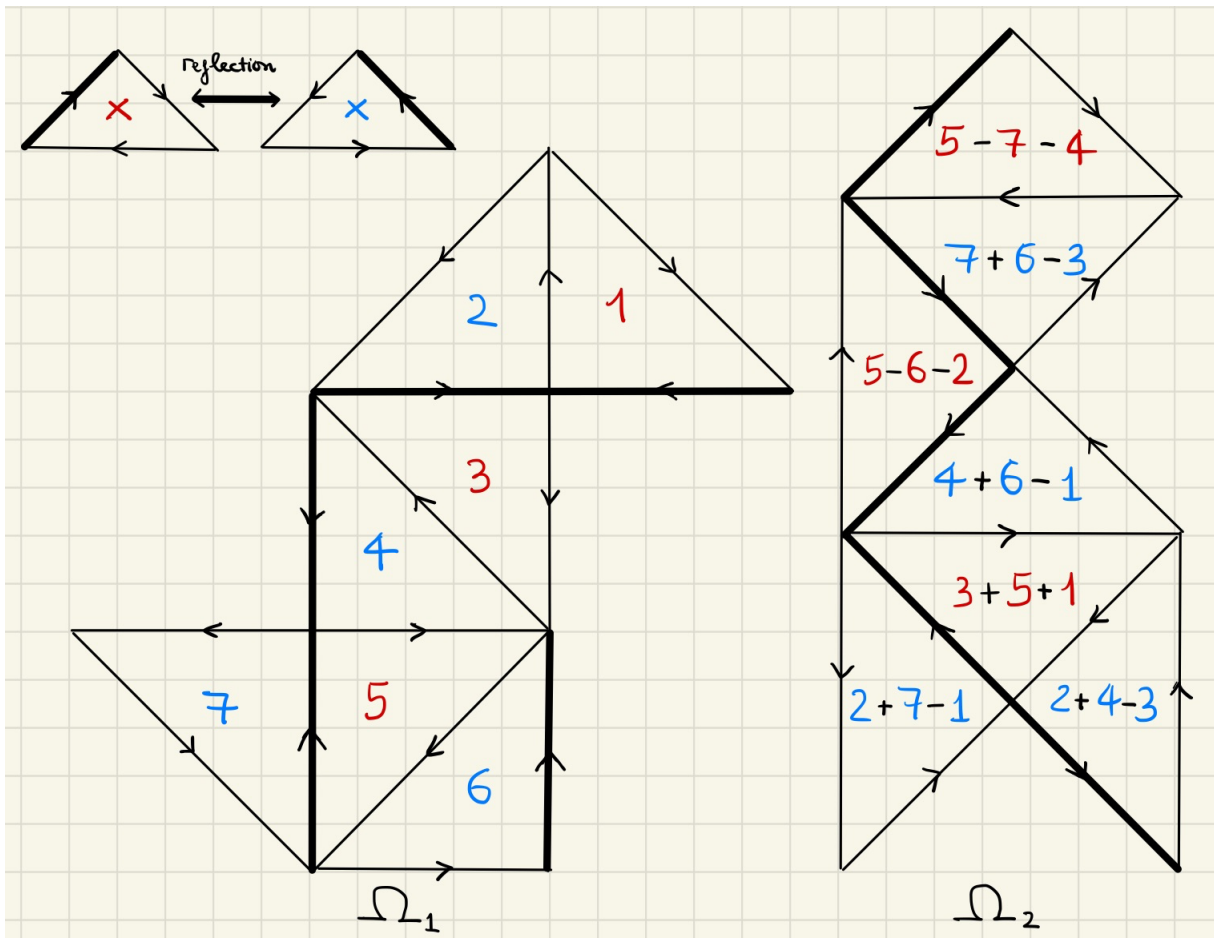


Figure: Two isospectral domains with different shapes

Why  $v$  constructed this way is an eigenfunction on  $\Omega_2$  with eigenvalue  $\mu$ ?

- First, we have  $\Delta u = \mu u$  in the interior part of each triangle in  $\Omega_1$ , so obviously  $\Delta v = \mu v$  on the interior part of each triangle of  $\Omega_2$ .
- Second, we have to check that  $v$  vanishes on the boundary of  $\Omega_2$ .
- Third, we have to check that  $v$  is smooth inside  $\Omega_2$ , namely it is smoothly connected on the joint boundary of the triangles of  $\Omega_2$ .

Finally, we have to check that the multiplicity of eigenvalues are the same in  $\Omega_1$  and  $\Omega_2$ . This means that the transplantation is invertible. Fortunately, this is true for this example.

On the other direction, **Zelditch** proved a **positive answer** for 2D sets which are **convex with analytic boundary**. An **open question** is that whether the convexity can be relaxed?

# Chapter 6

## Neumann Laplacian

**Definition** (Neumann Laplacian). *Let  $\Omega$  be an open set in  $\mathbb{R}^d$ . The **Neumann Laplacian**  $-\Delta_N$  on  $L^2(\Omega)$  is defined by Friedrich's extension via the quadratic form formula*

$$\langle u, -\Delta_N u \rangle = \int_{\Omega} |\nabla u|^2 \geq 0, \quad \forall u \in C^\infty(\mathbb{R}^d).$$

*The quadratic form domain of  $-\Delta_N$  is  $H^1(\Omega)$ .*

- Recall that  $H^1(\Omega)$  is a Hilbert space with the norm

$$\|u\|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\Omega)}^2.$$

Here the derivatives are interpreted in the **distributional sense**, namely

$$\int_{\Omega} (\partial_{x_d} u) \varphi dx = - \int_{\Omega} u \partial_{x_d} \varphi dx, \quad \forall \varphi \in C_c^\infty(\Omega).$$

More generally, we can define

$$H^m(\Omega) = \{u \in L^2(\Omega), \quad D^\alpha u \in L^2(\Omega) \text{ for all } |\alpha| \leq m\}$$

where the weak derivatives are defined by

$$\int_{\Omega} (D^\alpha u) \varphi = (-1)^{|\alpha|} \int_{\Omega} u D^\alpha \varphi, \quad \forall \varphi \in C_c^\infty(\Omega).$$

- If  $\mu$  is an eigenvalue of the Neumann Laplacian with eigenfunction  $u$

$$-\Delta_N u = \mu u$$

then by the definition of the quadratic form,

$$\int_{\Omega} \nabla f \cdot \nabla u = \mu \int_{\Omega} f u, \quad \forall f \in H^1(\Omega).$$

On the other hand, by simply integrating the equation  $-\Delta_N u = \mu u$  and using Green's formula we have

$$\mu \int_{\Omega} f u = \int_{\Omega} f (-\Delta_N u) = \int_{\Omega} \nabla f \cdot \nabla u - \int_{\partial\Omega} f \frac{\partial u}{\partial \mathbf{n}} d\sigma$$

where  $d\sigma$  is the surface measure on the boundary  $\partial\Omega$  and

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n}$$

is called the **outward normal derivative**. Here  $\mathbf{n} = (\mathbf{n}_1, \dots, \mathbf{n}_d)$  is the unit normal vector to  $\partial\Omega$ . Thus we obtain

$$\int_{\partial\Omega} f \frac{\partial u}{\partial \mathbf{n}} d\sigma = 0, \quad \forall f \in C^2(\mathbb{R}^d)$$

which implies that

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega.$$

The latter is called the **Neumann boundary condition**. In comparison, the **Dirichlet boundary condition** is  $u = 0$  on  $\partial\Omega$ .

- It is easy to see that the eigenvalues (or min-max values) of the Neumann Laplacian  $-\Delta_N$  are smaller than that of the Dirichlet Laplacian  $-\Delta_D$ . We always have  $\mu_1(-\Delta_N) = 0 < \mu_1(-\Delta_D)$ . More generally, by the min-max principle, for all  $N \geq 1$  we have

$$\mu_n(-\Delta_D) = \inf_{\substack{M \subset H_0^1(\Omega) \\ \dim M = n}} \max_{\substack{u \in M \\ \|u\|_{L^2} = 1}} \int_{\Omega} |\nabla u|^2 \geq \inf_{\substack{M \subset H^1(\Omega) \\ \dim M = n}} \max_{\substack{u \in M \\ \|u\|_{L^2} = 1}} \int_{\Omega} |\nabla u|^2 = \mu_n(-\Delta_N)$$

If  $-\Delta_N$  has compact resolvent, then all  $\mu_n(-\Delta_N)$  are eigenvalues, and the stronger inequality  $\mu_n(-\Delta_D) \geq \mu_{n+1}(-\Delta_N)$  holds (called **Friedlander's inequality**).

## 6.1 Essential spectrum vs. Compact resolvent

Unlike the Dirichlet Laplacian on any bounded set always has compact resolvent, the boundedness of the domain is not enough to ensure that the Neumann Laplacian has compact resolvent. Putting differently, when  $\Omega$  is bounded, the embedding  $H_0^1(\Omega) \subset L^2(\Omega)$  is always compact, but the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is not necessarily compact.

**Trivial example:** We can take  $\Omega = \cup_{n=1}^{\infty} B_n$  be a union of disjoint balls. The functions

$$u_n = \frac{\mathbb{1}_{B_n}}{|B_n|^{1/2}}$$

are orthonormal in  $L^2(\Omega)$  and all have  $\int_{\Omega} |\nabla u_n|^2 = 0$ . Thus  $\mu_n(-\Delta_N) = 0$  for all  $n \geq 1$ . In this example, however,  $\Omega$  is not connected.

**Example “Rooms and Passages”** (Courant and Hilbert). Let  $\Omega \subset \mathbb{R}^2$  be a union of a sequence of “rooms” and “passages” where

- the  $n$ -th room is a rectangle of size  $a_n \times b_n$ .
- the  $n$ -th passage is a rectangle of size  $\varepsilon_n \times b_n$ .

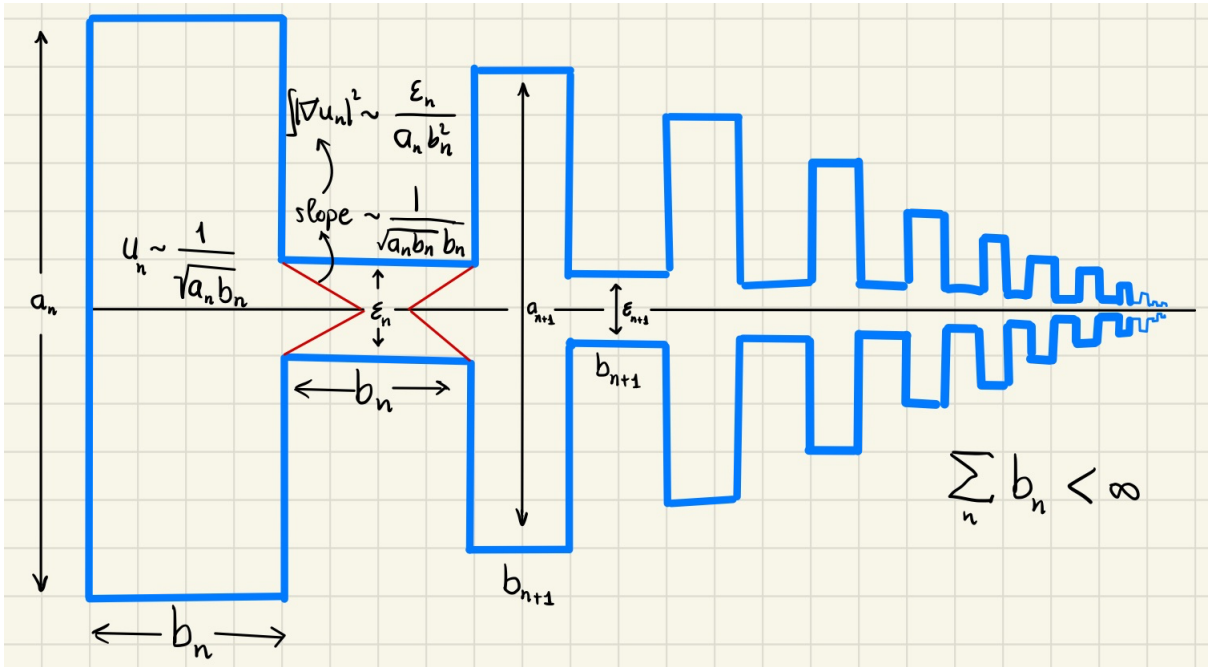


Figure: Rooms and Passages

Now the set  $\Omega$  is simply connected and it is bounded if  $\sum_n b_n < \infty$ . On the  $n$ -th room, we can take the function  $u_n \sim (a_n b_n)^{-1/2}$  and interpolate it to 0 linearly up to the middle points



of the connected passages. Thus  $\{u_n\}_{n \geq 1}$  are orthonormal in  $L^2(\Omega)$  and

$$\int_{\Omega} |\nabla u_n|^2 \leq \left( \frac{(a_n b_n)^{-1/2}}{b_n} \right)^2 \times \varepsilon_n b_n = \frac{\varepsilon_n}{a_n b_n^2}.$$

This can be made small uniformly in  $n$  by taking  $\varepsilon_n \ll a_n b_n^2$ .

By modifying this example, **Hempel-Seco-Simon** (1990) proved

**Theorem.** *For any  $d \geq 1$  and any close set  $S \subset [0, \infty)$ , there exists an open, connected, bounded subset  $\Omega \subset \mathbb{R}^d$  such that the spectrum of the Neumann Laplacian on  $L^2(\Omega)$  is equal to  $S$ .*

Thus to ensure that  $-\Delta_N$  has compact resolvent, i.e. the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is compact, we need something more than the usual requirement of the boundedness of  $\Omega$ . A sufficient condition is that  $\partial\Omega$  is sufficiently smooth (e.g.  $C^1$ ), or more generally that  $\Omega$  is an **extension domain**.

## 6.2 Extension domains

Clearly, if  $u \in H^1(\mathbb{R}^d)$ , then  $u|_{\Omega} \in H^1(\Omega)$ . More generally, if  $\Omega \subset \tilde{\Omega}$ , then we have the obvious **restriction**  $H^1(\Omega) \subset H^1(\tilde{\Omega})$ . Reversely, the **extension** is less trivial and requires some conditions on the boundary  $\partial\Omega$ .

**Definition** (Extension domains). *Let  $\Omega \subset \mathbb{R}^d$  be an open set. We call  $\Omega$  an extension domain if there exists an extension operator  $E : H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  such that*

$$(Eu)|_{\Omega} = u, \quad \|Eu\|_{L^2(\mathbb{R}^d)} \leq C\|u\|_{L^2(\Omega)}, \quad \|Eu\|_{H^1(\mathbb{R}^d)} \leq C\|u\|_{H^1(\Omega)}.$$

To illustrate the idea, we have

**Lemma** (Extension by reflection). *Consider the half-space  $\Omega = \mathbb{R}^{d-1} \times \mathbb{R}_+ = \{x = (x_1, \dots, x_d) : x_d > 0\}$ . For any function  $u : \Omega \rightarrow \mathbb{C}$  we define the extension  $Eu : \mathbb{R}^d \rightarrow \mathbb{C}$*

by

$$Eu(x) = \begin{cases} u(x), & \text{if } x_d > 0, \\ u(x_1, \dots, x_{d-1}, -x_d), & \text{if } x_d < 0. \end{cases}$$

If  $u \in H^1(\Omega)$ , then  $Eu \in H^1(\mathbb{R}^d)$ .

*Proof.* Clearly  $Eu \in L^2(\mathbb{R}^d)$  and  $\|Eu\|_{L^2(\mathbb{R}^d)}^2 = 2\|u\|_{L^2(\Omega)}^2$ .

For the derivatives, let us denote another extension  $\tilde{E}$  by

$$\tilde{E}u(x) = \begin{cases} u(x), & \text{if } x_d > 0, \\ -u(x_1, \dots, x_{d-1}, -x_d), & \text{if } x_d < 0. \end{cases}$$

Then we have

$$\partial_{x_j} Eu = \begin{cases} E\partial_{x_j} u, & \text{if } j = 1, 2, \dots, d-1, \\ \tilde{E}\partial_{x_j} u, & \text{if } j = d. \end{cases}$$

In fact, for every test function  $\varphi \in C_c^\infty(\mathbb{R}^d)$  we can write: for any  $j = 1, 2, \dots, d-1$

$$\begin{aligned} \int_{\mathbb{R}^d} (Eu)(x) \partial_{x_j} \varphi(x) dx &= \int_{\Omega} u(x) \partial_{x_j} \varphi(x) dx + \int_{\Omega^c} u(x_1, \dots, x_{d-1}, -x_d) \partial_{x_j} \varphi(x) dx \\ &= - \int_{\Omega} (\partial_{x_j} u)(x) \varphi(x) dx - \int_{\Omega^c} (\partial_{x_j} u)(x_1, \dots, x_{d-1}, -x_d) \varphi(x) dx \\ &= - \int_{\Omega} (E\partial_{x_j} u)(x) \varphi(x) dx - \int_{\Omega^c} (E\partial_{x_j} u)(x_1, \dots, x_{d-1}, x_d) \varphi(x) dx \\ &= - \int_{\mathbb{R}^d} (E\partial_{x_j} u)(x) \varphi(x) dx \end{aligned}$$

and for  $j = d$ ,

$$\begin{aligned} \int_{\mathbb{R}^d} (Eu)(x) \partial_{x_d} \varphi(x) dx &= \int_{\Omega} u(x) \partial_{x_d} \varphi(x) dx + \int_{\Omega^c} u(x_1, \dots, x_d, -x_d) \partial_{x_d} \varphi(x) dx \\ &= - \int_{\Omega} (\partial_{x_d} u)(x) \varphi(x) dx - \int_{\Omega^c} (-\partial_{x_d} u)(y, -x_d) \varphi(x) dx \\ &= - \int_{\Omega} (\tilde{E}\partial_{x_d} u)(x) \varphi(x) dx - \int_{\Omega^c} (\tilde{E}\partial_{x_d} u)(y, x_d) \varphi(x) dx \\ &= - \int_{\mathbb{R}^d} (\tilde{E}\partial_{x_d} u)(x) \varphi(x) dx. \end{aligned}$$

Hence

$$\begin{aligned} \int_{\mathbb{R}^d} |\nabla(Eu)|^2 &= \sum_{j=1}^{d-1} \int_{\mathbb{R}^d} |E\partial_{x_j}u|^2 + \int_{\mathbb{R}^d} |\tilde{E}\partial_{x_d}u|^2 \\ &= 2 \sum_{j=1}^{d-1} \int_{\Omega} |\partial_{x_j}u|^2 + 2 \int_{\Omega} |\partial_{x_d}u|^2 = 2 \int_{\mathbb{R}^d} |\nabla u|^2. \end{aligned}$$

Thus  $Eu \in H^1(\mathbb{R}^d)$  since  $\|Eu\|_{H^1(\mathbb{R}^d)}^2 = 2\|u\|_{H^1(\Omega)}^2 < \infty$ .  $\square$

Similarly, if  $\Omega \subset \mathbb{R}^{d-1} \times \mathbb{R}_+$ , we can extend functions in  $H^1(\Omega)$  to functions in  $H^1(\tilde{\Omega})$  with

$$\tilde{\Omega} = \Omega \cup \{(x_1, \dots, x_{d-1}, -x_d) : (x_1, \dots, x_d) \in \Omega\}$$

by setting for any  $u : \Omega \rightarrow \mathbb{C}$  the extension  $\tilde{u} : \tilde{\Omega} \rightarrow \mathbb{C}$  by

$$\tilde{u}(x) = \begin{cases} u(x), & \text{if } x \in \Omega_+, \\ u(x_1, \dots, x_{d-1}, -x_d), & \text{if } x \in \Omega_- < 0. \end{cases}$$

By repeating this procedure a few times. Moreover, note that  $\varphi u \in H^1(\mathbb{R}^d)$  if  $\varphi \in C_c^\infty(\mathbb{R}^d)$  and  $u \in H^1(\text{supp } \varphi)$ . Thus we can get  $H^1(\mathbb{R}^d)$ -extension for several simple domains, e.g. cubes.

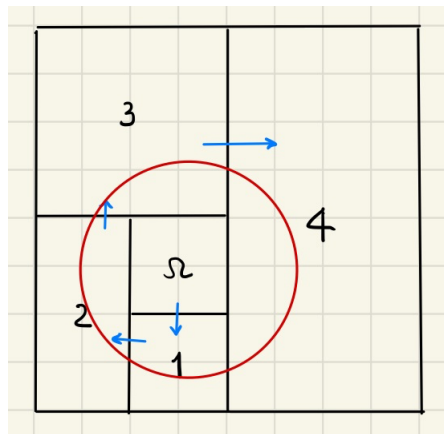


Figure: Extension by reflection

This technique can be made general by

**Definition.** Let  $Q = B_{\mathbb{R}^{d-1}}(0, 1) \times (-1, 1) = Q_+ \cup Q_- \cup Q_0$  with

$$Q_+ = B_{\mathbb{R}^{d-1}}(0, 1) \times (0, 1), Q_- = B_{\mathbb{R}^{d-1}}(0, 1) \times (-1, 0), Q_0 = B_{\mathbb{R}^{d-1}}(0, 1) \times \{0\}.$$

Let  $\Omega \subset \mathbb{R}^d$  be an open set. We say that  $\partial\Omega \in C^1$  if for every  $x \in \partial\Omega$ , there exists an open set  $x \in U \subset \mathbb{R}^d$  and a bijective map  $\Theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that

$$\Theta, \Theta^{-1} \in C^1, \quad \Theta(U) = Q, \quad \Theta(U \cap \Omega) = Q_-, \quad \Theta(U \cap \partial\Omega) = Q_0.$$

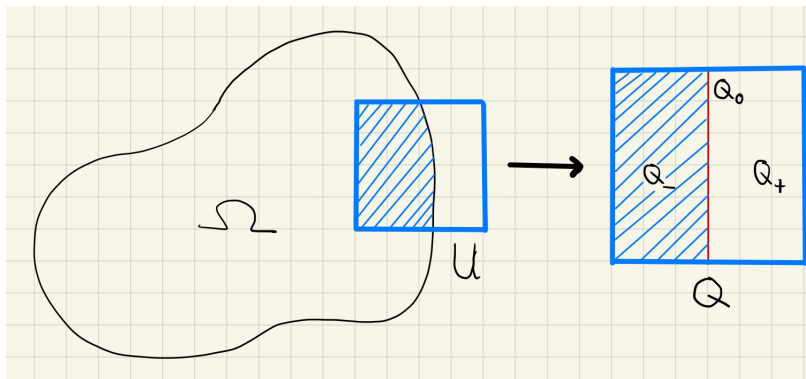


Figure:  $C^1$  boundary

**Theorem.** Let  $\Omega \subset \mathbb{R}^d$  be an open bounded set with  $\partial\Omega \in C^1$ . Then  $\Omega$  is an extension domain.

*Proof.* Since  $\partial\Omega$  is compact and  $C^1$  smooth, it can be covered by finitely many open sets  $U_1, \dots, U_k \subset \mathbb{R}^d$  such that in each set  $U_i$  we can find a bijective map  $\Theta_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that

$$\Theta_i, \Theta_i^{-1} \in C^1, \quad \Theta_i(U_i) = Q, \quad \Theta_i(U_i \cap \Omega) = Q_-, \quad \Theta_i(U_i \cap \partial\Omega) = Q_0.$$

Moreover, we can find smooths function  $\{\varphi_i\}_{i=0}^k \subset C^\infty(\mathbb{R}^d)$  such that

- $0 \leq \varphi_i \leq 1$  for all  $i$ ,  $\sum_{i=0}^k \varphi_i \equiv 1$ .
- $\text{supp } \varphi_i \subset\subset U_i$  for all  $i \neq 0$ ,  $\text{supp } \varphi_0 \subset\subset \mathbb{R}^d \setminus \partial\Omega$ .

Then for every  $u \in H^1(\Omega)$ , we can decompose

$$u = \sum_{i=0}^k \varphi_i u.$$

Let us extend each functions  $\varphi_i u$  to  $\mathbb{R}^d$ .

**Extension of  $\varphi_i u$  with  $i \neq 0$ .** Define

$$v_i = u(\Theta_i^{-1}(y)), \quad \forall y \in Q_-.$$

Then  $v_i \in H^1(Q_-)$  (we can compute the derivative by the chain rule since  $\Theta_i^{-1} \in C^1$ ). By reflection, we can extend  $v_i$  to  $\tilde{v}_i \in H^1(Q)$ . Then we can transfer back the variable using  $\Theta$  and define the extension  $E(\varphi_i u) : \mathbb{R}^d \rightarrow \mathbb{C}$

$$E(\varphi_i u) = \begin{cases} \varphi_i(x) \tilde{v}_i(\Theta_i(x)), & \text{if } x \in U_i, \\ 0, & \text{if } x \in \mathbb{R}^d \setminus U_i. \end{cases}$$

Clearly  $E(\varphi_i u) \in H^1(\mathbb{R}^d)$  since  $\tilde{v}_i(\Theta_i(x)) \in H^1(U_i)$  and  $\varphi_i \in C_c^\infty(U_i)$ .

**Extension of  $\varphi_0 u$ .** We simply define the extension  $E(\varphi_0 u) : \mathbb{R}^d \rightarrow \mathbb{C}$  by

$$E(\varphi_0 u) = \begin{cases} \varphi_0(x) u(x), & \text{if } x \in \Omega, \\ 0, & \text{if } x \in \mathbb{R}^d \setminus \Omega. \end{cases}$$

Clearly  $E(\varphi_0 u) \in H^1(\mathbb{R}^d)$  since  $u \in H^1(\Omega)$  and  $\text{supp } \varphi_0 \subset \subset \mathbb{R}^d \setminus \partial\Omega$ .

**Conclusion:** The function

$$Eu = \sum_{i=0}^k E(\varphi_i u) \in H^1(\mathbb{R}^d)$$

is an extension of  $u$ . It is easy to check that

$$\|Eu\|_{L^2(\mathbb{R}^d)} \leq C \|u\|_{L^2(\Omega)}, \quad \|Eu\|_{H^1(\mathbb{R}^d)} \leq \|u\|_{H^1(\Omega)}$$

for a constant  $C$  independent of  $u$ . □

The concept of extension domains is important because it gives a sufficient condition for the compact embedding  $H^1(\Omega) \subset L^2(\Omega)$ .

**Theorem.** *Let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. If  $\Omega$  is an extension domain, then the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is compact. Consequently, the Neumann Laplacian on  $L^2(\Omega)$  has compact resolvent.*

*Proof.* We assume that  $u_n \rightharpoonup 0$  weakly in  $H^1(\Omega)$  and prove that  $u_n \rightarrow 0$  strongly in  $L^2(\Omega)$ . The condition  $u_n \rightharpoonup 0$  weakly in  $H^1(\Omega)$  implies that

- $u_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$ , since  $H^1$ -norm is stronger than  $L^2$ -norm;
- $u_n$  is bounded in  $H^1(\Omega)$ , by Banach-Steinhaus theorem.

Let  $E : H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  be an extension operator. Then  $Eu_n$  is bounded in  $H^1(\mathbb{R}^d)$ . Thus up to a subsequence, we can assume that

$$Eu_n \rightharpoonup g \quad \text{weakly in } H^1(\mathbb{R}^d)$$

by Banach-Alaoglu theorem. By Sobolev embedding theorem, we obtain  $\mathbf{1}_\Omega Eu_n \rightarrow \mathbf{1}_\Omega g$  strongly in  $L^2(\mathbb{R}^d)$ , namely

$$u_n \rightarrow g|_\Omega \quad \text{strongly in } L^2(\Omega).$$

Since  $u_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$ , we conclude that  $g|_\Omega = 0$ , and hence  $u_n \rightarrow 0$  strongly in  $L^2(\Omega)$ . Since the limit is unique, this convergence holds for the whole sequence. Thus the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is compact.

This implies that  $(-\Delta_N + 1)^{-1}$  is a compact operator. Indeed, if  $u_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$ , then

$$v_n := (-\Delta_N + 1)^{-1}u_n \rightarrow 0$$

strongly in  $L^2(\Omega)$  as follows. Since  $u_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$  and  $(-\Delta_N + 1)^{-1}$  is bounded,  $v_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$ . Moreover,

$$\|v_n\|_{H^1(\Omega)}^2 = \langle v_n, (-\Delta_N + 1)v_n \rangle_{L^2(\Omega)} = \langle v_n, u_n \rangle_{L^2(\Omega)} \leq \|u_n\|_{L^2(\Omega)} \|v_n\|_{L^2(\Omega)}$$

is bounded. Since  $H^1(\Omega) \subset L^2(\Omega)$  is compact, up to a subsequence,  $v_n \rightarrow v$  strongly in  $L^2(\Omega)$ . Since  $v_n \rightharpoonup 0$  weakly in  $L^2(\Omega)$ , we must have  $v = 0$ . This completes the proof.  $\square$

Remark: For any open bounded set  $\Omega \subset \mathbb{R}^d$ , the embedding  $H_0^1(\Omega) \subset L^2(\Omega)$  is always compact since  $H_0^1(\Omega)$  can be always extended to  $H^1(\mathbb{R}^d)$  “by 0 from outside”.

**Exercise.** Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Define the extension  $\tilde{u} : \mathbb{R}^d \rightarrow \mathbb{C}$  by

$$\tilde{u}(x) = \begin{cases} u(x), & \text{if } x \in \Omega, \\ 0, & \text{if } x \notin \Omega. \end{cases}$$

Prove that if  $u \in H_0^1(\Omega)$ , then  $\tilde{u} \in H^1(\mathbb{R}^d)$  and

$$\nabla \tilde{u}(x) = \begin{cases} \nabla u(x), & \text{if } x \in \Omega, \\ 0, & \text{if } x \notin \Omega. \end{cases}$$

### 6.3 Kröger's inequality

Many semiclassical estimates for Dirichlet eigenvalues change the direction for Neumann eigenvalues. For example, we have the following **reverse Berezin-Li-Yau inequality**.

**Theorem** (Kröger's inequality). *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set such that the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is compact. Then the eigenvalues  $\mu_1 < \mu_2 \leq \dots$  of the Neumann Laplacian  $-\Delta_N$  on  $L^2(\Omega)$  satisfy*

$$\sum_{n=1}^N \mu_n \leq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}}.$$

*Proof.* Let  $\{u_n\}_{n \geq 1}$  be an orthonormal basis of eigenfunctions of  $-\Delta_N$ . Then

$$-\Delta_N u_n = \mu_n u_n$$

and hence

$$\langle \nabla f, \nabla u_n \rangle_{L^2(\Omega)} = \mu_n \langle f, u_n \rangle_{L^2(\Omega)}, \quad \forall f \in H^1(\Omega).$$

By the min-max principle, we know that

$$\mu_N = \inf_{\substack{u \perp u_1, \dots, u_{N-1} \\ \|u\|_{L^2(\Omega)}=1}} \int_{\Omega} |\nabla u|^2.$$

For any function  $f \in H^1(\Omega)$ , clearly

$$v := f - \sum_{n=1}^{N-1} \langle u_n, f \rangle_{L^2(\Omega)} u_n$$

is orthogonal to all  $u_1, \dots, u_{N-1}$  in  $L^2(\Omega)$ . Hence

$$\int_{\Omega} |\nabla v|^2 \geq \mu_N \int_{\Omega} |v|^2.$$

By the definition of  $v$ , we can compute

$$\int_{\Omega} |v|^2 = \int_{\Omega} |f|^2 - \sum_{n=1}^{N-1} \left| \langle u_n, f \rangle_{L^2(\Omega)} \right|^2$$

and

$$\begin{aligned} \int_{\Omega} |\nabla v|^2 &= \int_{\Omega} \left| \nabla f - \sum_{n=1}^{N-1} \langle u_n, f \rangle_{L^2(\Omega)} \nabla u_n \right|^2 \\ &= \int_{\Omega} |\nabla f|^2 - 2\Re \left( \sum_{n=1}^{N-1} \langle u_n, f \rangle_{L^2(\Omega)} \langle \nabla f, \nabla u_n \rangle_{L^2(\Omega)} \right) + \sum_{n=1}^{N-1} |\langle u_n, f \rangle_{L^2(\Omega)}|^2 \int_{\Omega} |\nabla u_n|^2 \\ &= \int_{\Omega} |\nabla f|^2 - 2\Re \left( \sum_{n=1}^{N-1} \langle u_n, f \rangle_{L^2(\Omega)} \mu_n \langle f, u_n \rangle_{L^2(\Omega)} \right) + \sum_{n=1}^{N-1} |\langle u_n, f \rangle_{L^2(\Omega)}|^2 \mu_n \\ &= \int_{\Omega} |\nabla f|^2 - \sum_{n=1}^{N-1} \mu_n |\langle u_n, f \rangle_{L^2(\Omega)}|^2. \end{aligned}$$

Here we have used the fact that  $u_n$  are eigenfunctions of  $-\Delta_N$ . Thus in summary,

$$\int_{\Omega} |\nabla f|^2 - \sum_{n=1}^{N-1} \mu_n |\langle u_n, f \rangle_{L^2(\Omega)}|^2 \geq \mu_N \left( \int_{\Omega} |f|^2 - \sum_{n=1}^{N-1} |\langle u_n, f \rangle_{L^2(\Omega)}|^2 \right), \quad \forall f \in H^1(\Omega).$$

In particular, we can choose  $f(x) = e^{2\pi i k \cdot x}$  for  $k \in \mathbb{R}^d$  and obtain

$$|2\pi k|^2 |\Omega| - \sum_{n=1}^{N-1} \mu_n |\widehat{\mathbb{1}_{\Omega} u_n}(k)|^2 \geq \mu_N \left( |\Omega| - \sum_{n=1}^{N-1} |\widehat{\mathbb{1}_{\Omega} u_n}(k)|^2 \right), \quad \forall k \in \mathbb{R}^d.$$

Integrating over  $k \in B_R(0) \subset \mathbb{R}^d$  we get

$$\int_{B_R} dk |2\pi k|^2 |\Omega| - \sum_{n=1}^{N-1} \mu_n \int_{B_R} dk |\widehat{\mathbb{1}_{\Omega} u_n}(k)|^2 \geq \mu_N \left( |B_R| |\Omega| - \int_{B_R} dk \sum_{n=1}^{N-1} |\widehat{\mathbb{1}_{\Omega} u_n}(k)|^2 \right).$$

Now we choose  $R$  such that

$$N = |\Omega| |B_R| = |\Omega| |B_1| R^d \iff R = \left( \frac{N}{|\Omega| |B_1|} \right)^{\frac{1}{d}}.$$



This gives

$$\int_{B_R} |2\pi k|^2 |\Omega| dk = |\Omega| R^{d+2} \int_{|k| \leq 1} |2\pi k|^2 dk = \underbrace{\frac{d}{d+2} \frac{4\pi^2}{|B_1|^{\frac{2}{d}}}}_{\frac{1}{|\Omega|^{\frac{2}{d}}}} N^{1+\frac{2}{d}} = \frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}}.$$

In conclusion, we have proved that

$$\frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}} - \sum_{n=1}^{N-1} \mu_n \int_{B_R} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 \geq \mu_N \left( N - \sum_{n=1}^{N-1} \int_{B_R} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 \right)$$

which is equivalent to

$$\frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}} - \sum_{n=1}^N \mu_n \geq \mu_N \sum_{n=1}^{N-1} \left( 1 - \int_{B_R} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 \right) - \sum_{n=1}^{N-1} \mu_n \left( 1 - \int_{B_R} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 \right).$$

The right side is  $\geq 0$  because for every  $n = 1, 2, \dots, N-1$  we have  $\mu_N \geq \mu_n$  and

$$\int_{B_R} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 \leq \int_{\mathbb{R}^d} dk |\widehat{\mathbb{1}_\Omega u_n}(k)|^2 = \int_{\mathbb{R}^d} dx |(\mathbb{1}_\Omega u_n)(x)|^2 = \int_{\Omega} dx |u_n(x)|^2 = 1.$$

Thus

$$\frac{K_d^{\text{cl}}}{|\Omega|^{\frac{2}{d}}} N^{1+\frac{2}{d}} - \sum_{n=1}^N \mu_n \geq 0.$$

□

## 6.4 Lieb–Thirring inequality for Neumann Laplacian

Recall the standard Lieb–Thirring kinetic inequality: For every  $d \geq 1$ , there exists a constant  $K_d > 0$  such that for all  $N \geq 1$  and for all **orthonormal functions**  $\{u_n\}_{n=1}^N$  in  $L^2(\mathbb{R}^d)$ , then

$$\sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla u_n|^2 \geq K_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}}, \quad \rho(x) = \sum_n |u_n(x)|^2.$$

The following generalization will be useful



**Exercise.** Let  $d \geq 1$ . Let  $N \geq 1$  and let  $\{u_n\}_{n=1}^N \subset H^1(\mathbb{R}^d)$  satisfy

$$\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1 \quad \text{on } L^2(\mathbb{R}^d).$$

Then

$$\sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla u_n|^2 \geq K_d \int_{\mathbb{R}^d} \rho^{1+\frac{2}{d}}, \quad \rho(x) = \sum_n |u_n(x)|^2.$$

The constant  $K_d > 0$  is the same as in the case of orthonormal functions.

This inequality can be also written in the compact form: for any trace class operator  $0 \leq \gamma \leq 1$  on  $L^2(\mathbb{R}^d)$ , then

$$\text{Tr}(-\Delta\gamma) \geq K_d \int_{\mathbb{R}^d} \rho_\gamma^{1+2/d}.$$

Here  $\text{Tr}(-\Delta\gamma) := \text{Tr}(\sqrt{-\Delta}\gamma\sqrt{-\Delta})$  and  $\rho_\gamma(x) = \sum_n |u_n(x)|^2$  if  $\gamma = \sum_n |u_n\rangle\langle u_n|$ .

Now we want to extend the above inequality for functions in  $L^2(\Omega)$  with an open bounded set  $\Omega \subset \mathbb{R}^d$ .

- This inequality extended immediately to the Dirichlet Laplacian, namely if  $\{u_n\}_n \subset H_0^1(\Omega)$  and  $\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1$  on  $L^2(\Omega)$ , then

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq K_d \int_{\Omega} \rho^{1+\frac{2}{d}}, \quad \rho(x) = \sum_{n=1}^N |u_n(x)|^2.$$

The reason is that the extension  $H_0^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  is trivial (we simply set  $u|_{\Omega^c} \equiv 0$ ).

- The analogue for Neumann Laplacian is less obvious since the extension  $H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  is more complicated.

**Theorem** (Lieb–Thirring kinetic inequality for Neumann Laplacian). Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set with  $\partial\Omega \in C^1$ . Let  $N \geq 1$  and let  $\{u_n\}_{n=1}^N \subset H^1(\Omega)$  such that  $\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1$  on  $L^2(\Omega)$ . Then

$$\sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq K_\Omega \int_{\Omega} \rho^{1+\frac{2}{d}} - \int_{\mathbb{R}^d} \rho, \quad \rho(x) = \sum_n |u_n(x)|^2$$

for a constant  $K_\Omega > 0$  independent of  $N$  and  $\{u_n\}_{n=1}^N$ .

Our idea is to use the extension operator  $E : H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  (which requires  $\partial\Omega \in C^1$ ) and then apply the standard Lieb–Thirring inequality for  $L^2(\mathbb{R}^d)$ . The key observation is that the extension operator does not destroy the orthogonality too much.

**Lemma.** *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set with  $\partial\Omega \in C^1$ . Let  $E : H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  be the extension operator constructed in a previous section. Then for every  $N \geq 1$ , if*

$$\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1 \quad \text{on } L^2(\Omega),$$

then

$$\sum_{n=1}^N |Eu_n\rangle\langle Eu_n| \leq C_\Omega \quad \text{on } L^2(\mathbb{R}^d)$$

Here the constant  $C_\Omega$  is independent of  $N$  and  $\{u_n\}_{n=1}^N$ .

*Proof.* First let us quickly recall the definition of  $E$ . We cover  $\partial\Omega$  by  $J$  open sets  $\{U_j\}_{j=1}^J$  such that in each set  $U_i$  we can find a bijective map  $\Theta_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that

$$\Theta_i, \Theta_i^{-1} \in C^1, \quad \Theta_i(U_i) = Q, \quad \Theta_i(U_i \cap \Omega) = Q_-, \quad \Theta_i(U_i \cap \partial\Omega) = Q_0.$$

Then we use a partition of unity  $1 = \sum_{j=0}^J \varphi_j$  on  $\mathbb{R}^d$  with  $\{\varphi_j\}_{j=0}^J \subset C^\infty(\mathbb{R}^d)$  such that

$$0 \leq \varphi_j \leq 1, \quad \text{supp } \varphi_j \subset\subset U_j \quad \forall j \neq 0, \quad \text{supp } \varphi_0 \subset\subset \mathbb{R}^d \setminus \partial\Omega.$$

We decompose

$$u = \sum_{j=0}^J \varphi_j u$$

and extend  $\varphi_j u$  as follows:

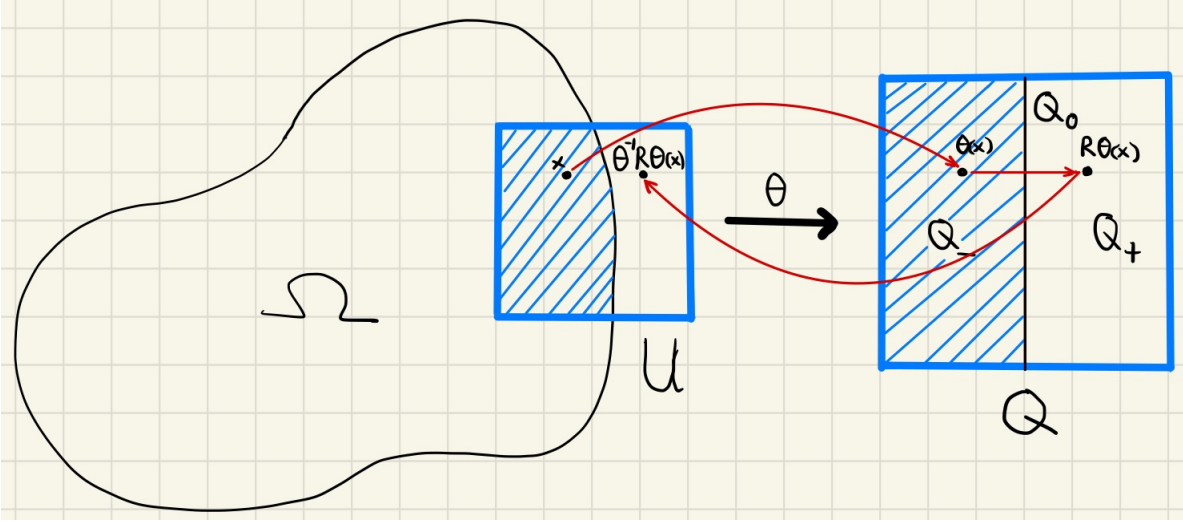


Figure: The mapping  $h_i = \Theta_i^{-1} R \Theta_i : U_i \cap \Omega \rightarrow U_i \setminus \bar{\Omega}$

- For  $j = 0$ ,  $\varphi_0 u \in H_0^1(\Omega) \subset H^1(\mathbb{R}^d)$  (by setting  $(\varphi_0 u)(y) = 0$  with  $y \notin \Omega$ );
- For  $j \neq 0$ , we set  $T_j u : U_j \rightarrow \mathbb{C}$  by  $T_j u = u$  on  $U_j \cap \Omega$  and  $T_j u(y) = u(h_j^{-1}(y))$  with  $y \in U_j \setminus \bar{\Omega}$  where

$$h_j = \Theta_j^{-1} R \Theta_j : U_j \cap \Omega \rightarrow U_j \setminus \bar{\Omega}$$

(here  $R : Q_- \rightarrow Q_+$  is the usual reflection). Then  $\varphi_j T_j u \in H_0^1(U_j) \subset H^1(\mathbb{R}^d)$ .

Thus in summary

$$Eu = \varphi_0 u + \sum_{j=1}^J \varphi_j T_j u, \quad \forall u \in L^2(\Omega).$$

Now let us show that if  $\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1$  on  $L^2(\Omega)$ , then

$$\sum_{n=1}^N |Eu_n\rangle\langle Eu_n| \leq C \quad \text{on } L^2(\mathbb{R}^d)$$

namely

$$\sum_{n=1}^N \left| \int_{\mathbb{R}^d} (Eu_n)(x) g(x) dx \right|^2 \leq C \int_{\mathbb{R}^d} |g|^2, \quad \forall g \in L^2(\mathbb{R}^d).$$

By the Cauchy-Schwarz inequality for complex numbers

$$\begin{aligned} \sum_{n=1}^N \left| \int_{\mathbb{R}^d} (Eu_n) g \right|^2 &= \sum_{n=1}^N \left| \int_{\mathbb{R}^d} \left( \varphi_0 u + \sum_{j=1}^J \varphi_j T_j u_n \right) g \right|^2 \\ &\leq \sum_{n=1}^N (J+1) \left[ \left| \int_{\mathbb{R}^d} (\varphi_0 u_n) g \right|^2 + \sum_{j=1}^J \left| \int_{\mathbb{R}^d} (\varphi_j T_j u_n) g \right|^2 \right]. \end{aligned}$$

For  $j = 0$ , we use  $\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1$  on  $L^2(\Omega)$  to bound

$$\sum_{n=1}^N \left| \int_{\mathbb{R}^d} (\varphi_0 u_n) g \right|^2 = \sum_{n=1}^N \left| \int_{\Omega} u_n (\varphi_0 g) \right|^2 \leq \int_{\Omega} |\varphi_0 g|^2 \leq \int_{\mathbb{R}^d} |g|^2.$$

For  $j \neq 0$ , we decompose further

$$\begin{aligned} \int_{\mathbb{R}^d} (\varphi_j T_j u_n) g &= \int_{U_j \cap \Omega} (\varphi_j T_j u_n) g + \int_{U_j \setminus \Omega} (\varphi_j T_j u_n) g \\ &= \int_{U_j \cap \Omega} \varphi_j(x) u_n(x) g(x) dx + \int_{U_j \setminus \Omega} \varphi_j(y) u_n(h_j^{-1}(y)) g(y) dy \\ &= \int_{U_j \cap \Omega} \varphi_j(x) u_n(x) g(x) dx + \int_{U_j \cap \Omega} \varphi_j(h_j(x)) u_n(x) g(h_j(x)) |\text{Det} \mathbf{J} h_j(x)| dx \\ &= \int_{U_j \cap \Omega} u_n(x) \left[ \varphi_j(x) g(x) + \varphi_j(h_j(x)) g(h_j(x)) |\text{Det} \mathbf{J} h_j(x)| \right] dx \\ &= \int_{\Omega} u_n(x) \mathbf{1}_{U_j}(x) \left[ \varphi_j(x) g(x) + \varphi_j(h_j(x)) g(h_j(x)) |\text{Det} \mathbf{J} h_j(x)| \right] dx. \end{aligned}$$

Here we have changed the variable  $y = h_j(x)$  and  $\text{Det} \mathbf{J} h_j$  is the Jacobian determinant of  $h_j$ .

Thus using again  $\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1$  on  $L^2(\Omega)$  we have

$$\begin{aligned} \sum_{n=1}^N \left| \int_{\mathbb{R}^d} (\varphi_j T_j u_n) g \right|^2 &= \sum_{n=1}^N \left| \int_{\Omega} u_n(x) \mathbf{1}_{U_j}(x) \left[ \varphi_j(x) g(x) + \varphi_j(h_j(x)) g(h_j(x)) |\text{Det} \mathbf{J} h_j(x)| \right] dx \right|^2 \\ &\leq \int_{U_j \cap \Omega} \left| \varphi_j(x) g(x) + \varphi_j(h_j(x)) g(h_j(x)) (\text{Det} \mathbf{J} h_j)(x) \right|^2 dx \\ &\leq 2 \int_{U_j \cap \Omega} |\varphi_j(x)|^2 |g(x)|^2 + 2 \int_{U_j \cap \Omega} |\varphi_j(h_j(x))|^2 |g(h_j(x))|^2 |\text{Det} \mathbf{J} h_j(x)|^2 dx \\ &\leq 2 \int_{U_j \cap \Omega} |\varphi_j(x)|^2 |g(x)|^2 + 2 \|\text{Det} \mathbf{J} h_j\|_{L^\infty} \int_{U_j \cap \Omega} |\varphi_j(h_j(x))|^2 |g(h_j(x))|^2 |\text{Det} \mathbf{J} h_j(x)| dx \\ &= \leq 2 \int_{U_j \cap \Omega} |\varphi_j(x)|^2 |g(x)|^2 + 2 \|\text{Det} \mathbf{J} h_j\|_{L^\infty} \int_{U_j \setminus \Omega} |\varphi_j(y)|^2 |g(y)|^2 dy \leq C \int_{\mathbb{R}^d} |g|^2. \end{aligned}$$

Thus in summary,

$$\sum_{n=1}^N \left| \int_{\mathbb{R}^d} (E u_n)(x) g(x) dx \right|^2 \leq C \int_{\mathbb{R}^d} |g|^2, \quad \forall g \in L^2(\mathbb{R}^d).$$

This completes the proof.  $\square$

Now we are ready to give

**Proof of the Lieb–Thirring kinetic inequality for Neumann Laplacian.** Let  $\{u_n\}_{n=1}^N \subset H^1(\Omega)$  satisfy

$$\sum_{n=1}^N |u_n\rangle\langle u_n| \leq 1 \quad \text{on } L^2(\Omega).$$

Let  $E : H^1(\Omega) \rightarrow H^1(\mathbb{R}^d)$  be the extension operator we discussed above. Then

$$\|Eu_n\|_{L^2(\mathbb{R}^d)} \leq C\|u_n\|_{L^2(\Omega)}, \quad \|Eu_n\|_{H^1(\mathbb{R}^d)} \leq C\|u_n\|_{H^1(\Omega)}, \quad \forall n = 1, 2, \dots, N$$

and

$$\sum_{n=1}^N |Eu_n\rangle\langle Eu_n| \leq C_\Omega \quad \text{on } L^2(\mathbb{R}^d).$$

Therefore,

$$\sum_{n=1}^N \|u_n\|_{H^1(\Omega)}^2 \geq \frac{1}{C} \sum_{n=1}^N \|Eu_n\|_{H^1(\mathbb{R}^d)}^2 \geq \frac{1}{C} \sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla(Eu_n)|^2.$$

On the other hand, applying the Lieb–Thirring inequality for functions  $Eu_n/\sqrt{C_\Omega}$  in  $L^2(\mathbb{R}^d)$  we have

$$\sum_{n=1}^N \int_{\mathbb{R}^d} |\nabla(Eu_n)|^2 \geq \frac{K_d}{C_\Omega^{2/d}} \int_{\mathbb{R}^d} \left( \sum_{n=1}^N |Eu_n|^2 \right)^{1+2/d} \geq \frac{K_d}{C_\Omega^{2/d}} \int_{\Omega} \left( \sum_{n=1}^N |u_n|^2 \right)^{1+2/d} = \frac{K_d}{C_\Omega^{2/d}} \int_{\Omega} \rho^{1+2/d}.$$

Thus we conclude that

$$\int_{\Omega} \rho + \sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 = \sum_{n=1}^N \|u_n\|_{H^1(\Omega)}^2 \geq K_\Omega \int_{\Omega} \rho^{1+2/d}.$$

□

## 6.5 Weyl's law

**Theorem** (Weyl's law for distribution of Neumann eigenvalues). *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set with  $\partial\Omega \in C^1$ . Then the eigenvalues  $\mu_1 < \mu_2 \leq \dots$  of the Neumann Laplacian  $-\Delta_N$  on  $L^2(\Omega)$  satisfy*

$$\mu_N = \frac{(2\pi)^2}{|B_1|^{2/d} |\Omega|^{2/d}} N^{2/d} + o(N^{2/d})_{N \rightarrow \infty}.$$

Equivalently, if we denote by  $N(\lambda)$  the number of eigenvalues  $< \lambda$ , then

$$N(\lambda) = \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}} + o(\lambda^{\frac{d}{2}})_{\lambda \rightarrow \infty}.$$

*Proof.* Thanks the Tauberian lemma, it suffices to prove Weyl's law for sum of Neumann eigenvalues

$$\sum_{n=1}^N \mu_n = \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+\frac{2}{d}} + o(N^{1+\frac{2}{d}})_{N \rightarrow \infty}.$$

Thanks to Kröger's inequality, it remains to prove the lower bound of the sum. The idea is to compare with the Dirichlet Laplacian on a smaller set.

**Step 1.** For every  $\varepsilon > 0$  small, by the inner regularity of Lebesgue measure, we can find an open set  $\Omega_\varepsilon \subset\subset \Omega$  such that

$$|\Omega \setminus \Omega_\varepsilon| \leq \varepsilon.$$

Since  $\text{dist}(\overline{\Omega_\varepsilon}, \Omega^c) > 0$ , we can find a smooth function  $\varphi_\varepsilon \in C_c^\infty$  such that

$$0 \leq \varphi_\varepsilon \leq 1, \quad \varphi_\varepsilon \equiv 1 \text{ on } \Omega_\varepsilon, \quad \text{supp}(\varphi_\varepsilon) \subset\subset \Omega.$$

Now let  $\{u_n\}_{n \geq 1}$  be an orthonormal basis of eigenfunctions of  $-\Delta_N$  on  $L^2(\Omega)$ . By the Cauchy–Schwarz inequality we can bound

$$\begin{aligned} \int_{\Omega} |\nabla(\varphi_\varepsilon u_n)|^2 &= \int_{\Omega} |(\nabla \varphi_\varepsilon)u_n + \varphi_\varepsilon(\nabla u_n)|^2 \\ &= \int_{\Omega} |\nabla \varphi_\varepsilon|^2 |u_n|^2 + \int_{\Omega} |\varphi_\varepsilon|^2 |\nabla u_n|^2 + 2\Re \int_{\Omega} \overline{(\nabla \varphi_\varepsilon)u_n} \varphi_\varepsilon(\nabla u_n) \\ &\leq (1 + \delta^{-1}) \int_{\Omega} |\nabla \varphi_\varepsilon|^2 |u_n|^2 + (1 + \delta) \int_{\Omega} |\varphi_\varepsilon|^2 |\nabla u_n|^2 \\ &\leq C_{\varepsilon, \delta} + (1 + \delta)\mu_n. \end{aligned}$$

for all  $n \geq 1$  and  $\delta > 0$ . Thus

$$(1 + \delta) \sum_{n=1}^N \mu_n = (1 + \delta) \sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq \sum_{n=1}^N \int_{\Omega} |\nabla(\varphi_\varepsilon u_n)|^2 - C_{\varepsilon, \delta} N.$$

**Step 2.** Now we estimate  $\sum_{n=1}^N \int_{\Omega} |\nabla(\varphi_\varepsilon u_n)|^2$  from below. Since  $\text{supp}(\varphi_\varepsilon) \subset\subset \Omega$ , we know

that  $\varphi_\varepsilon u_n \in H_0^1(\Omega) \subset H^1(\mathbb{R}^d)$ . Define

$$\gamma := \sum_{n=1}^N |u_n\rangle\langle u_n|, \quad \rho = \rho_\gamma = \sum_{n=1}^N |u_n|^2.$$

Since  $\{u_n\}$  are orthonormal in  $L^2(\Omega)$ , we have  $0 \leq \gamma \leq 1$  on  $L^2(\Omega)$ . Consequently,

$$0 \leq \varphi_\varepsilon \gamma \varphi_\varepsilon = \sum_{n=1}^N |\varphi_\varepsilon u_n\rangle\langle \varphi_\varepsilon u_n| \leq 1 \quad \text{on } L^2(\Omega).$$

Hence, we can use the Berezin-Li-Yau inequality to bound  $\text{Tr}(-\Delta_D \varphi_\varepsilon \gamma \varphi_\varepsilon)$  from below.

**Exercise.** Let  $\Omega \subset \mathbb{R}^d$  be an open bounded set. Let  $\{v_n\} \subset H_0^1(\Omega)$  such that

$$0 \leq \tilde{\gamma} := \sum_{n=1}^N |v_n\rangle\langle v_n| \leq 1 \quad \text{on } L^2(\Omega).$$

Prove that

$$\text{Tr}(-\Delta_D \tilde{\gamma}) \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( \text{Tr} \tilde{\gamma} \right)^{1+\frac{2}{d}},$$

namely

$$\sum_{n=1}^N \int_\Omega |\nabla v_n|^2 \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( \sum_{n=1}^N \int_\Omega |v_n|^2 \right)^{1+\frac{2}{d}}.$$

Thus in summary, we obtain

$$\sum_{n=1}^N \int_\Omega |\nabla(\varphi_\varepsilon u_n)|^2 \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( \sum_{n=1}^N \int_\Omega \varphi_\varepsilon^2 |u_n|^2 \right)^{1+\frac{2}{d}} = \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( N - \int_\Omega (1 - \varphi_\varepsilon^2) \rho \right)^{1+\frac{2}{d}}.$$

Combining with the bound from Step 1, we find that

$$(1 + \delta) \sum_{n=1}^N \mu_n \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( N - \int_\Omega (1 - \varphi_\varepsilon^2) \rho \right)^{1+\frac{2}{d}} - C_{\varepsilon, \delta} N.$$

**Step 3.** Now we bound

$$\int_\Omega (1 - \varphi_\varepsilon^2) \rho \leq \int_{\Omega \setminus \Omega_\varepsilon} \rho.$$

Using Kröger's inequality and the Lieb-Thirring inequality for orthonormal functions  $\{u_n\} \subset$



$H^1(\Omega)$ , we have

$$\frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} N^{1+2/d} \geq \sum_{n=1}^N \mu_n = \sum_{n=1}^N \int_{\Omega} |\nabla u_n|^2 \geq K_{\Omega} \int_{\Omega} \rho^{1+2/d} - N.$$

Hence,

$$\|\rho\|_{L^{1+2/d}(\Omega)} \leq C_{\Omega} N.$$

By Hölder's inequality and the choice  $|\Omega \setminus \Omega_{\varepsilon}| \leq \varepsilon$ ,

$$\int_{\Omega \setminus \Omega_{\varepsilon}} \rho \leq \|1\|_{L^{1+d/2}(\Omega \setminus \Omega_{\varepsilon})} \|\rho\|_{L^{1+2/d}(\Omega \setminus \Omega_{\varepsilon})} \leq C_{\Omega} \varepsilon^{2/(d+2)} N.$$

**Step 4.** In conclusion, for every  $\varepsilon > 0$  and  $\delta > 0$  we have

$$(1 + \delta) \sum_{n=1}^N \mu_n \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( N - C_{\Omega} \varepsilon^{2/(d+2)} N \right)^{1+\frac{2}{d}} - C_{\varepsilon, \delta} N.$$

Consequently,

$$(1 + \delta) \liminf_{N \rightarrow \infty} N^{-1-2/d} \sum_{n=1}^N \mu_n \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}} \left( 1 - C_{\Omega} \varepsilon^{2/(d+2)} \right)^{1+\frac{2}{d}}.$$

Taking  $\varepsilon \rightarrow 0$ , and then  $\delta \rightarrow 0$ , we conclude that

$$\liminf_{N \rightarrow \infty} N^{-1-2/d} \sum_{n=1}^N \mu_n \geq \frac{K_d^{\text{cl}}}{|\Omega|^{2/d}}.$$

This completes the desired lower bound for the sum of Neumann eigenvalues, and completes the proof of Weyl's law for the distribution of eigenvalues.  $\square$

## 6.6 Pólya conjecture

As we have seen, many inequalities change their directions when we turn Dirichlet to Neumann eigenvalues. **Pólya's conjecture** states that

$$\mu_n(-\Delta_D) \geq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}} |\Omega|^{2/d}} n^{\frac{2}{d}} \geq \mu_{n+1}(-\Delta_N), \quad \forall n \geq 1.$$

**Proof of Pólya's conjecture for Neumann eigenvalues of cubes.** Let  $\Omega \subset \mathbb{R}^d$  be a cube. By scaling, it suffices to consider the case  $\Omega = [0, \pi]^d$  where the Neumann eigenvalues are given explicitly by

$$\{|x|^2 = (x_1^2 + \dots + x_d^2) \mid x = (x_1, \dots, x_d) \in \mathbb{N}_0^d\}, \quad \mathbb{N}_0 = \{0, 1, 2, \dots\}.$$

If we denote for any point  $x = (x_1, \dots, x_d)$  the cube

$$Q_x = (x_1, x_1 + 1) \times (x_2, x_2 + 1) \times \dots \times (x_d, x_d + 1),$$

then

$$|\mathbb{N}_0^d \cap B_R(0)| = \left| \bigcup_{x \in \mathbb{N}_0^d \cap B_R(0)} Q_x \right| \geq 2^{-d} |B_R| = 2^{-d} |B_1| R^d.$$

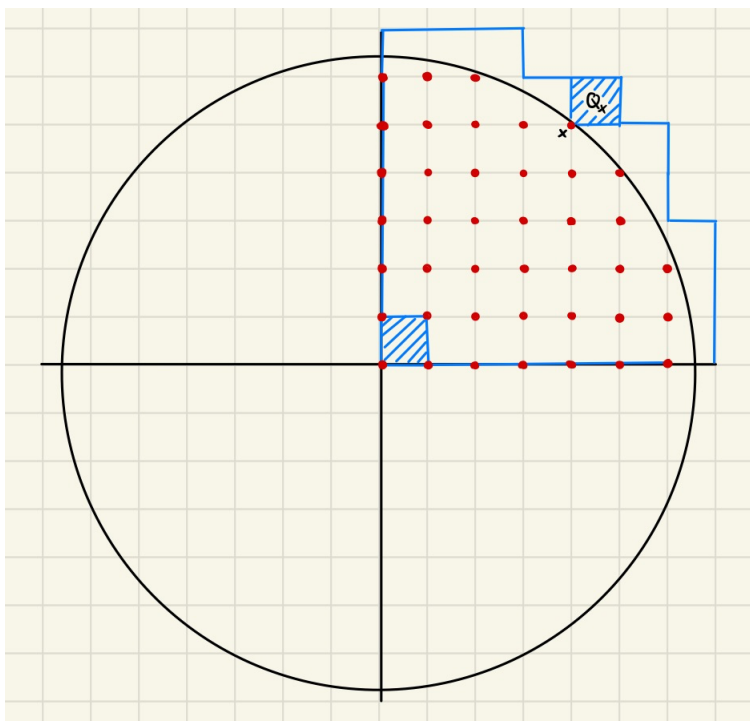


Figure: Nonnegative integer points inside a circle

Thus for every  $\lambda > 0$ , the number of Neumann eigenvalues  $< \lambda$ , which is equal to the number of non-negative integer points inside  $B(0, \sqrt{\lambda})$ , satisfies

$$N(\lambda) = |\mathbb{N}_0^d \cap B_{\sqrt{\lambda}}(0)| \geq 2^{-d} |B_1| \lambda^{d/2} = \frac{|B_1| |\Omega|}{(2\pi)^d} \lambda^{d/2}, \quad \forall \lambda > 0.$$

Since there are at most  $n$  eigenvalues  $< \mu_{n+1}$ , we have

$$n \geq N(\mu_{n+1}) \geq \frac{|B_1||\Omega|}{(2\pi)^d} \mu_{n+1}^{\frac{d}{2}}$$

which is equivalent to

$$\mu_{n+1} \leq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} n^{\frac{2}{d}}, \quad \forall n \geq 1.$$

**Pólya** (1961) extended this result for a sub-class of **tiling domains**. A proof for all tiling domains was just found recently by **Filonov** (June 2020).

**Theorem.** *Let  $d \geq 1$  and let  $\Omega \subset \mathbb{R}^d$  be an open bounded set such that the embedding  $H^1(\Omega) \subset L^2(\Omega)$  is compact. Assume that  $\Omega$  is a **tiling domain**, namely we can cover  $\mathbb{R}^d$  (up to a set of 0 measure) by a union of disjoint copies of  $\Omega$  (each copy is obtained from  $\Omega$  up to translation, rotation and reflection). Then the eigenvalues  $\mu_1 \leq \mu_2 \leq \dots$  of the Neumann Laplacian  $-\Delta_N$  on  $L^2(\Omega)$  satisfy*

$$\mu_{n+1} \leq \frac{(2\pi)^2}{|B_1|^{\frac{2}{d}}|\Omega|^{2/d}} n^{\frac{2}{d}}, \quad \forall n \geq 1.$$

*Equivalently, the number of Neumann eigenvalues  $< \lambda$  satisfies*

$$N(\lambda) \geq \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}}, \quad \forall \lambda > 0.$$

We will need the sub-additivity of  $\Omega \mapsto N(\lambda, \Omega)$ .

**Lemma.** *Let  $\{\Omega_j\}_{j=1}^J$  be disjoint open sets in  $\mathbb{R}^d$ . Then*

$$N(\lambda, \Omega) \leq \sum_{j=1}^J N(\lambda, \Omega_j), \quad \Omega = \text{interior of } \left( \bigcup_{j=1}^J \Omega_j \right).$$

*Proof.* First, we assume  $N(\lambda, \Omega) = k$ , namely

$$\mu_k < \lambda \leq \mu_{k+1}.$$

By the min-max principle,

$$\mu_k = \inf_{\substack{M_k \subset H^1(\Omega) \\ \dim M_k = k}} \sup_{u \in M_k} \frac{\int_{\Omega} |\nabla u|^2}{\int_{\Omega} |u|^2}.$$

Note that

$$H^1(\Omega) \subset X = \bigoplus_{j=1}^J H^1(\Omega_j) = \{u \in L^2(\Omega) \mid u = \sum_{j=1}^J \mathbb{1}_{\Omega_j} u_j \text{ such that } u_j \in H^1(\Omega_j) \text{ for all } j\}.$$

Hence,

$$\mu_k \geq \tilde{\mu}_k = \inf_{\substack{M_k \subset X \\ \dim M_k = k}} \sup_{u \in M_k} \frac{\sum_{j=1}^J \int_{\Omega_j} |\nabla u|^2}{\sum_{j=1}^J \int_{\Omega_j} |u|^2}.$$

On the other hand, for every subspace  $M_k \subset X$  with  $\dim M_k = k$ , we can decompose

$$M_k = \bigoplus_{j=1}^J M_{k,j}, \quad M_{k,j} \subset H^1(\Omega_j), \quad \dim M_{k,j} = \ell(k,j), \quad \sum_{j=1}^J \ell(k,j) = k.$$

Hence, by the min-max principle for the Neumann Laplacian on each  $\Omega_j$ ,

$$\sup_{u \in M_k} \frac{\sum_{j=1}^J \int_{\Omega_j} |\nabla u|^2}{\sum_{j=1}^J \int_{\Omega_j} |u|^2} \geq \sup_{1 \leq j \leq J} \left( \sup_{u \in M_{k,j}} \frac{\int_{\Omega_j} |\nabla u|^2}{\int_{\Omega_j} |u|^2} \right) \geq \sup_{1 \leq j \leq J} \mu_{\ell(k,j)}(\Omega_j).$$

Thus

$$\lambda > \mu_k \geq \tilde{\mu}_k \geq \inf_{\sum_{j=1}^J \ell(k,j) = k} \sup_{1 \leq j \leq J} \mu_{\ell(k,j)}(\Omega_j).$$

The infimum is taken over a finite set, so it must be attained for some  $\{\ell(k,j)\}_{j=1}^J$ . Thus for all  $j$

$$\lambda > \mu_{\ell(k,j)}(\Omega_j) \iff \ell(k,j) \leq N(\lambda, \Omega_j).$$

Therefore

$$N(\lambda, \Omega) = k = \sum_{j=1}^J \ell(k,j) \leq \sum_{j=1}^J N(\lambda, \Omega_j).$$

□

The above lemma allows to prove Pólya's conjecture for unions of cubes. Next, we have the “almost monotonicity” of  $\Omega \mapsto N(\lambda, \Omega)$ .

**Lemma.** Consider open bounded sets  $\Omega \subset \tilde{\Omega} \subset \mathbb{R}^d$ . Assume that there exists an extension operator  $E : H^1(\Omega) \rightarrow H^1(\tilde{\Omega})$ . Then

$$N(\lambda, \Omega) \leq N(\|E\|^2(\lambda + 1), \tilde{\Omega}).$$

*Proof.* We use the min-max principle again

$$\mu_k(\Omega) = \inf_{\substack{M_k \subset H^1(\Omega) \\ \dim M_k = k}} \sup_{u \in M_k} \frac{\int_{\Omega} |\nabla u|^2}{\int_{\Omega} |u|^2}.$$

Actually we know that the infimum is attained at

$$M_k = \text{Span}\{u_1, \dots, u_k\}$$

where  $\{u_n\}_n \subset H^1(\Omega)$  is an orthonormal basis of eigenfunctions of  $-\Delta_N(\Omega)$ .

Then we define

$$\widetilde{M}_k = EM_k \subset H^1(\widetilde{\Omega}), \quad \dim \widetilde{M}_k = k.$$

By the min-max principle on  $\widetilde{\Omega}$

$$\mu_k(\widetilde{\Omega}) \leq \sup_{v \in \widetilde{M}_k} \frac{\int_{\widetilde{\Omega}} |\nabla v|^2}{\int_{\widetilde{\Omega}} |v|^2} = \sup_{u \in M_k} \frac{\int_{\widetilde{\Omega}} |\nabla(Eu)|^2}{\int_{\widetilde{\Omega}} |Eu|^2}.$$

For every  $u \in M_k$ , we can bound

$$\begin{aligned} & \int_{\widetilde{\Omega}} |\nabla(Eu)|^2 + \int_{\widetilde{\Omega}} |Eu|^2 = \|Eu\|_{H^1(\widetilde{\Omega})}^2 \\ & \leq \|E\|^2 \|u\|_{H^1(\Omega)}^2 = \|E\|^2 \left( \int_{\Omega} |\nabla u|^2 + \int_{\Omega} |u|^2 \right) \\ & \leq \|E\|^2 (\mu_k(\Omega) + 1) \int_{\Omega} |u|^2 \leq \|E\|^2 (\mu_k(\Omega) + 1) \int_{\widetilde{\Omega}} |u|^2. \end{aligned}$$

Hence,

$$\mu_k(\widetilde{\Omega}) \leq \sup_{u \in M_k} \frac{\int_{\widetilde{\Omega}} |\nabla(Eu)|^2}{\int_{\widetilde{\Omega}} |Eu|^2} \leq \|E\|^2 (\mu_k(\Omega) + 1).$$

In particular, if  $N(\lambda, \Omega) = k$ , then  $\mu_k(\Omega) < \lambda$ , and hence

$$\mu_k(\widetilde{\Omega}) \leq \|E\|^2 (\mu_k(\Omega) + 1) < \|E\|^2 (\lambda + 1).$$

Thus  $N(\lambda, \Omega) = k \leq N(\|E\|^2(\lambda + 1), \widetilde{\Omega})$ . □

In order to put the previous lemma in a good use, we need to control the norm of the extension operator in some simple cases.

**Lemma.** Consider three cubes  $Q^- \subset Q \subset Q^+$  where  $Q^+ = [-L^+, L^+]^d$ ,  $Q = [-L, L]^d$ ,  $Q^- = [-L^-, L^-]$  with  $L^+ - L = L - L^-$ . Take a closed set  $U \subset Q^-$ . Then there exists an extension operator

$$E : H^1(Q \setminus U) \rightarrow H^1(Q^+ \setminus U)$$

with  $\|E\|^2 \leq 2^d$ .

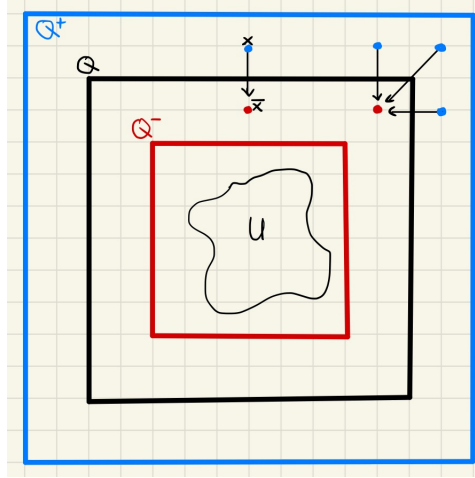


Figure: Reflection  $x \mapsto \bar{x}$ .

*Proof.* For every  $x \in Q^+ \setminus Q$ , we define the “reflection point”  $\bar{x} \in Q \setminus Q^-$  by

$$(\bar{x})_j = \begin{cases} x_j, & \text{if } x_j \in [-L, L], \\ 2L - x_j, & \text{if } x_j \in [L, L^+], \\ -2L - x_j, & \text{if } x_j \in [-L^+, -L]. \end{cases}, \quad \forall j = 1, 2, \dots, d.$$

Note that each  $\bar{x}$  has at most  $2^d - 1$  preimages. Then for every  $f \in H^1(Q \setminus U)$  we define the extension  $Ef \in H^1(Q^+ \setminus U)$  by

$$(Ef)(x) = \begin{cases} f(x), & \text{if } x \in Q \setminus U, \\ f(\bar{x}), & \text{if } x \in Q^+ \setminus Q. \end{cases}$$

Then it is straightforward to check that  $Ef \in H^1(Q^+ \setminus U)$  and

$$\|Ef\|_{H^1(Q^+ \setminus U)}^2 \leq 2^d \|f\|_{H^1(Q \setminus U)}^2.$$

□

Finally we can give

**Filonov's proof of Pólya's conjecture for tiling domains.** We know that  $R^d$  can be covered by disjoint sets  $\{\Omega_n\}_{n=1}^\infty$  where each  $\Omega_n$  is a copy of  $\Omega$ . Let  $R = \text{diam}(\Omega)$ . Take a big number  $L > 2R$  and denote

$$Q = [-L, L]^d, \quad Q^- = [-(L-R), L-R]^d, \quad Q^+ = [-(L+R), L+R]^d.$$

Let  $\{\Omega_j\}_{j \in J}$  be all copies inside  $Q^-$ . Denote

$$U = \overline{\bigcup_{j \in J} \Omega_j} \subset Q^-.$$

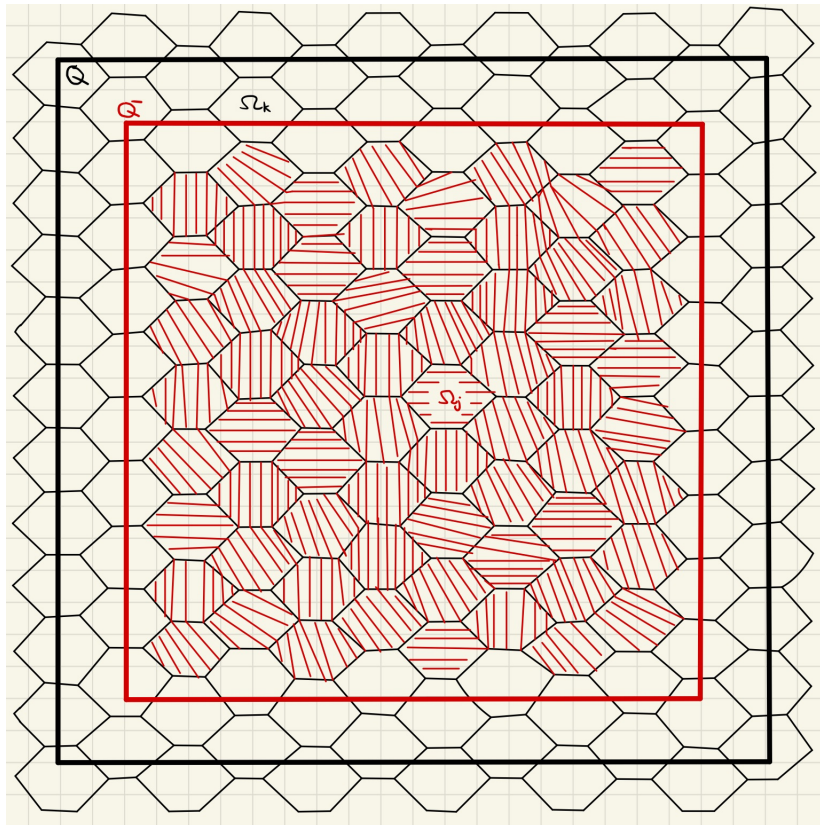


Figure:  $\{\Omega_j\}_{j \in J} \subset Q^- \subset Q$

By the sub-additivity of  $N(\lambda, \cdot)$  we have

$$N(\lambda, Q) \leq \sum_{j \in J} N(\lambda, \Omega_j) + N(\lambda, Q \setminus U).$$

On the other hand, by the extension lemma, we can find an extension operator

$$E : H^1(Q \setminus U) \rightarrow H^1(Q^+ \setminus U), \quad \|E\|^2 \leq 2^d.$$

Let  $\{\Omega_k\}_{k \in K}$  be all copies inside  $Q^+ \setminus U$ , namely

$$Q \setminus U \subset V = \overline{\bigcup_{k \in K} \Omega_k} \subset Q^+ \setminus U.$$

Then  $E$  also defines an extension operator

$$E_1 : H^1(Q \setminus U) \rightarrow H^1(\text{int}(V)) \subset H^1(Q^+ \setminus U), \quad \|E_1\|^2 \leq \|E\|^2 \leq 2^d.$$

Thus by the ‘‘almost monotonicity’’ and the sub-additivity of  $N(\lambda, \cdot)$ , we have

$$N(\lambda, Q \setminus U) \leq N(2^d(\lambda + 1), V) \leq \sum_{k \in K} N(2^d(\lambda + 1), \Omega_k).$$

In conclusion, we already proved

$$\begin{aligned} N(\lambda, Q) &\leq \sum_{j \in J} N(\lambda, \Omega_j) + N(\lambda, Q \setminus U) \\ &\leq \sum_{j \in J} N(\lambda, \Omega_j) + \sum_{k \in K} N(2^d(\lambda + 1), \Omega_k) = |J|N(\lambda, \Omega) + |K|N(2^d(\lambda + 1), \Omega). \end{aligned}$$

Since Pólya’s conjecture holds for cubes, we have

$$N(\lambda, Q) \geq \frac{|B_1||Q|}{(2\pi)^d} \lambda^{\frac{d}{2}} = \frac{|B_1|(2L)^d}{(2\pi)^d} \lambda^{\frac{d}{2}}$$

while

$$|J| \leq \frac{|Q^-|}{|\Omega|} \leq \frac{(2L - 2R)^d}{|\Omega|}, \quad |K| \leq \frac{|Q^+ \setminus U|}{|\Omega|} \leq \frac{(2L + 2R)^d - (2L - 4R)^d}{|\Omega|} \leq \frac{CL^{d-1}R}{|\Omega|}$$

Thus

$$\frac{|B_1|(2L)^d}{(2\pi)^d} \lambda^{\frac{d}{2}} \leq \frac{(2L - 2R)^d}{|\Omega|} N(\lambda, \Omega) + \frac{CL^{d-1}R}{|\Omega|} N(2^d(\lambda + 1), \Omega).$$

Dividing both sides for  $(2L)^d$  and sending  $L \rightarrow \infty$  we obtain

$$\frac{|B_1|}{(2\pi)^d} \lambda^{\frac{d}{2}} \leq \frac{1}{|\Omega|} N(\lambda, \Omega)$$



which is equivalent to

$$N(\lambda, \Omega) \geq \frac{|B_1||\Omega|}{(2\pi)^d} \lambda^{\frac{d}{2}}.$$

This completes the proof of Pólya's conjecture for Neumann eigenvalues in tiling domains.

□

# Chapter 7

## Many-body quantum systems

We consider a system of  $N$  **identical fermions** in  $\mathbb{R}^d$ . From first principles of quantum mechanics, the total energy of the system is described by a self-adjoint operator  $H_N$  on **the anti-symmetric space**  $L_a^2(\mathbb{R}^{dN})$ , which is a subspace of  $L^2(\mathbb{R}^{dN})$  containing wave functions satisfying

$$\Psi_N(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -\Psi_N(x_1, \dots, x_j, \dots, x_i, \dots, x_N), \quad \forall i \neq j, \quad \forall x_i \in \mathbb{R}^d,$$

or equivalently

$$\Psi_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}) = \text{sign}(\sigma)\Psi_N(x_1, \dots, x_N), \quad \forall \sigma \in S_N$$

where  $S_N$  is the permutation group of  $\{1, 2, \dots, N\}$ . A typical **many-body Schrödinger operator** has the form

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

where  $h_i$  is the copy of the operator  $h$  on  $L^2(\mathbb{R}^d)$  acting on the  $i$ -th variable  $x_i \in \mathbb{R}^d$ , namely

$$h_i = 1 \otimes \cdots \otimes 1 \otimes \underbrace{h}_{i\text{-th variable}} \otimes 1 \otimes \cdots \otimes 1,$$

and similarly for the two-body interaction  $W_{ij}$ . The ground state energy is

$$E_N = \inf_{\|\Psi\|_{L^2}=1} \langle \Psi, H_N \Psi \rangle.$$

If a ground state exists, then it solves **the Schrödinger equation**

$$H_N \Psi = E_N \Psi.$$

## 7.1 Slater determinants

**Definition** (Slater determinants). *For any functions  $\{u_i\}_{i=1}^N$  in  $L^2(\mathbb{R}^d)$ , define*

$$\begin{aligned} (u_1 \wedge u_2 \wedge \dots \wedge u_N)(x_1, \dots, x_N) &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \text{sign}(\sigma) u_1(x_{\sigma(1)}) \dots u_N(x_{\sigma(N)}) \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \text{sign}(\sigma) u_{\sigma(1)}(x_1) \dots u_{\sigma(N)}(x_N) \\ &= \frac{1}{\sqrt{N!}} \det \left[ (u_i(x_j))_{1 \leq i, j \leq N} \right]. \end{aligned}$$

*Clearly it is an anti-symmetric function in  $L_a^2(\mathbb{R}^{dN})$ .*

**Theorem.** *Let  $\{u_i\}_{i=1}^\infty$  be an orthonormal basis for  $L^2(\mathbb{R}^d)$ . Then the Slater determinants*

$$\{u_{i_1} \wedge u_{i_2} \wedge \dots \wedge u_{i_N} \mid i_1, \dots, i_N \in \mathbb{N}, i_1 < i_2 < \dots < i_N\}$$

*form an orthonormal basis for  $L_a^2(\mathbb{R}^{dN})$ .*

*Proof. Step 1.* First, we check that

$$\{u_{i_1} \wedge u_{i_2} \wedge \dots \wedge u_{i_N} \mid i_1, \dots, i_N \in \mathbb{N}, i_1 < i_2 < \dots < i_N\}$$

are orthonormal functions in  $L_a^2(\mathbb{R}^{dN})$ . For  $i_1 < i_2 < \dots < i_N$  and  $j_1 < j_2 < \dots < j_N$  we can write

$$\begin{aligned} &\left\langle u_{i_1} \wedge u_{i_2} \wedge \dots \wedge u_{i_N}, u_{j_1} \wedge u_{j_2} \wedge \dots \wedge u_{j_N} \right\rangle \\ &= \left\langle (N!)^{-1/2} \sum_{\sigma \in S_N} \text{sign}(\sigma) u_{i_{\sigma(1)}}(x_1) \dots u_{i_{\sigma(N)}}(x_N), (N!)^{-1/2} \text{sign}(\tau) \sum_{\tau \in S_N} u_{j_{\tau(1)}}(x_1) \dots u_{j_{\tau(N)}}(x_N) \right\rangle \\ &= (N!)^{-1} \sum_{\sigma, \tau \in S_N} \text{sign}(\sigma) \text{sign}(\tau) \langle u_{i_{\sigma(1)}}, u_{j_{\tau(1)}} \rangle \dots \langle u_{i_{\sigma(N)}}, u_{j_{\tau(N)}} \rangle \\ &= (N!)^{-1} \sum_{\sigma, \tau \in S_N} \text{sign}(\sigma) \text{sign}(\tau) \delta_{i_{\sigma(1)}, j_{\tau(1)}} \dots \delta_{i_{\sigma(N)}, j_{\tau(N)}} \end{aligned}$$

$$= \begin{cases} 1, & \text{if } (i_1, \dots, i_N) = (j_1, \dots, j_N), \\ 0, & \text{otherwise.} \end{cases}$$

It remains to prove that

$$L_a^2(\mathbb{R}^{dN}) = \overline{\text{Span} \{u_{i_1} \wedge u_{i_2} \wedge \dots \wedge u_{i_N} \mid i_1, \dots, i_N \in \mathbb{N}, i_1 < i_2 < \dots < i_N\}}.$$

**Step 2.** We prove that if  $\Omega_1, \Omega_2$  are two measure spaces, then

$$L^2(\Omega_1 \times \Omega_2) \cong L^2(\Omega_1) \otimes L^2(\Omega_2) := \overline{\text{Span} \{u \otimes v \mid u \in L^2(\Omega_1), v \in L^2(\Omega_2)\}}$$

where we used the usual notation of **tensor product**

$$(u \otimes v)(x, y) = u(x)v(y).$$

More precisely, we prove that if  $\{u_i\}_{i \in \mathbb{N}}$  is an orthonormal basis for  $L^2(\Omega_1)$  and  $\{v_i\}_{i \in \mathbb{N}}$  for  $L^2(\Omega_2)$ , then  $\{u_i \otimes v_j\}_{i, j \in \mathbb{N}}$  is an orthonormal basis for  $L^2(\Omega_1 \times \Omega_2)$ .

- $\{u_i \otimes v_j\}_{i, j \in \mathbb{N}}$  are orthonormal functions in  $L^2(\Omega_1 \times \Omega_2)$  as

$$\langle u_i \otimes v_j, u_\ell \otimes v_k \rangle = \langle u_i, u_\ell \rangle \langle v_j, v_k \rangle = \delta_{i\ell} \delta_{jk}.$$

- $\{u_i \otimes v_j\}_{i, j \in \mathbb{N}}$  is complete: Assume that  $f \in L^2(\Omega_1 \times \Omega_2)$  and  $f \perp u_i \otimes v_j$  for all  $i, j$ , then by Fubini's theorem

$$\begin{aligned} 0 &= \langle f, u_i \otimes v_j \rangle = \iint_{\Omega_1 \times \Omega_2} \overline{f(x, y)} u_i(x) v_j(y) d\mu_1(x) d\mu_2(y) \\ &= \int_{\Omega_1} u_i(x) \underbrace{\int_{\Omega_2} \overline{f(x, y)} v_j(y) d\mu_2(y)}_{g_j(x)} d\mu_1(x). \end{aligned}$$

Because  $\{u_i\}_{i \in \mathbb{N}}$  is an orthonormal basis for  $L^2(\Omega_1)$ , we must have  $g_j \equiv 0$  in  $L^2(\Omega_1)$ , namely for a.e.  $x \in \Omega_1$ ,

$$\int \overline{f(x, y)} v_j(y) = 0.$$

Since this holds for  $j \in \mathbb{N}$  and  $\{v_j\}_{j \in \mathbb{N}}$  is an orthonormal basis for  $L^2(\Omega)$ , we find that for a.e.  $x \in \Omega_1$ , for a.e.  $y \in \Omega_2$ ,  $f(x, y) = 0$ . Thus  $f \equiv 0$  in  $L^2(\Omega_1 \times \Omega_2)$ .

Using the above result and by induction, we find that

$$L^2(\mathbb{R}^{dN}) = L^2(\mathbb{R}^d) \otimes L^2(\mathbb{R}^d) \otimes \cdots \otimes L^2(\mathbb{R}^d) = L^2(\mathbb{R}^d)^{\otimes N}$$

and if  $\{u_i\}_{i \in \mathbb{N}}$  is an orthonormal basis for  $L^2(\mathbb{R}^d)$ , then

$$\{u_{i_1} \otimes \cdots \otimes u_{i_N} \mid i_1, \dots, i_N \in \mathbb{N}\}$$

is an orthonormal basis for  $L^2(\mathbb{R}^{dN})$ .

**Step 3.** We define the operator  $P_N$  on  $L^2(\mathbb{R}^{dN})$  by

$$(P_N \Psi_N)(x_1, \dots, x_N) = \frac{1}{N!} \sum_{\sigma \in S_N} \text{sign}(\sigma) \Psi_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}), \quad \forall \Psi_N \in L^2(\mathbb{R}^d).$$

Then  $P_N$  is a projection as  $(P_N)^2 = P_N$ :

$$\begin{aligned} (P_N)^2 \Psi_N(x_1, \dots, x_N) &= P_N \frac{1}{N!} \sum_{\sigma \in S_N} \text{sign}(\sigma) \Psi_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}) = \\ &= \frac{1}{N!} \sum_{\sigma \in S_N} \text{sign}(\sigma) \frac{1}{N!} \sum_{\tau \in S_N} \text{sign}(\tau) \Psi_N(x_{\tau \circ \sigma(1)}, \dots, x_{\tau \circ \sigma(N)}) = \\ &= \frac{1}{N!} \sum_{\sigma \in S_N} \frac{1}{N!} \underbrace{\sum_{\tau \in S_N} \text{sign}(\tau \circ \sigma) \Psi_N(x_{\tau \circ \sigma(1)}, \dots, x_{\tau \circ \sigma(N)})}_{\text{independent of } \sigma} = \\ &= \frac{1}{N!} \sum_{\tau \in S_N} \text{sign}(\tau) \Psi_N(x_{\tau(1)}, \dots, x_{\tau(N)}) = P_N \Psi_N(x_1, \dots, x_N). \end{aligned}$$

Moreover,  $P_N \Psi_N \in L_a^2(\mathbb{R}^{dN})$  for all  $\Psi_N \in L^2(\mathbb{R}^{dN})$  and  $P_N \Psi_N = \Psi_N$  if  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$ :

$$\begin{aligned} P_N \Psi_N(x_1, \dots, x_N) &= \frac{1}{N!} \sum_{\sigma \in S_N} \text{sign}(\sigma) \Psi_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}) \\ &= \frac{1}{N!} \sum_{\sigma \in S_N} \underbrace{(\text{sign}(\sigma))^2}_{=1} \Psi_N(x_1, \dots, x_N) \\ &= \Psi_N(x_1, \dots, x_N). \end{aligned}$$

Therefore,

$$L_a^2(\mathbb{R}^{dN}) = P_N L^2(\mathbb{R}^{dN})$$

**Step 4.** From Step 2 and Step 3, we obtain

$$L_a^2(\mathbb{R}^{dN}) = \overline{\text{Span} \{P_N(u_{i_1} \otimes \cdots \otimes u_{i_N}) \mid i_1, \dots, i_N \in \mathbb{N}\}}.$$

It remains to compute

$$\begin{aligned} P_N(u_{i_1} \otimes \cdots \otimes u_{i_N})(x_1, \dots, x_N) &= P_N(u_{i_1}(x_1) \cdots u_{i_N}(x_N)) \\ &= \frac{1}{N!} \sum_{\sigma \in S_N} \text{sign}(\sigma) u_{i_1}(x_{\sigma(1)}) \cdots u_{i_N}(x_{\sigma(N)}) \\ &= \frac{1}{N!} \det \left[ (u_{i_k}(x_\ell))_{1 \leq k, \ell \leq N} \right]. \end{aligned}$$

Thus

$$P_N(u_{i_1} \otimes \cdots \otimes u_{i_N}) = \frac{1}{\sqrt{N!}} u_{i_1} \wedge u_{i_2} \wedge \cdots \wedge u_{i_N}.$$

Consequently, if  $i_k = i_\ell$  for some  $k \neq \ell$ , then  $P_N(u_{i_1} \otimes \cdots \otimes u_{i_N}) = 0$ . Also,

$$P_N(u_{i_{\sigma(1)}} \otimes \cdots \otimes u_{i_{\sigma(N)}}) = \text{sign}(\sigma) P_N(u_{i_1} \otimes \cdots \otimes u_{i_N}).$$

Hence,  $L_a^2(\mathbb{R}^{dN})$  is equal to

$$\begin{aligned} &\text{Span} \{P_N(u_{i_1} \otimes \cdots \otimes u_{i_N}) \mid i_1, \dots, i_N \in \mathbb{N}\} \\ &= \text{Span} \{u_{i_1} \wedge u_{i_2} \wedge \cdots \wedge u_{i_N} \mid i_1, \dots, i_N \in \mathbb{N}, i_1 < i_2 < \cdots < i_N\}. \end{aligned}$$

Thus the Slater determinants form an orthonormal basis for  $L_a^2(\mathbb{R}^{dN})$ .  $\square$

## 7.2 Reduced density matrices

For many applications, the wave functions in  $L_a^2(\mathbb{R}^{dN})$  have too many variables for practical computations. Therefore, it is often useful to consider its **reduced density matrices** which are simpler to analyze.

**Definition.** Let  $\Psi_N$  be a normalized wave function in  $L_a^2(\mathbb{R}^{dN})$ . The **one-body density matrix**  $\gamma_{\Psi_N}^{(1)}$  of  $\Psi_N$  is a trace class operator on  $L^2(\mathbb{R}^d)$  with kernel

$$\gamma_{\Psi_N}^{(1)}(x, y) = N \int_{\mathbb{R}^{d(N-1)}} \Psi_N(x, x_2, \dots, x_N) \overline{\Psi_N(y, x_2, \dots, x_N)} dx_2 \cdots dx_N.$$

Thus  $\gamma_{\Psi_N}^{(1)} \geq 0$  and  $\text{Tr} \gamma_{\Psi_N}^{(1)} = N$ . The “diagonal part” of  $\gamma_{\Psi_N}^{(1)}$  is the one-body density

$$\rho_{\Psi_N}(x) = \gamma_{\Psi_N}^{(1)}(x, x) = N \int_{\mathbb{R}^{d(N-1)}} |\Psi_N(x, x_2, \dots, x_N)|^2 dx_2 \dots dx_N$$

which satisfies  $\rho_{\Psi_N} \geq 0$ ,  $\int_{\mathbb{R}^d} \rho_{\Psi_N} = N$ .

In application, the one-body density matrix  $\gamma_{\Psi_N}^{(1)}$  is sufficient to encode the expectation against every one-body observable:

$$\left\langle \Psi_N, \sum_{i=1}^N h_i \Psi_N \right\rangle = \text{Tr}(h \gamma_{\Psi_N}^{(1)}).$$

Here as usual  $\text{Tr}(h \gamma_{\Psi_N}^{(1)}) = \text{Tr}((\gamma_{\Psi_N}^{(1)})^{1/2} h (\gamma_{\Psi_N}^{(1)})^{1/2})$  makes sense when  $h$  is bounded from below on  $L^2(\mathbb{R}^d)$ . In particular, if  $V$  is a multiplication operator on  $L^2(\mathbb{R}^d)$ , then

$$\left\langle \Psi_N, \sum_{i=1}^N V(x_i) \Psi_N \right\rangle = \text{Tr}(V \gamma_{\Psi_N}^{(1)}) = \int_{\mathbb{R}^d} V(x) \rho_{\Psi_N}(x) dx.$$

Historically, the formalism of density matrices was introduced by **John von Neumann** in 1927. In his development of quantum statistical mechanics, the name “density matrix” is related to its analogue in classical statistical mechanics, namely a probability measure on the phase-space  $\mathbb{R}^d \times \mathbb{R}^d$ . In this general setting, a **mixed quantum state** of  $N$  fermions is a trace class operator  $\Gamma_N$  on  $L_a^2(\mathbb{R}^{dN})$  with

$$\Gamma_N \geq 0, \quad \text{Tr} \Gamma_N = 1.$$

Its one-body density matrix is obtained by taking the partial trace of all but 1 particle

$$\Gamma_N^{(1)} = N \text{Tr}_{2 \rightarrow N} \Gamma_N.$$

Thus  $\Gamma_N^{(1)}$  is a trace class operator on  $L^2(\mathbb{R}^d)$  with

$$\Gamma_N^{(1)} \geq 0, \quad \text{Tr} \Gamma_N^{(1)} = N.$$

In terms of kernels, we can write

$$\Gamma_N^{(1)}(x, y) = N \int_{\mathbb{R}^{d(N-1)}} \Gamma_N(x, x_2, \dots, x_N; y, x_2, \dots, x_N) dx_2 \dots dx_N$$

which is conceptually related to the **the marginal distribution** in probability theory. In particular, if  $\Gamma_N$  is a **pure state**, namely  $\Gamma_N = |\Psi_N\rangle\langle\Psi_N|$  with a normalized function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$ , then

$$\Gamma_N(x_1, \dots, x_N; y_1, \dots, y_N) = \Psi_N(x_1, \dots, x_N) \overline{\Psi_N(y_1, \dots, y_N)}$$

and  $\Gamma_N^{(1)}$  boils down to the operator  $\gamma_{\Psi_N}^{(1)}$  we defined before.

A key consequence of the anti-symmetry assumption is

**Theorem** (Pauli's exclusion principle). *For every normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$ , we have*

$$0 \leq \gamma_{\Psi_N}^{(1)} \leq 1 \quad \text{on } L^2(\mathbb{R}^d).$$

Remarks:

- Without the anti-symmetry assumption,  $\gamma_{\Psi_N}^{(1)}$  may have an eigenvalue as large as  $N$ . In fact, if

$$\Psi_N(x_1, \dots, x_N) = u^{\otimes N}(x_1, \dots, x_N) = u(x_1) \dots u(x_N)$$

with a normalized function  $u \in L^2(\mathbb{R}^d)$ , then  $\gamma_{\Psi_N}^{(1)} = N|u\rangle\langle u|$ .

- The physical interpretation of Pauli's exclusion principle is that

**“two quantum particles cannot occupy the same quantum state”.**

A less precise version of this principle can be seen easily from the anti-symmetry assumption: if  $x_i = x_j$  for  $i \neq j$ , then

$$\Psi_N(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -\Psi_N(x_1, \dots, x_j, \dots, x_i, \dots, x_N) = 0.$$

However, the operator inequality  $0 \leq \gamma_{\Psi_N}^{(1)} \leq 1$  is much deeper than the fact that the wave functions vanish on the diagonal set.

It is easy to verify Pauli's exclusion principle for Slater determinants

**Exercise.** *Let  $\{u_i\}_{i=1}^N$  be orthonormal functions in  $L^2(\mathbb{R}^d)$  and consider the Slater determinant  $\Psi_N = u_1 \wedge u_2 \wedge \dots \wedge u_N$ . Prove that the one-body density matrix of  $\Psi_N$*



is

$$\gamma_{\Psi_N}^{(1)} = \sum_{i=1}^N |u_i\rangle\langle u_i|.$$

**Proof of Pauli's exclusion principle in the general case.** We want to prove that for every normalized function  $u \in L^2(\mathbb{R}^d)$ , then

$$\langle u, \gamma_{\Psi_N}^{(1)} u \rangle \leq 1.$$

By the definition of the one-body density matrix  $\gamma_{\Psi_N}^{(1)}$ , we can write

$$\langle u, \gamma_{\Psi_N}^{(1)} u \rangle = \left\langle \Psi_N, \sum_{j=1}^N (P_u)_{x_j} \Psi_N \right\rangle, \quad P_u = |u\rangle\langle u|.$$

Thus we need to prove that

$$A = \sum_{j=1}^N (P_u)_{x_j} \leq 1 \quad \text{on } L_a^2(\mathbb{R}^{dN}).$$

Take an orthonormal basis  $\{u_i\}_{i=1}^\infty$  for  $L^2(\mathbb{R}^d)$  such that  $u_1 = u$ . We claim that

$$A = \sum_{1 \leq i_1 < i_2 < \dots < i_N} |u_{i_1} \wedge \dots \wedge u_{i_N}\rangle\langle u_{i_1} \wedge \dots \wedge u_{i_N}|$$

and the desired result follows as the Slater determinants form an orthonormal basis for  $L_a^2(\mathbb{R}^{dN})$ . Indeed, for every  $1 \leq i_1 < i_2 < \dots < i_N$  we have

$$\begin{aligned} Au_{i_1} \wedge \dots \wedge u_{i_N} &= \sum_{j=1}^N (P_u)_{x_j} \sum_{\sigma \in S_N} \frac{1}{\sqrt{N!}} \text{sign}(\sigma) u_{i_{\sigma(1)}}(x_1) \dots u_{i_{\sigma(N)}}(x_N) = \\ &= \sum_{j=1}^N \sum_{\sigma \in S_N} \frac{1}{\sqrt{N!}} \text{sign}(\sigma) u_{i_{\sigma(1)}}(x_1) \dots \left( P_u u_{i_{\sigma(j)}}(x_j) \right) \dots u_{i_{\sigma(N)}}(x_N) \\ &= \sum_{j=1}^N \sum_{\sigma \in S_N} \frac{1}{\sqrt{N!}} \text{sign}(\sigma) u_{i_{\sigma(1)}}(x_1) \dots \left( \delta_{1, i_{\sigma(j)}} u_{i_{\sigma(j)}}(x_j) \right) \dots u_{i_{\sigma(N)}}(x_N) \\ &= \mathbb{1}(1 \in \{i_1, \dots, i_N\}) \sum_{\sigma \in S_N} \frac{1}{\sqrt{N!}} \text{sign}(\sigma) u_{i_{\sigma(1)}}(x_1) \dots u_{i_{\sigma(j)}}(x_j) \dots u_{i_{\sigma(N)}}(x_N) \\ &= \begin{cases} u_{i_1} \wedge \dots \wedge u_{i_N} & 1 = i_1, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

This completes the proof of Pauli's exclusion principle.  $\square$

The following result of **Coleman** (1963) tells us that the condition  $0 \leq \gamma \leq 1$  in Pauli's exclusion principle is optimal.

**Theorem** (Admissible one-body density matrices). *Let  $\gamma$  be a trace class operator on  $L^2(\mathbb{R}^d)$  such that*

$$0 \leq \gamma \leq 1 \quad \text{on } L^2(\mathbb{R}^d), \quad \text{Tr } \gamma = N \in \mathbb{N}.$$

*Then there exists a mixed state  $\Gamma_N$  on  $L_a^2(\mathbb{R}^{dN})$ , namely a non-negative operator on  $L_a^2(\mathbb{R}^{dN})$  with  $\text{Tr } \Gamma_N = 1$ , such that its one-body density matrix is  $\Gamma_N^{(1)} = \gamma$ .*

Remarks:

- If  $\gamma$  is a projection, namely  $\gamma = \gamma^2$ , then  $\Gamma_N$  can be chosen to be a pure state  $\Gamma_N = |\Psi_N\rangle\langle\Psi_N|$  and  $\Psi_N$  is simply a Slater determinant.
- In general, it might be not possible to choose  $\Gamma_N$  to be a pure state (see an exercise below).

**Exercise.** *Let  $\gamma$  be a trace class operator on  $L^2(\mathbb{R}^d)$  such that*

$$0 \leq \gamma \leq 1 \quad \text{on } L^2(\mathbb{R}^d), \quad \text{Tr } \gamma = N \in \mathbb{N}.$$

*Assume further that  $\gamma$  has  $N - 1$  eigenvalues equal to 1, but  $\gamma$  is not a projection. Prove that there exists no normalized function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $\gamma_{\Psi_N}^{(1)} = \gamma$ .*

*Proof of Coleman's theorem.* We claim that  $\gamma$  can be written as a convex combination of rank- $N$  projections, namely

$$\gamma = \sum_{k=1}^{\infty} c_k \gamma_k, \quad c_k \geq 0, \quad \sum_{k=1}^{\infty} c_k = 1, \quad \gamma_k \text{ is a rank-}N \text{ projection for all } k.$$

Then any  $\gamma_k$  is the one-body density matrix of a Slater determinant  $\Psi_{N,k} \in L_a^2(\mathbb{R}^{dN})$ , and we can simply take

$$\Gamma_N = \sum_{k=1}^{\infty} c_k |\Psi_{N,k}\rangle\langle\Psi_{N,k}|.$$

We construct the sequences  $\{c_k\}$  and  $\{\gamma_k\}$  by induction.

**For**  $k = 1$ , by Spectral Theorem we can write

$$\gamma = \sum_{n=1}^{\infty} \lambda_n(\gamma) |u_n\rangle\langle u_n|$$

where  $\{u_n\}_{n=1}^{\infty}$  is an orthonormal family in  $L^2(\mathbb{R}^d)$  and

$$\lambda_1(\gamma) \geq \lambda_2(\gamma) \geq \dots \geq 0, \quad \sum_{n=1}^{\infty} \lambda_n(\gamma) = N.$$

If  $\lambda_{N+1}(\gamma) = 0$ , then  $\gamma$  is a rank- $N$  projection and we can stop. Otherwise, we take

$$\varepsilon_1 = \min\{\lambda_N(\gamma), 1 - \lambda_{N+1}(\gamma)\} \in (0, 1)$$

and write

$$\gamma = \varepsilon_1 \gamma_1 + (1 - \varepsilon_1) \tilde{\gamma}_1, \quad \gamma_1 = \sum_{n=1}^N |u_n\rangle\langle u_n|, \quad 0 \leq \tilde{\gamma}_1 \leq 1, \quad \text{Tr } \tilde{\gamma}_1 = N.$$

More precisely,

$$\tilde{\gamma}_1 = \sum_{n=1}^N \frac{\lambda_n(\gamma) - \varepsilon_1}{1 - \varepsilon_1} |u_n\rangle\langle u_n| + \sum_{n=N+1}^{\infty} \frac{\lambda_n(\gamma)}{1 - \varepsilon_1} |u_n\rangle\langle u_n|.$$

The choice of  $\varepsilon_1$  ensures that  $0 \leq \tilde{\gamma}_1 \leq 1$ . Of course, we take  $c_1 = \varepsilon_1$ .

**For**  $k = 2$ , we can repeat the above argument with  $\gamma$  replaced by  $\tilde{\gamma}_1$ . More precisely, if  $\tilde{\gamma}_1$  is a rank- $N$  projection, then we can stop. Otherwise, we can write

$$\tilde{\gamma}_1 = \varepsilon_2 \gamma_2 + (1 - \varepsilon_2) \tilde{\gamma}_2, \quad \gamma_2 \text{ rank-}N \text{ projection}, \quad 0 \leq \tilde{\gamma}_2 \leq 1, \quad \text{Tr } \tilde{\gamma}_2 = N$$

and  $\varepsilon_2 = \min\{\lambda_N(\tilde{\gamma}_1), 1 - \lambda_{N+1}(\tilde{\gamma}_1)\} \in (0, 1)$ . Thus

$$\gamma = \varepsilon_1 \gamma_1 + (1 - \varepsilon_1) \tilde{\gamma}_1 = \varepsilon_1 \gamma_1 + (1 - \varepsilon_1) \varepsilon_2 \gamma_2 + (1 - \varepsilon_1)(1 - \varepsilon_2) \tilde{\gamma}_2.$$

Hence, we take  $c_2 = (1 - \varepsilon_1) \varepsilon_2$ .

**For every**  $k \geq 3$ , by induction we have

$$\gamma = c_1 \gamma_1 + c_2 \gamma_2 + \dots + c_{k-1} \gamma_{k-1} + (1 - \varepsilon_1) \dots (1 - \varepsilon_{k-1}) \tilde{\gamma}_{k-1}.$$

If  $\tilde{\gamma}_{k-1}$  is a rank- $N$  projection, then we can stop. Otherwise, we can write

$$\tilde{\gamma}_{k-1} = \varepsilon_k \gamma_k + (1 - \varepsilon_k) \tilde{\gamma}_k, \quad \gamma_k \text{ rank-}N \text{ projection,} \quad 0 \leq \tilde{\gamma}_k \leq 1, \quad \text{Tr } \tilde{\gamma}_k = N$$

and  $\varepsilon_k = \min\{\lambda_N(\tilde{\gamma}_{k-1}), 1 - \lambda_{N+1}(\tilde{\gamma}_{k-1})\} \in (0, 1)$ . Thus

$$\begin{aligned} \gamma &= c_1 \gamma_1 + \dots + c_{k-1} \gamma_{k-1} + (1 - \varepsilon_1) \dots (1 - \varepsilon_{k-1}) \varepsilon_k \gamma_k + (1 - \varepsilon_1) \dots (1 - \varepsilon_{k-1}) (1 - \varepsilon_k) \tilde{\gamma}_k \\ &= c_1 \gamma_1 + \dots + c_{k-1} \gamma_{k-1} + c_k \gamma_k + \prod_{i=1}^k (1 - \varepsilon_i) \tilde{\gamma}_k \end{aligned}$$

with

$$c_k = (1 - \varepsilon_1) \dots (1 - \varepsilon_{k-1}) \varepsilon_k = \varepsilon_k \prod_{i=1}^{k-1} (1 - \varepsilon_i).$$

**Conclusion:** In order to conclude

$$\gamma = \sum_{k=1}^{\infty} c_k \gamma_k$$

it remains to show that

$$\sum_{k=1}^{\infty} c_k = 1 \iff \prod_{k=1}^{\infty} (1 - \varepsilon_k) = 0$$

Assume by contradiction that

$$\sum_{k=1}^{\infty} c_k = 1 - \delta < 1 \iff \prod_{k=1}^{\infty} (1 - \varepsilon_k) = \delta > 0.$$

Then we can write

$$\gamma = \sum_{k=1}^{\infty} c_k \gamma_k + \delta \tilde{\gamma}_{\infty}, \quad 0 \leq \tilde{\gamma}_{\infty} \leq 1, \quad \text{Tr } \tilde{\gamma}_{\infty} = N.$$

From the induction formula

$$\gamma = \sum_{k=1}^M c_k \gamma_k + \prod_{k=1}^M (1 - \varepsilon_k) \tilde{\gamma}_M, \quad 0 \leq \tilde{\gamma}_M \leq 1, \quad \text{Tr } \tilde{\gamma}_M = N$$

we find that  $\tilde{\gamma}_M \rightarrow \tilde{\gamma}_{\infty}$  strongly in trace class as  $M \rightarrow \infty$ . Consequently,

$$\varepsilon_{M+1} = \min\{\lambda_N(\tilde{\gamma}_M), 1 - \lambda_{N+1}(\tilde{\gamma}_M)\} \rightarrow \min\{\lambda_N(\tilde{\gamma}_{\infty}), 1 - \lambda_{N+1}(\tilde{\gamma}_{\infty})\} > 0.$$

Here in the latter inequality, we have used that

$$\lambda_N(\tilde{\gamma}_\infty) > 0, \quad \lambda_{N+1}(\tilde{\gamma}_\infty) \leq \frac{1}{N+1} \sum_{i=1}^{N+1} \lambda_i(\tilde{\gamma}_\infty) \leq \frac{1}{N+1} \text{Tr} \tilde{\gamma}_\infty = \frac{N}{N+1}.$$

But the fact that  $\lim_{M \rightarrow \infty} \varepsilon_M > 0$  just contradicts the assumption

$$\prod_{k=1}^{\infty} (1 - \varepsilon_k) > 0.$$

This completes the proof. □

Remark: In general, we can also define higher reduced density matrices. For example, if  $\Psi_N$  is a normalized wave function in  $L_a^2(\mathbb{R}^{dN})$ , then **the two-body density matrix**  $\gamma_{\Psi_N}^{(2)}$  is a trace class operator on  $L_a^2(\mathbb{R}^{2d})$  with kernel

$$\gamma_{\Psi_N}^{(2)}(x_1, x_2; y_1, y_2) = \frac{N(N-1)}{2} \int_{\mathbb{R}^{d(N-2)}} \Psi_N(x_1, x_2, z_3, \dots, z_N) \overline{\Psi_N(y_1, y_2, z_3, \dots, z_N)} dz_3 \dots dz_N.$$

Then for every two-body operator  $W$  on  $L_a^2(\mathbb{R}^{2d})$ , we can write

$$\left\langle \Psi_N, \sum_{1 \leq i < j \leq N} W_{ij} \Psi_N \right\rangle = \text{Tr}(W \gamma_{\Psi_N}^{(2)}).$$

However, this formulation is hard to use in practice because there is no complete characterization for the two-body density matrices. This so-called **N-representability problem** is one reason making the interacting systems much harder than the non-interacting ones.

## 7.3 Ideal Fermi gas

We consider a non-interacting system of  $N$  fermions in  $\mathbb{R}^d$ . In principle, the idea gas is “solvable”.

**Theorem.** *Let  $h$  be a self-adjoint operator on  $L^2(\mathbb{R}^d)$ . Assume that  $h$  is bounded from below. Then for every  $N \geq 1$ , the Hamiltonian*

$$H_N = \sum_{i=1}^N h_i \quad \text{on } L_a^2(\mathbb{R}^{dN})$$

is bounded from below with the core domain  $D(h) \otimes \cdots \otimes D(h)$ , and hence it can be defined as a self-adjoint operator by Friedrichs' method. Moreover, its ground state energy is

$$E_N = \sum_{i=1}^N \mu_i(h)$$

where  $\mu_1(h) \leq \mu_2(h) \leq \dots$  are the min-max values of  $h$ .

*Proof.* Using  $h \geq \mu_1(h)$ , we have the obvious lower bound

$$H_N = \sum_{i=1}^N h_i \geq N\mu_1(h).$$

Thus  $H_N$  is bounded from below and hence it can be defined as a self-adjoint operator by Friedrichs' method. It remains to compute the ground state energy  $E_N$  of  $H_N$ .

**Lower bound.** For every normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  we can write

$$\left\langle \Psi_N, \sum_{i=1}^N h_i \Psi_N \right\rangle = \text{Tr}(h\gamma_{\Psi_N}^{(1)}).$$

Since  $0 \leq \gamma_{\Psi_N}^{(1)} \leq 1$  (by Pauli's exclusion principle) and  $\text{Tr} \gamma_{\Psi_N}^{(1)} = N$ , we obtain

$$\begin{aligned} \text{Tr}(h\gamma_{\Psi_N}^{(1)}) &\geq \inf \left\{ \sum_{n=1}^{\infty} \nu_n \langle u_n, h u_n \rangle \mid \{u_n\}_{n=1}^{\infty} \text{ ONF}, 0 \leq \nu_n \leq 1, \sum_{n=1}^{\infty} \nu_n = N \right\} \\ &= \inf \left\{ \sum_{n=1}^N \langle u_n, h u_n \rangle \mid \{u_n\}_{n=1}^N \text{ ONF} \right\} = \sum_{i=1}^N \mu_i(h) \end{aligned}$$

thanks to the min-max principle.

**Upper bound.** Consider the Slater determinant

$$\Psi_N = u_1 \wedge \cdots \wedge u_N.$$

with orthonormal functions  $\{u_i\}_{i=1}^N$  in  $L^2(\mathbb{R}^d)$  (we can take  $\{u_i\}_{i=1}^N \subset D(h)$ ). Then using

$$\gamma_{\Psi_N}^{(1)} = \sum_{n=1}^N |u_n\rangle \langle u_n|$$

we have

$$\langle \Psi_N, H_N \Psi_N \rangle = \text{Tr}(h\gamma_{\Psi_N}^{(1)}) = \sum_{n=1}^N \langle u_n, h u_n \rangle.$$

Thus

$$E_N \leq \inf \left\{ \sum_{n=1}^N \langle u_n, h u_n \rangle \mid \{u_n\}_{n=1}^N \text{ ONF} \right\} = \sum_{i=1}^N \mu_i(h).$$

□

Let us consider an example of the hydrogen-like atom.

**Theorem.** For every  $N \in \mathbb{N}$ , consider the Hamiltonian

$$H_N = \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{N}{|x_i|} \right) \quad \text{on } L_a^2(\mathbb{R}^{3N}).$$

Then  $H_N$  is a self-adjoint operator with the quadratic form domain  $H_a^1(\mathbb{R}^{3N})$  and its ground state energy satisfies

$$E_N = -N^{7/3} \left( \frac{(3)^{1/3}}{4} + o(1)_{N \rightarrow \infty} \right).$$

By the general theory of the ideal gas, we know that

$$E_N = \sum_{i=1}^N \mu_i(-\Delta - N|x|^{-1})$$

where  $\mu_i(-\Delta - N|x|^{-1})$  is the  $i$ -th min-max value of the Schrödinger operator  $-\Delta - N|x|^{-1}$  on  $L^2(\mathbb{R}^3)$ . Actually, the spectrum of  $-\Delta - N|x|^{-1}$  is known completely: it has negative eigenvalues

$$-\frac{N^2}{4k^2} \text{ with multiplicity } k^2, \text{ with } k = 1, 2, \dots$$

Hence, if we can write

$$N = 1^2 + 2^2 + \dots + M^2 + M', \quad 0 \leq M' < (M+1)^2,$$

then

$$E_N = \sum_{i=1}^N \mu_i(-\Delta - N|x|^{-1}) = \sum_{k=1}^M -\frac{N^2}{4k^2} \times k^2 - \frac{M'N^2}{4(M+1)^2} = -\frac{MN^2}{4} - \frac{M'N^2}{4(M+1)^2}.$$

Using

$$1^2 + 2^2 + \dots + M^2 = \frac{M(M+1)(2M+1)}{6} = N + o(N)_{N \rightarrow \infty}$$

we find that

$$M = (3N)^{1/3} + o(N^{1/3})_{N \rightarrow \infty}.$$

Thus

$$E_N = -N^{7/3} \left( \frac{3^{1/3}}{4} + o(1)_{N \rightarrow \infty} \right).$$

On the other hand, we can also prove the above asymptotic formula for  $E_N$  using the semiclassical estimates developed previously, without having to compute all the negative eigenvalues explicitly.

**Proof of the asymptotic formula for  $E_N$  by semiclassical estimates.** First, by rescaling  $x \mapsto N^{1/3}x$  we can write

$$E_N = \sum_{i=1}^N \mu_i(-\Delta - N|x|^{-1}) = \sum_{i=1}^N \mu_i(-N^{2/3}\Delta - N^{4/3}|x|^{-1}) = N^{2/3} \sum_{i=1}^N \mu_i(-\Delta - N^{2/3}|x|^{-1}).$$

Thus we need to show that

$$\tilde{E}_N := \sum_{i=1}^N \mu_i(-\Delta - N^{2/3}|x|^{-1}) = -N^{5/3} \left( \frac{3^{1/3}}{4} + o(1)_{N \rightarrow \infty} \right).$$

**Lower bound.** For every constant  $a > 0$ , we can write

$$\begin{aligned} \tilde{E}_N &:= \sum_{i=1}^N \mu_i(-\Delta - N^{2/3}|x|^{-1}) = \sum_{i=1}^N \mu_i(-\Delta + N^{2/3}(a - |x|^{-1})) - aN^{5/3} \\ &\geq \text{Tr}(-\Delta + N^{2/3}(a - |x|^{-1}))_- - aN^{5/3}. \end{aligned}$$

By Weyl's law,

$$\begin{aligned} \text{Tr}(-\Delta + N^{2/3}(a - |x|^{-1}))_- &= -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} |N^{2/3}(a - |x|^{-1})_-|^{5/2} dx + o((N^{2/3})^{5/2})_{N \rightarrow \infty} \\ &= -N^{5/3} \left( L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} |(a - |x|^{-1})_-|^{5/2} dx + o(1)_{N \rightarrow \infty} \right) \end{aligned}$$

where

$$L_{1,3}^{\text{cl}} = \int_{\mathbb{R}^3} |(|2\pi k|^2 - 1)_-| dk = \left( \frac{2}{d+2} \cdot \frac{|B_1|}{(2\pi)^d} \right)_{|d=3} = \frac{2}{5} \cdot \frac{(4\pi/3)}{(2\pi)^3} = \frac{1}{15\pi^2}$$



and

$$\int_{\mathbb{R}^3} |(a - |x|^{-1})_-|^{5/2} dx = \frac{4\pi}{\sqrt{a}} \int_0^1 (r^{-1} - 1)^{5/2} r^2 dr = \frac{4\pi}{\sqrt{a}} \cdot \frac{5\pi}{16}.$$

Thus in summary,

$$\tilde{E}_N \geq -N^{5/3} \left( \frac{1}{15\pi^2} \cdot \frac{4\pi}{\sqrt{a}} \cdot \frac{5\pi}{16} + a + o(1)_{N \rightarrow \infty} \right) = -N^{5/3} \left( \frac{1}{12\sqrt{a}} + a + o(1)_{N \rightarrow \infty} \right)$$

We can optimize over  $a > 0$ , namely choose  $a > 0$  such that

$$\frac{1}{24\sqrt{a}} = \frac{1}{24\sqrt{a}} = a = \left( \frac{1}{24\sqrt{a}} \cdot \frac{1}{24\sqrt{a}} \cdot a \right)^{1/3} = (24)^{-2/3}.$$

Hence,

$$\tilde{E}_N \geq -N^{5/3} \left( \frac{3}{(24)^{2/3}} + o(1)_{N \rightarrow \infty} \right) = -N^{5/3} \left( \frac{3^{1/3}}{4} + o(1)_{N \rightarrow \infty} \right).$$

**Upper bound.** We need to show that with  $V(x) = a - |x|^{-1}$ ,  $a = (24)^{-2/3}$  and  $\lambda = N^{2/3}$

$$\sum_{i=1}^N \mu_i(-\Delta + \lambda V(x)) \leq -\lambda^{5/2} \left( L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} |V_-|^{5/2} + o(1)_{\lambda \rightarrow \infty} \right).$$

By the min-max principle,

$$\sum_{i=1}^N \mu_i(-\Delta + \lambda V(x)) \leq \text{Tr}((-\Delta + \lambda V(x))\gamma)$$

for any trace class operator  $\gamma$  on  $L^2(\mathbb{R}^3)$  satisfying

$$0 \leq \gamma \leq 1, \quad \text{Tr} \gamma = N.$$

We will construct a trial operator  $\gamma$  using the coherent state method. Take a radial function  $0 \leq G \in C_c^\infty(\mathbb{R}^3)$  satisfying  $\|G\|_{L^2} = 1$  and denote

$$F_{k,y}(x) = e^{2\pi i k \cdot x} G(x - y).$$

As in the proof of Weyl's law, we choose

$$\gamma := \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 + \lambda V(y) < 0) dk dy.$$

Then  $0 \leq \gamma \leq 1$  on  $L^2(\mathbb{R}^3)$  by the resolution of identity and

$$\begin{aligned} \text{Tr } \gamma &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \mathbf{1}(|2\pi k|^2 + \lambda V(y) < 0) dk dy = L_{0,3}^{\text{cl}} \int_{\mathbb{R}^3} |\lambda V_-|^{3/2} \\ &= \frac{|B_1|}{(2\pi)^3} \lambda^{3/2} \int_{\mathbb{R}^3} |(a - |x|^{-1})_-|^{3/2} dx = \frac{(4\pi/3)}{(2\pi)^3} \lambda^{3/2} 4\pi a^{-3/2} \frac{\pi}{16} = N. \end{aligned}$$

In the last equality we have used  $\lambda = N^{2/3}$  and  $a = (24)^{-2/3}$ . Thus proceeding as in the Weyl's law upper bound, we find that

$$\begin{aligned} \text{Tr}((-\Delta + \lambda V(x))\gamma) &= -L_{1,3}^{\text{cl}} \lambda^{5/2} \int_{\mathbb{R}^3} |V_-|^{5/2} + L_{0,3}^{\text{cl}} \lambda^{5/2} \int_{\mathbb{R}^3} |V_-|^{\frac{3}{2}} (V * G^2 - V) dy \\ &\quad + \|\nabla G\|_{L^2}^2 L_{0,3}^{\text{cl}} \lambda^{3/2} \int_{\mathbb{R}^3} |V_-|^{3/2}. \end{aligned}$$

Note that  $V_- = (a - |x|^{-1})_- \in L^1 \cap L^{3-}(\mathbb{R}^3)$  and  $\text{supp } V_- = \{|x| \leq a^{-1}\}$ . Hence

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-5/2} \text{Tr}((-\Delta + \lambda V(x))\gamma) \leq -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} |V_-|^{5/2} + L_{0,3}^{\text{cl}} \int_{\mathbb{R}^3} |V_-|^{\frac{3}{2}} (V * G^2 - V) dy.$$

Moreover,

$$V * G^2 - V = (a - |x|^{-1}) * G^2 - (a - |x|^{-1}) = |x|^{-1} - |x|^{-1} * G^2$$

Hence,

$$|V_-|^{3/2} (V * G^2 - V) = |V_-|^{3/2} (|x|^{-1} - |x|^{-1} * G^2) \leq |V_-|^{3/2} (f - f * G^2)$$

where

$$f(x) = |x|^{-1} \mathbf{1}(|x| \leq a^{-1}).$$

Since  $f \in L^{3-}(\mathbb{R}^3)$ , by choosing  $G_n(x) = n^{3/2} G_1(nx)$  for a fixed function  $\|G_1\|_{L^2} = 1$ , we obtain

$$f - f * G^2 \rightarrow 0$$

strongly in  $L^{3^-}(\mathbb{R}^3)$ . Combining with  $|V_-|^{3/2} \in L^1 \cap L^{2^-}$ , we get

$$|V_-|^{3/2}(V * G_n^2 - V) = |V_-|^{3/2}(f - f * G_n^2) \rightarrow 0$$

strongly in  $L^1(\mathbb{R}^3)$ . Thus we conclude that

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-5/2} \operatorname{Tr}((-\Delta + \lambda V(x))\gamma) \leq -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} |V_-|^{5/2}.$$

This completes the proof of the upper bound. □

# Chapter 8

## Thomas–Fermi theory

### 8.1 Density functional theory

In density functional theory, instead of considering a complicated wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  one simply looks at its one–body density

$$\rho_{\Psi_N}(x) = N \int_{(\mathbb{R}^d)^{N-1}} |\Psi_N(x, x_2, \dots, x_N)|^2 dx_2 \dots dx_N,$$

which satisfies the simple constraints

$$\rho_{\Psi_N}(x) \geq 0, \quad \int_{\mathbb{R}^d} \rho_{\Psi_N}(x) dx = N.$$

**Theorem** (Representability). *Let  $d \geq 1$  and  $0 \leq \rho \in L^1(\mathbb{R}^d)$ ,  $\int_{\mathbb{R}^d} \rho = N \in \mathbb{N}$ . Then there exists a normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $\rho_{\Psi_N} = \rho$ . We can choose  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  to be a Slater determinant. Moreover, we can choose  $\Psi_N \in H_a^1(\mathbb{R}^{dN})$  if and only if  $\sqrt{\rho} \in H^1(\mathbb{R}^d)$ .*

*Proof.* **Step 1.** For every  $0 \leq \rho \in L^1(\mathbb{R}^d)$ ,  $\int_{\mathbb{R}^d} \rho = N \in \mathbb{N}$ , we can take

$$\Psi_N = u_1 \wedge u_2 \wedge \dots \wedge u_N, \quad u_k = \frac{\sqrt{\rho(x)}}{\sqrt{N}} e^{2\pi i k f(x)}$$

with a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . Then all of  $u_k$  are normalized in  $L^2(\mathbb{R}^d)$  and

$$\sum_{k=1}^N |u_k(x)|^2 = \rho(x).$$

It remains to choose  $f$  such that  $\{u_k\}_{k=1}^N$  are orthogonal. We use an idea of Harriman and Lieb (1981). Using the notation  $x = (x^1, y) \in \mathbb{R} \times \mathbb{R}^{d-1}$ , we define

$$f(x) = f(x^1) = \frac{1}{N} \int_{-\infty}^{x^1} \left( \int_{\mathbb{R}^{d-1}} \rho(t, y) dy \right) dt.$$

Then

$$f'(x^1) = \frac{1}{N} \int_{\mathbb{R}^{d-1}} \rho(x^1, y) dy.$$

Thus  $f' \in L^1(\mathbb{R})$ , hence  $f$  is at least continuous. Moreover,  $f$  is increasing (as  $\rho \geq 0$ ) and

$$\lim_{x^1 \rightarrow -\infty} f(x^1) = 0, \quad \lim_{x^1 \rightarrow \infty} f(x^1) = 1.$$

Moreover, for every  $k \neq \ell$ , we have

$$\begin{aligned} \int_{\mathbb{R}^d} \overline{u_k(x)} u_\ell(x) dx &= \int_{\mathbb{R}^d} \frac{\rho(x)}{N} e^{2\pi i(\ell-k)f(x)} dx = \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{d-1}} \frac{\rho(x^1, y)}{N} e^{2\pi i(\ell-k)f(x^1)} dy \right) dx^1 \\ &= \int_{\mathbb{R}} f'(x^1) e^{2\pi i(\ell-k)f(x^1)} dx^1 = \int_0^1 e^{2\pi i(\ell-k)s} ds = 0. \end{aligned}$$

Thus  $\{u_k\}_{k=1}^N$  are orthonormal, and hence  $\Psi_N$  is a Slater determinant with

$$\rho_{\Psi_N} = \sum_{k=1}^N |u_k(x)|^2 = \rho.$$

**Step 2.** If  $\sqrt{\rho} \in H^1(\mathbb{R}^d)$ , then

$$\nabla u_k(x) = \left( \nabla(\sqrt{\rho(x)}) + \sqrt{\rho(x)} 2\pi i k f'(x^1) \right) \frac{e^{2\pi i k f(x^1)}}{\sqrt{N}} \in L^2(\mathbb{R}^d).$$

Here  $\nabla(\sqrt{\rho(x)}) \in L^2(\mathbb{R}^d)$  by the assumption  $\sqrt{\rho} \in H^1(\mathbb{R}^d)$ . For the second term,  $f'(x^1)\sqrt{\rho(x)} \in L^2(\mathbb{R}^d)$  because  $\sqrt{\rho} \in H^1(\mathbb{R}^d) \in L^p(\mathbb{R}^d)$  for some  $p > 2$  by Sobolev inequality and  $f'(x^1) \in$

$L^q(\mathbb{R})$  for all  $1 \leq q < \infty$ . The latter fact can be seen from

$$\begin{aligned} f''(x^1) &= \frac{1}{N} \int_{\mathbb{R}^{d-1}} \partial_{x_1} \rho(x^1, y) dy = \frac{1}{N} \int_{\mathbb{R}^{d-1}} \partial_{x_1} (\sqrt{\rho(x^1, y)})^2 dy \\ &= \frac{2}{N} \int_{\mathbb{R}^{d-1}} \sqrt{\rho(x^1, y)} \partial_{x_1} \sqrt{\rho(x^1, y)} dy \in L^1(\mathbb{R}) \end{aligned}$$

and Sobolev inequality  $W^{1,1}(\mathbb{R}) \subset L^q(\mathbb{R})$  for all  $1 \leq q < \infty$ . Thus we conclude that if  $\sqrt{\rho} \in H^1(\mathbb{R}^d)$ , then all  $u_k$  belong to  $H^1(\mathbb{R}^d)$ , and hence the Slater determinant  $\Psi_N$  constructed above belongs to  $H_a^1(\mathbb{R}^{dN})$ .

**Step 3.** If  $\Psi_N \in H_a^1(\mathbb{R}^{dN})$  (not necessarily a Slater determinant), then  $\sqrt{\rho_{\Psi_N}} \in H^1(\mathbb{R}^d)$  because we have the **Hoffmann-Ostenhof<sup>2</sup> inequality**

$$\left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \geq \int_{\mathbb{R}^d} |\nabla \sqrt{\rho_{\Psi_N}}|^2.$$

Using the one-body density matrix  $\gamma_{\Psi_N}^{(1)}$ , this inequality can be written as

$$\mathrm{Tr}(-\Delta \gamma_{\Psi_N}^{(1)}) \geq \int_{\mathbb{R}^d} |\nabla \sqrt{\rho_{\Psi_N}}|^2.$$

Using the spectral decomposition

$$\gamma_{\Psi_N}^{(1)} = \sum_{n \geq 1} |f_n\rangle \langle f_n|$$

(here  $f_n$ 's not necessarily normalized), the above inequality can be written as

$$\sum_{n \geq 1} \int_{\mathbb{R}^d} |\nabla f_n|^2 \geq \int_{\mathbb{R}^d} \left| \nabla \left( \sum_{n \geq 1} |f_n|^2 \right)^{1/2} \right|^2.$$

For the sum of two functions, this follows directly from the diamagnetic inequality  $|\nabla|\varphi(x)|| \leq |\nabla\varphi(x)|$  for every  $\varphi \in H^1(\mathbb{R}^d)$ . For the general sum, we can do induction.  $\square$

The idea of describing a quantum state using only its one-body density goes back to **Thomas** and **Fermi** in 1927. It was conceptually pushed forward by a variational principle of **Hohenberg** and **Kohn** in 1964. Here we will follow the approach by **Levy** (1979) and **Lieb** (1983). In general, given any Hamiltonian  $H_N$  on  $L_a^2(\mathbb{R}^{dN})$ , the ground state energy can be

rewritten as

$$E_N = \inf_{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1} \langle \Psi_N, H_N \Psi_N \rangle = \inf_{\substack{\rho \geq 0 \\ \int_{\mathbb{R}^d} \rho = N}} \inf_{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1, \rho_{\Psi_N}=\rho} \langle \Psi_N, H_N \Psi_N \rangle.$$

This motivates the definition of the **Levy–Lieb density functional**

$$\mathcal{L}_N(\rho) = \inf_{\substack{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1 \\ \rho_{\Psi_N}=\rho}} \langle \Psi_N, H_N \Psi_N \rangle, \quad \forall \rho \geq 0, \int_{\mathbb{R}^d} \rho = N.$$

Thus the ground state problem of  $H_N$  becomes

$$E_N = \inf_{\substack{\rho \geq 0 \\ \int_{\mathbb{R}^d} \rho = N}} \mathcal{L}_N(\rho).$$

This looks simple, but of course the complication of the many-body problem is now hidden in the determination of  $\mathcal{L}_N$ . In principle, computing  $\mathcal{L}_N$  is very hard. However, we may try to develop approximations which capture some properties of  $\mathcal{L}_N$  when  $N \rightarrow \infty$ .

Consider a typical Hamiltonian  $H_N$  on  $L_a^2(\mathbb{R}^{dN})$  of the form

$$H_N = \sum_{i=1}^N \left( -h^2 \Delta_{x_i} + V(x_i) \right) + \lambda \sum_{1 \leq i < j \leq N} w(x_i - x_j).$$

Here  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  is an external potential and  $w : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $w(x) = w(-x)$ , is an interaction potential. The parameter  $h > 0$  plays the role of Planck's constant and  $\lambda > 0$  corresponds to the strength of the interaction.

For the external potential, we have the exact formula

$$\left\langle \Psi_N, \sum_{i=1}^N V(x_i) \Psi_N \right\rangle = \int_{\mathbb{R}^d} V(x) \rho_{\Psi_N}(x) dx.$$

For the kinetic and interaction terms, there are no exact expression in terms of  $\rho_{\Psi_N}$ . However, the **semiclassical approximation** suggests that

$$\left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \approx K_d^{\text{cl}} \int_{\mathbb{R}^d} \rho_{\Psi_N}^{1+2/d}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{2/d}}$$

while the **mean-field approximation** tell us

$$\left\langle \Psi_N, \sum_{1 \leq i < j \leq N}^N w(x_i - x_j) \Psi_N \right\rangle \approx \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho(x) \rho(y) w(x - y) dx dy.$$

Putting all this together, we arrive at the **Thomas-Fermi approximation**

$$\left\langle \Psi_N, H_N \Psi_N \right\rangle \approx K_d^{\text{cl}} h^2 \int_{\mathbb{R}^d} \rho_{\Psi_N}^{1+2/d} + \int_{\mathbb{R}^d} V \rho_{\Psi_N} + \frac{\lambda}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho_{\Psi_N}(x) \rho_{\Psi_N}(y) w(x - y) dx dy.$$

In particular, this suggests that

$$\mathcal{L}_N(\rho) \approx K_d^{\text{cl}} h^2 \int_{\mathbb{R}^d} \rho^{1+2/d} + \int_{\mathbb{R}^d} V \rho + \frac{\lambda}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho(x) \rho(y) w(x - y) dx dy.$$

We will justify this approximation in the next section (following my joint work with **Nina Gottschling** (2018)). As we will see, the Thomas-Fermi theory is correct to the leading order in the **semiclassical mean-field regime**

$$h \sim N^{-1/d}, \quad \lambda \sim N^{-1}, \quad N \rightarrow \infty.$$

This is the choice making all three terms on the Thomas-Fermi density functional comparable (all are of order  $N$ ).

Historically, the Thomas-Fermi approximation was proposed for the atomic Hamiltonian, when  $V$  and  $w$  are Coulomb potentials in  $\mathbb{R}^3$ . We will consider it in a more general context, and then pay a special attention to the Coulomb case at the end.

## 8.2 Convergence of the kinetic density functional

Recall the semiclassical approximation

$$\left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \approx K_d^{\text{cl}} \int_{\mathbb{R}^d} \rho_{\Psi_N}^{1+2/d}, \quad K_d^{\text{cl}} = \frac{d}{d+2} \cdot \frac{(2\pi)^2}{|B_1|^{2/d}}.$$

It is convenient to introduce

$$f_{\Psi_N} = \frac{\rho_{\Psi_N}}{N} = \int_{(\mathbb{R}^d)^{N-1}} |\Psi_N(x, x_2, \dots, x_N)|^2 dx_2 \dots dx_N, \quad f \geq 0, \quad \int_{\mathbb{R}^d} f = 1.$$



Thus the above approximation becomes

$$\frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \approx K_d^{\text{cl}} \int_{\mathbb{R}^d} f_{\Psi_N}^{1+2/d}.$$

We can justify this approximation as follows.

**Theorem.** *For all  $d \geq 1$ , the followings hold true when  $N \rightarrow \infty$ .*

- (i) (Lower bound) *If the normalized wave functions  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  satisfy that  $f_{\Psi_N} \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$ , then*

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \geq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}.$$

- (ii) (Upper bound) *For every*

$$0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f = 1,$$

*there exist Slater determinants  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $f_{\Psi_N} \rightarrow f$  strongly in  $L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  and*

$$\limsup_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}.$$

Remark: For upper bound, we do not assume  $\sqrt{f} \in H^1(\mathbb{R}^d)$ , and hence it is not always possible to choose  $f_{\Psi_N} = f$  (due to the Hoffmann-Ostenhof<sup>2</sup> inequality).

*Proof.* **Lower bound.** For every function  $0 \leq U \in C_c^\infty(\mathbb{R}^d)$ , we can write

$$\frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle = \frac{1}{N^{1+2/d}} \text{Tr} \left[ (-\Delta - N^{2/d}U) \gamma_{\Psi_N}^{(1)} \right] + \int_{\mathbb{R}^d} U f_{\Psi_N}.$$

By the Pauli's exclusion principle  $0 \leq \gamma_{\Psi_N}^{(1)} \leq 1$  and Weyl's law on the sum of negative eigenvalues, we can estimate

$$\text{Tr} \left[ (-\Delta - N^{2/d}U) \gamma_{\Psi_N}^{(1)} \right] \geq \text{Tr}[-\Delta - N^{2/d}U]_- = -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} |N^{2/d}U|^{1+d/2} + o((N^{2/d})^{1+d/2})$$

$$= N^{1+2/d} \left( -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} U^{1+d/2} + o(1)_{N \rightarrow \infty} \right).$$

Moreover, using  $f_{\Psi_N} \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$  we find that

$$\int_{\mathbb{R}^d} U f_{\Psi_N} \rightarrow \int_{\mathbb{R}^d} U f.$$

Thus

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle \geq -L_{1,d}^{\text{cl}} \int_{\mathbb{R}^d} U^{1+d/2} + \int_{\mathbb{R}^d} U f.$$

Optimizing over  $U$  (i.e. choosing  $U = \text{const.} \cdot f^{2/d}$ ) we conclude that

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle \geq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}.$$

**Upper bound.** We can follow the coherent state approach in the proof of Weyl's law to deduce the upper bound, but in this way it is not easy to keep the important constraint that  $\Psi_N$  are Slater determinants. In the following, we will follow a more direct approach, which is close to Weyl's original method and Thomas–Fermi heuristic argument.

**Step 1 (Slater determinants of Dirichlet Laplacian on a cube).** Consider the Dirichlet Laplacian  $-\Delta$  on  $Q = [0, L]^d$ . Recall that it has eigenvalues  $|\pi k/L|^2$ ,  $k \in \mathbb{N}^d$ , with eigenfunctions

$$u_k(x) = \prod_{i=1}^d \left[ \sqrt{\frac{2}{L}} \sin \left( \frac{\pi k^i x^i}{L} \right) \right], \quad k = (k^i)_{i=1}^d, \quad x = (x^i)_{i=1}^d \in \mathbb{R}^d.$$

The ground state of the  $M$ -body kinetic operator  $\sum_{j=1}^M (-\Delta_{x_j})$  is the Slater determinant  $\Psi_M^S$  made of the first  $M$  eigenfunctions  $\{u_k\}$ . It is straightforward to see that when  $M \rightarrow \infty$ ,

$$\frac{1}{M^{1+2/d}} \left\langle \Psi_M^S, \sum_{i=1}^M (-\Delta_{x_i}) \Psi_M^S \right\rangle = \frac{1}{M^{1+2/d}} \sum_{k \in S_M} \left| \frac{\pi k}{L} \right|^2 \rightarrow \frac{K_d^{\text{cl}}}{|Q|^{2/d}}$$

and

$$f_{\Psi_M^S} = \frac{1}{M} \sum_{\substack{\text{first } M \\ \text{eigenfunctions}}} |u_k|^2 \rightarrow \frac{\mathbb{1}_Q}{|Q|} \quad \text{strongly in } L^p(Q), \quad \forall p \in [1, \infty).$$

**Step 2 (Slater determinants of step-function densities).** Let

$$0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} f = 1.$$

Let  $\{Q\}$  be a finite family of disjoint cubes, whose construction will be specified in the next step. In the following we only consider cubes  $Q$  such that

$$\int_Q f > 0.$$

We can find an integer number

$$M_Q \in \left( N \int_Q f \right) \pm [-1, 1]$$

such that

$$\sum_Q M_Q = N \int_{\mathbb{R}^d} f = N.$$

Now for every  $Q$ , consider the first  $M_Q$  eigenfunctions  $\{u_j^Q\}_{j=1}^{M_Q}$  of the Dirichlet Laplacian  $-\Delta$  on  $Q$ . These functions can be trivially extended to zero outside  $Q$  to become a function in  $H_0^1(\mathbb{R}^d)$ . Since the cubes  $\{Q\}$  are disjoint, the  $N$  functions  $\bigcup_Q \{u_j^Q\}_{j=1}^{M_Q} \subset H_0^1(\mathbb{R}^d)$  are orthonormal in  $L^2(\mathbb{R}^d)$ . Let  $\Psi_N^S \in L_a^2(\mathbb{R}^{dN})$  be the Slater determinant made of this orthogonal family. Then in the limit  $N \rightarrow \infty$ , using the fact that

$$\frac{M_Q}{N} \rightarrow \int_Q f > 0$$

and the calculation in Step 1 for each cube, we get

$$\begin{aligned} & \frac{1}{N^{1+2/d}} \left\langle \Psi_N^S, \sum_{i=1}^N -\Delta_{x_i} \Psi_N^S \right\rangle = \frac{1}{N^{1+2/d}} \sum_Q \sum_{i=1}^{M_Q} \|\nabla u_i^Q\|^2 \\ & = \sum_Q \left[ \frac{1}{M_Q^{1+2/d}} \sum_{i=1}^{M_Q} \|\nabla u_i^Q\|^2 \right] \left| \frac{M_Q}{N} \right|^{1+2/d} \rightarrow \sum_Q \frac{K_d^{\text{cl}}}{|Q|^{2/d}} \cdot \left| \int_Q f \right|^{1+2/d} \\ & = K_d^{\text{cl}} \sum_Q |Q| \left| \sum_Q \frac{1}{|Q|} \int_Q f \right|^{1+2/d} \leq K_d^{\text{cl}} \sum_Q \int_Q f^{1+2/d} \quad (\text{by Jensen's inequality}) \\ & \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d} \end{aligned}$$

and

$$f_{\Psi_N^S} = \frac{1}{N} \sum_Q \sum_{i=1}^{M_Q} |u_i^Q|^2 = \sum_Q \sum_{i=1}^{M_Q} \frac{|u_i^Q|^2}{M_Q} \cdot \frac{M_Q}{N} \rightarrow \sum_Q \frac{\mathbf{1}_Q}{|Q|} \int_Q f$$

strongly in  $L^p(\mathbb{R}^d)$  for all  $1 \leq p < \infty$ .

**Step 3 (Conclusion).** Since  $0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$ , for every  $k \geq 1$  we can find a finite family of disjoint cubes  $\{Q\}$  such that

$$\left\| f - \sum_Q \mathbf{1}_Q \bar{f}^Q \right\|_{L^1(\mathbb{R}^d)} + \left\| f - \sum_Q \mathbf{1}_Q \bar{f}^Q \right\|_{L^{1+2/d}(\mathbb{R}^d)} \leq k^{-1}, \quad \bar{f}^Q := \frac{1}{|Q|} \int_Q f.$$

Using this collection of cubes, for every  $N \geq 1$  we can construct a Slater determinant  $\Psi_N^k \in L_a^2(\mathbb{R}^{dN})$  as in Step 2. Thus there exists  $N_k > 0$  such that for every  $N \geq N_k$ ,

$$\frac{1}{N^{1+2/d}} \left\langle \Psi_N^k, \sum_{i=1}^N -\Delta_{x_i} \Psi_N^k \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d} + k^{-1}$$

and

$$\left\| f_{\Psi_N^k} - \sum_Q \mathbf{1}_Q \bar{f}^Q \right\|_{L^1} + \left\| f_{\Psi_N^k} - \sum_Q \mathbf{1}_Q \bar{f}^Q \right\|_{L^{1+2/d}} \leq k^{-1}.$$

By the triangle inequality, for every  $N \geq N_k$ ,

$$\left\| f_{\Psi_N^k} - f \right\|_{L^1(\mathbb{R}^d)} + \left\| f_{\Psi_N^k} - f \right\|_{L^{1+2/d}(\mathbb{R}^d)} \leq 2k^{-1}.$$

Now we conclude using a standard diagonal argument. By induction in  $k$ , we can choose the above sequence  $N_k$  such that  $N_{k+1} > N_k$ . Since  $\lim_{k \rightarrow \infty} N_k = \infty$ , we can find  $\lim_{N \rightarrow \infty} k_N = \infty$  slowly such that

$$N \geq N_{k_N}.$$

Thus the Slater determinant  $\Psi_N = \Psi_N^{k_N} \in S_N$  constructed as above satisfies, as  $N \rightarrow \infty$ ,

$$\frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d} + k_N^{-1} \rightarrow K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}$$

and

$$\left\| f_{\Psi_N} - f \right\|_{L^1(\mathbb{R}^d)} + \left\| f_{\Psi_N} - f \right\|_{L^{1+2/d}(\mathbb{R}^d)} \leq 2k_N^{-1} \rightarrow 0.$$

This completes the proof of the theorem.  $\square$

Note that the above theorem is conceptually equivalent to Weyl's law for the sum of eigenvalues. For example, we can use this theorem to give another analysis for the hydrogen-like atom

$$H_N = \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{N}{|x_i|} \right) \quad \text{on } L_a^2(\mathbb{R}^{3N}).$$

We want to show that the ground state energy of  $H_N$  is

$$E_N = -N^{7/3} \left( \frac{(3)^{1/3}}{4} + o(1)_{N \rightarrow \infty} \right).$$

By rescaling  $x \mapsto N^{1/3}x$ , it is equivalent to prove that the Hamiltonian

$$\tilde{H}_N = \sum_{i=1}^N \left( -\frac{1}{N^{5/3}} \Delta_{x_i} - \frac{1}{N|x_i|} \right) \quad \text{on } L_a^2(\mathbb{R}^{3N})$$

has the ground state energy  $\tilde{E}_N = -(3)^{1/3}/4 + o(1)_{N \rightarrow \infty}$ .

**Another look at  $\tilde{H}_N$ . Lower bound.** Take an arbitrary normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that

$$\langle \Psi_N, H_N \Psi_N \rangle = \tilde{E}_N + O(N^{-1}).$$

By the Lieb–Thirring inequality

$$\langle \Psi_N, H_N \Psi_N \rangle = \frac{1}{N^{5/3}} \langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \rangle - \int_{\mathbb{R}^3} \frac{f_{\Psi_N}}{|x|} dx \geq K \int_{\mathbb{R}^3} f_{\Psi_N}^{5/3} - \int_{\mathbb{R}^3} \frac{f_{\Psi_N}}{|x|} dx.$$

with a constant  $K > 0$  independent of  $N$ . Moreover, using  $\int_{\mathbb{R}^3} f_{\Psi_N} = 1$  and Hölder's inequality we find that

$$\begin{aligned} \int_{\mathbb{R}^3} \frac{f_{\Psi_N}}{|x|} dx &\leq \int_{|x| \geq 1} \frac{f_{\Psi_N}}{|x|} dx + \int_{|x| < 1} \frac{f_{\Psi_N}}{|x|} dx \\ &\leq 1 + \|f_{\Psi_N}\|_{L^{5/3}(\mathbb{R}^3)} \| |x|^{-1} \mathbf{1}(|x| \leq 1) \|_{L^{5/2}(\mathbb{R}^3)} \leq 1 + C \|f_{\Psi_N}\|_{L^{5/3}(\mathbb{R}^3)}. \end{aligned}$$

Therefore,

$$K \int_{\mathbb{R}^3} f_{\Psi_N}^{5/3} - \int_{\mathbb{R}^3} \frac{f_{\Psi_N}}{|x|} dx \geq \frac{K}{2} \int_{\mathbb{R}^3} f_{\Psi_N}^{5/3} - C.$$

Thus  $\tilde{E}_N \geq -C$  and  $f_{\Psi_N}$  is bounded in  $L^{5/3}(\mathbb{R}^3)$ . Up to a subsequence, we can assume that

$f_{\Psi_N} \rightharpoonup f$  weakly in  $L^{5/3}(\mathbb{R}^3)$ . Then

$$\int_{\mathbb{R}^3} f \leq \liminf_{N \rightarrow \infty} \int_{\mathbb{R}^3} f_{\Psi_N} = 1.$$

Moreover, by the above theorem, we have

$$\liminf_{N \rightarrow \infty} \langle \Psi_N, H_N \Psi_N \rangle \geq K_3^{\text{cl}} \int_{\mathbb{R}^3} f^{5/3} - \int_{\mathbb{R}^3} \frac{f(x)}{|x|} dx$$

where

$$K_3^{\text{cl}} = \left( \frac{d}{d+2} \frac{(2\pi)^d}{|B_1|^{2/d}} \right)_{|d=3} = \frac{3}{5} \frac{(2\pi)^2}{(4\pi/3)^{2/3}} = \frac{3}{5} (6\pi^2)^{2/3}.$$

**Exercise.** Consider the Thomas–Fermi functional

$$\mathcal{E}^{\text{TF}}(f) = K_3^{\text{cl}} \int_{\mathbb{R}^3} f^{5/3} - \int_{\mathbb{R}^3} \frac{f(x)}{|x|} dx, \quad K_3^{\text{cl}} = \frac{3}{5} (6\pi^2)^{2/3}.$$

Prove that the variational problem

$$E = \inf \left\{ \mathcal{E}^{\text{TF}}(f) \mid 0 \leq f \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3), \int_{\mathbb{R}^3} f \leq 1 \right\}$$

has a unique minimizer  $f_0$ . Moreover,  $\int_{\mathbb{R}^3} f_0 = 1$  and  $E = -(3)^{1/3}/4$ .

This leads to the lower bound

$$\liminf_{N \rightarrow \infty} \tilde{E}_N = \liminf_{N \rightarrow \infty} \langle \Psi_N, H_N \Psi_N \rangle \geq -(3)^{1/3}/4$$

for a subsequence as  $N \rightarrow \infty$ . We then obtain the convergence for the whole sequence by a standard contradiction argument.

**Upper bound.** Let  $f_0$  be the Thomas–Fermi minimizer from the above exercise,  $\int_{\mathbb{R}^3} f_0 = 1$ . By the above theorem, we can find Slater determinants  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $f_{\Psi_N} \rightarrow f_0$  strongly in  $L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3)$  and

$$\limsup_{N \rightarrow \infty} \tilde{E}_N \leq \limsup_{N \rightarrow \infty} \langle \Psi_N, H_N \Psi_N \rangle = K_3^{\text{cl}} \int_{\mathbb{R}^3} f_0^{5/3} - \int_{\mathbb{R}^3} \frac{f_0(x)}{|x|} dx = -(3)^{1/3}/4.$$

□

It is **conjectured** that

$$K_d^{\text{cl}} \int_{\mathbb{R}^d} \rho_{\Psi_N}^{1+2/d} \leq \left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} \rho_{\Psi_N}^{1+2/d} + \int_{\mathbb{R}^d} |\nabla \sqrt{\rho_{\Psi_N}}|^2$$

$$\iff K_d^{\text{cl}} \int_{\mathbb{R}^d} f_{\Psi_N}^{1+2/d} \leq \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f_{\Psi_N}^{1+2/d} + N^{-2/d} \int_{\mathbb{R}^d} |\nabla \sqrt{f_{\Psi_N}}|^2,$$

where the lower bound holds for all  $d \geq 3$  (Lieb–Thirring conjecture) and the upper bound holds for all  $d \geq 1$ . The upper bound was proved by **March** and **Young** in 1958 for  $d = 1$ , but their proof cannot be extended to higher dimensions.

### 8.3 Convergence of the Levy–Lieb functional

Consider the Hamiltonian

$$H_N = \sum_{i=1}^N \left( -h^2 \Delta_{x_i} + V(x_i) \right) + \lambda \sum_{1 \leq i < j \leq N} w(x_i - x_j).$$

in the semiclassical mean–field regime

$$h = N^{-1/d}, \quad \lambda = N^{-1}.$$

We will prove that the rescaled Levy–Lieb density functional

$$\mathcal{E}_N(f) := \frac{\mathcal{L}_N(Nf)}{N} = \inf_{\substack{\|\Psi_N\|_{L^2_a(\mathbb{R}^{dN})} = 1 \\ f_{\Psi_N} = f}} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N}$$

converges to the Thomas–Fermi density functional

$$\mathcal{E}^{\text{TF}}(f) := K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d} + \int_{\mathbb{R}^d} V f + \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x) f(y) w(x-y) dx dy.$$

**Conditions on potentials.** *The potentials  $V, w : \mathbb{R}^d \rightarrow \mathbb{R}$  belong to  $L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$ . Moreover,  $w$  admits the decomposition*

$$w(x) = \int_0^\infty (\chi_r * \chi_r)(x) d\mu(r),$$

for a positive measure  $\mu$  on  $(0, \infty)$  and for a family of even functions  $0 \leq \chi_r \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [2 + d, \infty)$ .

The decomposition on  $w$  is equivalent to

$$\widehat{w}(k) = \int_0^\infty |\widehat{\chi}_r(k)|^2 d\mu(r).$$

So this essentially requires that  $\widehat{w}(k) \geq 0$ , plus some regularity. This holds for a large class of potentials, including Coulomb potentials. For example, in  $\mathbb{R}^3$  we have the **Fefferman-de la Llave formula**

$$\frac{1}{|x|} = \frac{1}{\pi} \int_0^\infty (\mathbb{1}_{B_r} * \mathbb{1}_{B_r})(x) \frac{dr}{r^5}, \quad \forall x \in \mathbb{R}^3 \setminus \{0\}$$

where  $\mathbb{1}_{B_r}$  is the characteristic function of the ball  $B(0, r)$  in  $\mathbb{R}^3$ .

**Exercise.** Let  $d \geq 1$  and let  $\mathbb{1}_{B_r}$  be the characteristic function of the ball  $B(0, r)$  in  $\mathbb{R}^d$ . Prove that for every  $0 < \lambda < d$ , there exists a constant  $C_{\lambda, d} > 0$  such that

$$\frac{1}{|x|^\lambda} = C_{\lambda, d} \int_0^\infty (\mathbb{1}_{B_r} * \mathbb{1}_{B_r})(x) \frac{dr}{r^{d+\lambda+1}}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}.$$

**Theorem** (Gamma convergence from Levy-Lieb to Thomas-Fermi functional). For all  $d \geq 1$ , when  $N \rightarrow \infty$ , the Levy-Lieb functional  $\mathcal{E}_N$  converges to the Thomas-Fermi functional  $\mathcal{E}^{\text{TF}}$  in the following sense:

(i) (Lower bound) For every sequence  $0 \leq f_N \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  such that  $\int_{\mathbb{R}^d} f_N = 1$  and  $f_N \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$ , then

$$\liminf_{N \rightarrow \infty} \mathcal{E}_N(f_N) \geq \mathcal{E}^{\text{TF}}(f).$$

(ii) (Upper bound) For every  $0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  such that  $\int_{\mathbb{R}^d} f = 1$ , there exists a sequence of Slater determinants  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $f_{\Psi_N} = f_N \rightarrow f$  strongly in  $L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$ , and

$$\limsup_{N \rightarrow \infty} \mathcal{E}_N(f_N) \leq \mathcal{E}^{\text{TF}}(f).$$

*Proof. Lower bound.* Consider a normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  with  $f_{\Psi_N} =$



$f_N \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$ . We have

$$\frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} = \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle + \int_{\mathbb{R}^d} V f_N + \frac{1}{N^2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi_N \right\rangle.$$

By the convergence of the kinetic functional, we have

$$\liminf_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N -\Delta_{x_i} \Psi_N \right\rangle \geq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}.$$

Moreover, since  $f_N \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$  and  $\|f_N\|_{L^1} = 1$ , by interpolation we have  $f_N \rightharpoonup f$  weakly in  $L^r(\mathbb{R}^d)$  for all  $r \in (1, 1 + 2/d]$ . Under the condition  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$ , we deduce that

$$\lim_{N \rightarrow \infty} N^{-1} \left\langle \Psi_N, \sum_{i=1}^N V(x_i) \Psi_N \right\rangle = \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} V f_N = \int_{\mathbb{R}^d} V f.$$

It remains to consider the interaction terms. Using

$$w(x - y) = \int_0^\infty d\mu(r) (\chi_r * \chi_r)(x - y) = \int_0^\infty d\mu(r) \int_{\mathbb{R}^d} dz \chi_r(x - z) \chi_r(y - z)$$

we find that

$$\left\langle \Psi_N, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi_N \right\rangle = \int_0^\infty d\mu(r) \int_{\mathbb{R}^d} dz \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} \chi_r(x_i - z) \chi_r(x_j - z) \Psi_N \right\rangle.$$

For every  $r > 0$  and  $z \in \mathbb{R}^d$ , by the Cauchy-Schwarz inequality we get

$$\begin{aligned} & \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} \chi_r(x_i - z) \chi_r(x_j - z) \Psi_N \right\rangle \\ &= \frac{1}{2} \left[ \left\langle \Psi_N, \left( \sum_{i=1}^N \chi_r(x_i - z) \right)^2 \Psi_N \right\rangle - \left\langle \Psi_N, \sum_{i=1}^N \chi_r^2(x_i - z) \Psi_N \right\rangle \right]_+ \\ &\geq \frac{1}{2} \left[ \left\langle \Psi_N, \sum_{i=1}^N \chi_r(x_i - z) \Psi_N \right\rangle^2 - \left\langle \Psi_N, \sum_{i=1}^N \chi_r^2(x_i - z) \Psi_N \right\rangle \right]_+ \\ &= \frac{1}{2} [N^2 (f_N * \chi_r)^2(z) - N (f_N * \chi_r^2)(z)]_+. \end{aligned}$$

Since  $f_N \rightharpoonup f$  weakly in  $L^r(\mathbb{R}^d)$  for all  $1 < r \leq 1 + 2/d$ , and  $\chi_r, \chi_r^2 \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$ , we find that

$$\begin{aligned}\lim_{N \rightarrow \infty} (f_N * \chi_r)(z) &= (f * \chi_r)(z), \\ \lim_{N \rightarrow \infty} (f_N * \chi_r^2)(z) &= (f * \chi_r^2)(z).\end{aligned}$$

Hence, for every  $r > 0$  and  $z \in \mathbb{R}^d$ ,

$$\begin{aligned}\liminf_{N \rightarrow \infty} N^{-2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} \chi_r(x_i - z) \chi_r(x_j - z) \Psi_N \right\rangle \\ \geq \liminf_{N \rightarrow \infty} N^{-2} \frac{1}{2} [N^2 (f_N * \chi_r)^2(z) - N (f_N * \chi_r^2)(z)]_+ = \frac{1}{2} (f * \chi_r)^2(z).\end{aligned}$$

Therefore, by Fatou's lemma,

$$\begin{aligned}\liminf_{N \rightarrow \infty} N^{-2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi_N \right\rangle \\ = \liminf_{N \rightarrow \infty} \int_0^\infty dr \int_{\mathbb{R}^d} dz N^{-2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} \chi_r(x_i - z) \chi_r(x_j - z) \Psi_N \right\rangle \\ \geq \int_0^\infty dr \int_{\mathbb{R}^d} dz \frac{1}{2} (f * \chi_r)^2(z) = \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x) f(y) w(x - y) dx dy.\end{aligned}$$

Here we have repeatedly use the decomposition of  $w$ . Thus in summary,

$$\liminf_{N \rightarrow \infty} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} \geq \mathcal{E}^{\text{TF}}(f).$$

Since  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  can be chosen arbitrarily under the sole condition  $f_{\Psi_N} = f_N$ , this leads the desired lower bond

$$\liminf_{N \rightarrow \infty} \mathcal{E}_N(f_N) \geq \mathcal{E}^{\text{TF}}(f).$$

**Upper bound.** Let  $0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  with  $\int_{\mathbb{R}^d} f = 1$ . Then by the convergence of the kinetic functional, there exist Slater determinants  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $f_{\Psi_N} = f_N \rightarrow f$  strongly in  $L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  and

$$\limsup_{N \rightarrow \infty} \frac{1}{N^{1+2/d}} \left\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \right\rangle \leq K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d}.$$

Since  $f_N \rightarrow f$  in  $L^r(\mathbb{R}^d)$  for all  $r \in [1, 1 + 2/d]$  and  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in$

$[1 + d/2, \infty)$ , we have

$$\lim_{N \rightarrow \infty} N^{-1} \left\langle \Psi_N, \sum_{i=1}^N V(x_i) \Psi_N \right\rangle = \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} V f_N = \int_{\mathbb{R}^d} V f.$$

Finally, for the interaction terms, since  $\Psi_N$  is a Slater determinants and  $w$  is non-negative, an explicit computation shows that

$$\begin{aligned} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi_N \right\rangle &= \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left[ \rho_{\Psi_N}(x) \rho_{\Psi_N}(y) - |\gamma_{\Psi_N}^{(1)}(x, y)|^2 \right] w(x - y) dx dy \\ &\leq \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \rho_{\Psi_N}(x) \rho_{\Psi_N}(y) w(x - y) dx dy \\ &= \frac{N^2}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f_N(x) f_N(y) w(x - y) dx dy. \end{aligned}$$

The convergence  $f_N \rightarrow f$  in  $L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  and the assumption  $w \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  imply that

$$f_N * w \rightarrow f * w \quad \text{in } L^\infty(\mathbb{R}^d)$$

by Young's inequality. Hence,

$$\begin{aligned} N^{-2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi_N \right\rangle &\leq \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f_N(x) f_N(y) w(x - y) dx dy \\ &\rightarrow \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x) f(y) w(x - y) dx dy. \end{aligned}$$

Putting all together we obtain the desired upper bound

$$\limsup_{N \rightarrow \infty} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} \leq \mathcal{E}^{\text{TF}}(f).$$

□

## 8.4 Convergence of ground state energy and ground states

The notion of Gamma convergence is useful for many applications. In particular, we can come back to the ground state problem of the Hamiltonian

$$H_N = \sum_{i=1}^N \left( -h^2 \Delta_{x_i} + V(x_i) \right) + \lambda \sum_{1 \leq i < j \leq N} w(x_i - x_j), \quad h = N^{-1/d}, \lambda = N^{-1}.$$

As above, we assume that  $V, w \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$  and

$$w(x) = \int_0^\infty (\chi_r * \chi_r)(x) d\mu(r),$$

for a positive measure  $\mu$  on  $(0, \infty)$  and for a family of even functions  $0 \leq \chi_r \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [2 + d, \infty)$ . Recall the Thomas–Fermi functional

$$\mathcal{E}^{\text{TF}}(f) := K_d^{\text{cl}} \int_{\mathbb{R}^d} f^{1+2/d} + \int_{\mathbb{R}^d} V f + \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x) f(y) w(x - y) dx dy.$$

and the TF energy

$$E^{\text{TF}} := \inf \left\{ \mathcal{E}^{\text{TF}}(f) : 0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d), \int_{\mathbb{R}^d} f \leq 1 \right\}.$$

**Exercise.** Given real-valued functions  $V, w \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$ . Prove that  $E^{\text{TF}}$  has a minimizer  $f^{\text{TF}}$  with  $\int_{\mathbb{R}^d} f^{\text{TF}} \leq 1$  and

$$E^{\text{TF}} = \inf \left\{ \mathcal{E}^{\text{TF}}(f) : 0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d), \int_{\mathbb{R}^d} f = 1 \right\}.$$

Prove that if we assume further  $\widehat{w} \geq 0$ , then  $f^{\text{TF}}$  is unique.

We have

**Theorem** (Convergence of ground state energy and ground states). *Let  $d \geq 1$ . The*

ground state energy  $E_N$  of  $H_N$  converges to the Thomas-Fermi energy:

$$\lim_{N \rightarrow \infty} \frac{E_N}{N} = E^{\text{TF}}.$$

Moreover, if  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  is a ground state for  $H_N$ , or more generally an approximate ground state in the sense that

$$\lim_{N \rightarrow \infty} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} = E^{\text{TF}},$$

then

$$f_{\Psi_N} \rightharpoonup f^{\text{TF}} \quad \text{weakly in } L^{1+2/d}(\mathbb{R}^d)$$

where  $f^{\text{TF}}$  is the unique Thomas-Fermi minimizer satisfying  $\int_{\mathbb{R}^d} f^{\text{TF}} \leq 1$ .

- It may happen that  $E_N$  has no minimizer, and/or  $E^{\text{TF}}$  has no minimizer satisfying  $\int_{\mathbb{R}^d} f^{\text{TF}} = 1$ . Nevertheless, the convergence of the ground state energy is always valid. In fact, the convergence of the ground state energy is valid under a very general condition on  $w$  (including negative potentials, e.g.  $w(x) = -|x|^{-1}$ ), as proved by **Fournais**, **Lewin** and **Solovej** (FLS-2018).
- This result justifies the validity of Thomas-Fermi in the atomic case, which was first proved by **Lieb** and **Simon** (1973). In this case we have Coulomb potentials in  $\mathbb{R}^3$

$$V(x) = -\frac{1}{|x|}, \quad w(x) = \frac{1}{|x|}.$$

In the literature, the atomic Hamiltonian is often written in the form

$$H_N^{\text{atom}} = \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{N}{|x_i|} \right) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \quad \text{on } L_a^2(\mathbb{R}^{3N}),$$

which corresponds to a **neutral atom** of  $N$  quantum electrons moving around a heavy nucleus fixed at  $0 \in \mathbb{R}^3$  of the nuclear charge  $Z = N$ , interacting via Coulomb forces. By changing the variable  $x \mapsto N^{-1/3}x$ ,  $H_N^{\text{atom}}$  is unitarily equivalent to

$$\begin{aligned} \tilde{H}_N^{\text{atom}} &= \sum_{i=1}^N \left( -N^{2/3} \Delta_{x_i} - \frac{N}{N^{-1/3}|x_i|} \right) + \sum_{1 \leq i < j \leq N} \frac{1}{N^{-1/3}|x_i - x_j|} \quad \text{on } L_a^2(\mathbb{R}^{3N}) \\ &= N^{4/3} \left[ \sum_{i=1}^N \left( -N^{-2/3} \Delta_{x_i} - \frac{1}{|x_i|} \right) + N^{-1} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \right]. \end{aligned}$$

Thus the above theorem tells us that

$$\frac{\inf \sigma(H_N^{\text{atom}})}{N^{7/3}} = \frac{\inf \sigma(\tilde{H}_N^{\text{atom}})}{N^{7/3}} \rightarrow E^{\text{TF}}$$

where

$$\mathcal{E}^{\text{TF}}(f) = K_3^{\text{cl}} \int_{\mathbb{R}^3} f^{5/3} - \int_{\mathbb{R}^3} \frac{f(x)}{|x|} dx + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{f(x)f(y)}{|x-y|} dx dy, \quad K_3^{\text{cl}} = \frac{3}{5}(6\pi^2)^{2/3}.$$

In this case,  $H_N^{\text{atom}}$  has a ground state and the unique minimizer of  $E^{\text{TF}}$  satisfies  $\int_{\mathbb{R}^d} f^{\text{TF}} = 1$  (we will come to that later).

*Proof. Energy upper bound.* Recall the variational principle

$$E_N = \inf_{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1} \langle \Psi_N, H_N \Psi_N \rangle = \inf_{\substack{f \geq 0 \\ \int_{\mathbb{R}^d} f = 1}} \inf_{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1, f_{\Psi_N}=f} \langle \Psi_N, H_N \Psi_N \rangle,$$

which can be rewritten as

$$\frac{E_N}{N} = \inf_{\substack{f \geq 0 \\ \int_{\mathbb{R}^d} f = 1}} \mathcal{E}_N(f)$$

thanks to the definition of the (rescaled) Levy–Lieb functional

$$\mathcal{E}_N(f) := \inf_{\substack{\|\Psi_N\|_{L_a^2(\mathbb{R}^{dN})}=1 \\ f_{\Psi_N}=f}} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N}$$

Recall also the following equivalent definition of the Thomas–Fermi energy (see an exercise above)

$$E^{\text{TF}} = \inf \left\{ \mathcal{E}^{\text{TF}}(f) : 0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d), \int_{\mathbb{R}^d} f = 1 \right\}.$$

For every  $0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  satisfying  $\int_{\mathbb{R}^d} f = 1$ , by the Gamma convergence (upper bound), we can find Slater determinants  $\Psi_N \in L_a^2(\mathbb{R}^{dN})$  such that  $f_{\Psi_N} = f_N \rightarrow f$  strongly in  $L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$ , and

$$\limsup_{N \rightarrow \infty} \frac{E_N}{N} \leq \limsup_{N \rightarrow \infty} \mathcal{E}_N(f_N) \leq \mathcal{E}^{\text{TF}}(f).$$

Then optimizing over  $f$  we obtain

$$\limsup_{N \rightarrow \infty} \frac{E_N}{N} \leq E^{\text{TF}}.$$

**Energy lower bound.** For any normalized wave function  $\Psi_N \in L_a^2(\mathbb{R}^{3N})$ , using  $w \geq 0$  and the Lieb–Thirring kinetic inequality we have

$$H_N \geq \sum_{i=1}^N (-N^{-2/d} \Delta_{x_i} + V(x_i)).$$

Hence,

$$\frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} \geq \frac{\langle \Psi_N, \sum_{i=1}^N (-\Delta_{x_i}) \Psi_N \rangle}{N^{1+2/d}} + \int_{\mathbb{R}^d} V f_{\Psi_N} \geq K_d \int_{\mathbb{R}^d} f_{\Psi_N}^{1+2/d} + \int_{\mathbb{R}^d} V f_{\Psi_N}$$

for a constant  $K_d > 0$ . Since  $V \in L^p(\mathbb{R}^d) + L^q(\mathbb{R}^d)$  with  $p, q \in [1 + d/2, \infty)$  we have

$$\frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} \geq \frac{K_d}{2} \int_{\mathbb{R}^d} f_{\Psi_N}^{1+2/d} - C.$$

Thus  $E_N/N$  is bounded from below. Moreover, if the wave function satisfies

$$\langle \Psi_N, H_N \Psi_N \rangle = E_N + o(N)_{N \rightarrow \infty},$$

then  $f_N := f_{\Psi_N}$  is bounded in  $L^{1+2/d}(\mathbb{R}^d)$ . Up to a subsequence, we can assume that  $f_N \rightharpoonup f$  in  $L^{1+2/d}(\mathbb{R}^d)$ . Hence, by the Gamma-convergence (lower bound) we have

$$\liminf_{N \rightarrow \infty} \frac{E_N}{N} = \liminf_{N \rightarrow \infty} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} \geq \liminf_{N \rightarrow \infty} \mathcal{E}_N(f_N) \geq \mathcal{E}^{\text{TF}}(f) \geq E^{\text{TF}}.$$

In the last inequality, we have used the variational definition of  $E^{\text{TF}}$ . Note that the weak convergence  $f_N \rightharpoonup f$  implies that  $0 \leq f \in L^1(\mathbb{R}^d) \cap L^{1+2/d}(\mathbb{R}^d)$  and

$$\int_{\mathbb{R}^d} f \leq \liminf_{N \rightarrow \infty} \int_{\mathbb{R}^d} f_N = 1.$$

Since the limit  $E^{\text{TF}}$  is unique, we can obtain the lower bound estimate for the whole sequence  $N \rightarrow \infty$ . Thus in conclusion we obtain the convergence of the ground state energy

$$\lim_{N \rightarrow \infty} \frac{E_N}{N} = E^{\text{TF}}.$$

**Convergence of ground states.** Let  $\Psi_N$  be an approximate ground state for  $H_N$  and let  $f_N = f_{\Psi_N}$ . Let  $f^{\text{TF}}$  be a minimizer for  $E^{\text{TF}}$  with  $\int_{\mathbb{R}^d} f^{\text{TF}} = 1$ . Then from the above proof of

the lower bound, we have  $f_N \rightharpoonup f$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$  and

$$\lim_{N \rightarrow \infty} \frac{E_N}{N} = \lim_{N \rightarrow \infty} \frac{\langle \Psi_N, H_N \Psi_N \rangle}{N} = \lim_{N \rightarrow \infty} \mathcal{E}_N(f_N) = \mathcal{E}^{\text{TF}}(f) = E^{\text{TF}} = \mathcal{E}^{\text{TF}}(f^{\text{TF}}).$$

Since the TF functional is strictly convex, it has a unique minimizer. Thus  $f_N \rightharpoonup f^{\text{TF}}$  weakly in  $L^{1+2/d}(\mathbb{R}^d)$ .  $\square$

## 8.5 Atomic Thomas–Fermi minimizer

We take a closer look at the Thomas–Fermi functional for Coulomb potentials

$$\mathcal{E}^{\text{TF}}(f) := \frac{3}{5}K \int_{\mathbb{R}^3} f^{5/3} - \int_{\mathbb{R}^3} \frac{f(x)}{|x|} dx + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{f(x)f(y)}{|x-y|} dx dy.$$

with a constant  $K > 0$  (the physical constant will be  $\frac{3}{5}K = K_3^{\text{cl}}$ , namely  $K = (6\pi^2)^{2/3}$ ). From an exercise of the previous section, we know that the minimization problem

$$E^{\text{TF}} := \inf \left\{ \mathcal{E}^{\text{TF}}(f) : 0 \leq f \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3), \int_{\mathbb{R}^3} f \leq 1 \right\}.$$

has a unique minimizer  $f^{\text{TF}}$  since  $\widehat{w}(k) = \text{const.}|k|^{-2} > 0$ .

**Theorem** (Atomic Thomas–Fermi minimizer). *The unique atomic TF minimizer  $f^{\text{TF}}$  is radially symmetric,  $\int_{\mathbb{R}^3} f^{\text{TF}} = 1$ , and it solves the TF equation*

$$K f^{\text{TF}}(x)^{2/3} = |x|^{-1} - f^{\text{TF}} * |x|^{-1}, \quad \forall x \in \mathbb{R}^3 \setminus \{0\}$$

Moreover,  $f^{\text{TF}}$  is the **unconstrained minimizer**, namely

$$\mathcal{E}(f^{\text{TF}}) \leq \mathcal{E}^{\text{TF}}(f), \quad \forall 0 \leq f \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3).$$

*Proof. Step 1.* Since  $f \mapsto \mathcal{E}^{\text{TF}}(f)$  is rotational invariant, the unique minimizer  $f^{\text{TF}}$  must be radially symmetric. Let us prove that  $\int_{\mathbb{R}^3} f^{\text{TF}} = 1$ . Assume by contradiction that  $\int_{\mathbb{R}^3} f^{\text{TF}} < 1$ . Then for every  $0 \leq \varphi \in C_c^\infty(\mathbb{R}^3)$  and  $t > 0$  small we have

$$f^{\text{TF}} + t\varphi \geq 0, \quad \int_{\mathbb{R}^3} (f^{\text{TF}} + t\varphi) \leq 1.$$



Thus by the minimality of  $f^{\text{TF}}$  we have

$$\mathcal{E}^{\text{TF}}(f^{\text{TF}}) \leq \mathcal{E}^{\text{TF}}(f^{\text{TF}} + t\varphi)$$

for all  $t \geq 0$  sufficiently small. Consequently,

$$0 \leq \frac{d}{dt} \left( \mathcal{E}^{\text{TF}}(f^{\text{TF}} + t\varphi) \right)_{|t=0^+} = \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) \varphi(x) dx.$$

Since this holds for all  $0 \leq \varphi \in C_c^\infty$ , we find that

$$K(f^{\text{TF}})^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \geq 0, \quad \text{for a.e. } x \in \mathbb{R}^3.$$

On the other hand, by Newton's theorem

$$f^{\text{TF}} * \frac{1}{|x|} = \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{|x-y|} dy = \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{\max\{|x|, |y|\}} dy \leq \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{|x|} dy = \frac{\int_{\mathbb{R}^3} f^{\text{TF}}}{|x|}.$$

Thus

$$K f^{\text{TF}}(x)^{2/3} \geq \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} \geq \left( 1 - \int_{\mathbb{R}^3} f^{\text{TF}} \right) \frac{1}{|x|} \quad \text{for a.e. } x \in \mathbb{R}^3,$$

which implies that

$$f^{\text{TF}}(x) \geq \frac{c_0}{|x|^{3/2}} \quad \text{for a.e. } x \in \mathbb{R}^3$$

with a constant  $c_0 > 0$ . However, the last inequality contradicts to the fact that  $\int_{\mathbb{R}^3} f^{\text{TF}} < \infty$ .

Thus we must have  $\int_{\mathbb{R}^3} f^{\text{TF}} = 1$ .

**Step 2.** Now we derive the TF equation. We can proceed similarly as above, but now we have to choose the test functions

$$\varphi \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3), \quad \varphi(x) \geq -f^{\text{TF}}(x) \text{ for a.e. } x \in \mathbb{R}^3, \quad \int_{\mathbb{R}^3} \varphi \leq 0.$$

Then as above, we have

$$\mathcal{E}(f^{\text{TF}}) \leq \mathcal{E}^{\text{TF}}(f^{\text{TF}} + t\varphi), \quad \forall t \in [0, 1],$$

and hence

$$0 \leq \frac{d}{dt} \left( \mathcal{E}^{\text{TF}}(f^{\text{TF}} + t\varphi) \right)_{|t=0^+} = \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) \varphi(x) dx.$$

By taking  $\varphi = -f^{\text{TF}}$ , we find that

$$\mu^{\text{TF}} := - \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) f^{\text{TF}}(x) \geq 0.$$

We can also take

$$\varphi(x) = g(x) - \left( \int_{\mathbb{R}^3} g \right) f^{\text{TF}}(x)$$

with

$$0 \leq g \in C_c^\infty(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} g \leq 1.$$

Then the above variational inequality reads

$$\begin{aligned} 0 &\leq \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) \left[ g(x) - \left( \int_{\mathbb{R}^3} g \right) f^{\text{TF}}(x) \right] dx \\ &= \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) g(x) dx + \mu^{\text{TF}} \int_{\mathbb{R}^3} g \\ &= \int_{\mathbb{R}^3} W(x) g(x) dx \end{aligned}$$

where

$$W(x) := K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} + \mu^{\text{TF}}.$$

Thus we have proved that

$$\int_{\mathbb{R}^3} W(x) g(x) dx \geq 0, \quad \forall 0 \leq g \in C_c^\infty(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} g \leq 1.$$

which implies that  $W(x) \geq 0$  for a.e.  $x \in \mathbb{R}^3$ . On the other hand, thanks to the definition of  $\mu^{\text{TF}}$  and the fact that  $\int_{\mathbb{R}^3} f^{\text{TF}} = 0$ , we obtain

$$\int_{\mathbb{R}^3} W f^{\text{TF}} = \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) f^{\text{TF}} + \mu^{\text{TF}} = 0.$$

Since  $W(x) \geq 0$  and  $f^{\text{TF}}(x) \geq 0$  for a.e.  $x \in \mathbb{R}^3$ , we conclude that when  $f^{\text{TF}}(x) > 0$ , we have  $W(x) = 0$  namely

$$K f^{\text{TF}}(x)^{2/3} = \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} - \mu^{\text{TF}}.$$

On the other hand, if  $f^{\text{TF}}(x) = 1$ , then using  $W(x) \geq 0$  we get

$$0 \geq \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} - \mu^{\text{TF}}.$$

These two formulas can be written in the compact form

$$Kf^{\text{TF}}(x)^{2/3} = \left[ \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} - \mu^{\text{TF}} \right]_+.$$

We will prove later that  $\mu^{\text{TF}} = 0$ , but that requires a preparation.

**Step 3.** Now we prove that  $f^{\text{TF}}$  is an unconstrained minimizer. For any  $m > 1$  (not necessarily an integer) we can consider the variational problem

$$E^{\text{TF}}(m) := \inf \left\{ \mathcal{E}^{\text{TF}}(f) : 0 \leq f \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3), \int_{\mathbb{R}^3} f \leq m \right\}.$$

Then proceeding exactly as before, we know that  $E^{\text{TF}}(m)$  has a unique minimizer  $f_m$  with  $\int_{\mathbb{R}^3} f_m \leq m$ , which is radially symmetric. Our goal is to show that  $f_m = f^{\text{TF}}$ . It suffices to prove

$$\int_{\mathbb{R}^3} f_m \leq 1.$$

We will need a TF equation for  $f_m$ . We can proceed similarly as in Step 2. More precisely, for all

$$\varphi \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3), \quad \varphi(x) \geq -f_m(x) \text{ for a.e. } x \in \mathbb{R}^3, \quad \int_{\mathbb{R}^3} \varphi \leq 0$$

we also have

$$0 \leq \frac{d}{dt} \left( \mathcal{E}^{\text{TF}}(f_m + t\varphi) \right)_{|t=0^+} = \int_{\mathbb{R}^3} \left( Kf_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} \right) \varphi(x) dx$$

and that implies

$$\mu_m := - \frac{\int_{\mathbb{R}^3} \left( Kf_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} \right) f_m(x)}{\int_{\mathbb{R}^3} f_m} \geq 0.$$

Now we take

$$\varphi(x) = g(x) - \left( \frac{\int_{\mathbb{R}^3} g}{\int_{\mathbb{R}^3} f_m} \right) f_m(x), \quad 0 \leq g \in C_c^\infty(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} g \leq 1$$

which gives

$$0 \leq \int_{\mathbb{R}^3} \left( Kf_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} \right) \left[ g(x) - \left( \frac{\int_{\mathbb{R}^3} g}{\int_{\mathbb{R}^3} f_m} \right) f_m(x) \right] dx$$

$$= \int_{\mathbb{R}^3} \left( K f_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} \right) g(x) dx + \mu_m \int_{\mathbb{R}^3} g = \int_{\mathbb{R}^3} W_m(x) g(x) dx$$

where

$$W_m(x) := K f_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} + \mu_m.$$

This implies that  $W_m(x) \geq 0$  for a.e.  $x \in \mathbb{R}^3$ . Moreover, by the choice of  $\mu_m$  we have

$$\int_{\mathbb{R}^3} W_m f_m = \int_{\mathbb{R}^3} \left( K f_m(x)^{2/3} - \frac{1}{|x|} + f_m * \frac{1}{|x|} \right) f_m + \mu_m \int_{\mathbb{R}^3} f_m = 0.$$

Thus  $W_m(x) = 0$  if  $f_m(x) > 0$ . All this gives the TF equation

$$K f_m(x)^{2/3} = \left[ \frac{1}{|x|} - f_m * \frac{1}{|x|} - \mu_m \right]_+.$$

It remains to deduce the bound  $\int_{\mathbb{R}^3} f_m \leq 1$  from the TF equation. Recall that  $\mu_m \geq 0$  and  $f_m$  is radial (hence, we can use Newton's theorem to simplify the convolution). Multiplying the TF equation with  $f_m(x)$  we have the pointwise inequality

$$\begin{aligned} 0 &\leq K f_m(x)^{5/3} = \left[ \frac{1}{|x|} - f_m * \frac{1}{|x|} - \mu_m \right] f_m(x) \\ &\leq \left[ \frac{1}{|x|} - f_m * \frac{1}{|x|} \right] f_m(x) \\ &= \left[ \frac{1}{|x|} - \int_{\mathbb{R}^3} \frac{f_m(y)}{\max\{|x|, |y|\}} dy \right] f_m(x), \quad \text{for a.e. } x \in \mathbb{R}^3. \end{aligned}$$

Integrating against  $\mathbb{1}(|x| \leq R) |x|^k dx$  with  $k = 2, 3, \dots$  we obtain

$$\begin{aligned} \int_{|x| \leq R} |x|^{k-1} f_m(x) dx &\geq \int_{|x| \leq R} \int_{|y| \leq R} \frac{|x|^k}{\max\{|x|, |y|\}} f_m(x) f_m(y) dx dy \\ &= \frac{1}{2} \int_{|x| \leq R} \int_{|y| \leq R} \frac{|x|^k + |y|^k}{\max\{|x|, |y|\}} f_m(x) f_m(y) dx dy. \end{aligned}$$

Note that any  $a \geq b \geq 0$  and  $k = 2, 3, \dots$  we have the elementary inequality

$$\frac{a^k + b^k}{\max\{a, b\}} = a^{k-1} + \frac{b^k}{a} \geq \frac{k-1}{k} (a^{k-1} + b^{k-1})$$

which is equivalent to

$$a^{k-1} + k \frac{b^k}{a} \geq (k-1) b^{k-1}.$$

The latter follows from the AM-GM inequality

$$a^{k-1} + (k-1)\frac{b^k}{a} \geq kb^{k-1}.$$

Coming back to the TF problem, we have

$$\begin{aligned} \int_{|x| \leq R} |x|^{k-1} f_m(x) dx &\geq \frac{1}{2} \int_{|x| \leq R} \int_{|y| \leq R} \frac{|x|^k + |y|^k}{\max\{|x|, |y|\}} f_m(x) f_m(y) dx dy \\ &\geq \frac{k-1}{2k} \int_{|x| \leq R} \int_{|y| \leq R} (|x|^{k-1} + |y|^{k-1}) f_m(x) f_m(y) dx dy \\ &= \frac{k-1}{k} \left( \int_{|x| \leq R} |x|^{k-1} f_m(x) dx \right) \left( \int_{|y| \leq R} f_m(y) dy \right). \end{aligned}$$

Thus for  $R > 0$  large we find that  $\int_{|x| \leq R} |x|^{k-1} f_m(x) dx > 0$ , and hence

$$1 \geq \frac{k-1}{k} \int_{|y| \leq R} f_m(y) dy.$$

The taking  $k \rightarrow \infty$  and  $R \rightarrow \infty$  we conclude that

$$1 \geq \int_{\mathbb{R}^3} f_m.$$

Thus  $f_m = f^{\text{TF}}$ , for all  $m > 1$ . This means that  $f^{\text{TF}}$  is an unconstrained minimizer for the TF functional.

**Step 4.** Since  $f^{\text{TF}}$  is an unconstrained minimizer for the TF functional, we have

$$\mathcal{E}(f^{\text{TF}}) \leq \mathcal{E}(t f^{\text{TF}}), \quad \forall t > 0.$$

Consequently,

$$0 = \frac{d}{dt} \mathcal{E}(t f^{\text{TF}})_{t=1} = \int_{\mathbb{R}^3} \left( K f^{\text{TF}}(x)^{2/3} - \frac{1}{|x|} + f^{\text{TF}} * \frac{1}{|x|} \right) f^{\text{TF}}(x) = -\mu^{\text{TF}}.$$

Thus the TF equation becomes

$$K f^{\text{TF}}(x)^{2/3} = \left[ \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} \right]_+ = \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|}.$$

Here in the second identity we have used Newton's theorem

$$f^{\text{TF}} * \frac{1}{|x|} = \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{|x-y|} dy = \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{\max\{|x|, |y|\}} dy \leq \int_{\mathbb{R}^3} \frac{f^{\text{TF}}(y)}{|x|} dy = \frac{1}{|x|}.$$

From the TF equation, we find that  $f^{\text{TF}}(x)$  is continuous away from 0, since both  $|x|^{-1}$  and the convolution  $f^{\text{TF}} * \frac{1}{|x|}$  are continuous away from 0. Hence, the TF equation holds true for all  $x \neq 0$ . This completes the proof of the theorem.  $\square$

Remark: The function

$$\Phi^{\text{TF}}(x) = \frac{1}{|x|} - f^{\text{TF}} * \frac{1}{|x|} \geq 0$$

is called the **Thomass–Fermi potential**. Note that

$$\Delta \Phi^{\text{TF}}(x) = 4\pi f^{\text{TF}}(x), \quad \forall x \in \mathbb{R}^3 \setminus \{0\}.$$

Hence, the TF equation can be written in the equivalent form

$$\Delta \Phi^{\text{TF}}(x) = 4\pi K^{-3/2} \Phi^{\text{TF}}(x)^{3/2}.$$

This nonlinear PDE is well-studied. We have the **Sommerfeld asymptotic formula**

$$\lim_{|x| \rightarrow \infty} |x|^4 \Phi^{\text{TF}}(x) = A^{\text{TF}}, \quad A^{\text{TF}} = \frac{(3K)^3}{3\pi^2}.$$

In fact, the following pointwise bound was proved by **Solovej** (2000).

$$A^{\text{TF}} |x|^{-4} \geq \Phi^{\text{TF}}(x) \geq A^{\text{TF}} |x|^{-4} - C |x|^{-4-\varepsilon}, \quad \varepsilon = \frac{\sqrt{73}-7}{2} > 0.$$

Consequently, the TF minimizer  $f^{\text{TF}}(x)$  decays as  $\text{const.}|x|^{-6}$  when  $|x| \rightarrow \infty$ .

# Chapter 9

## Hartree–Fock theory

Recall that the ground state energy of a Hamiltonian  $H_N$  on  $L_a^2(\mathbb{R}^{dN})$  is defined by

$$E_N = \inf_{\|\Psi\|_{L_a^2(\mathbb{R}^{dN})}=1} \langle \Psi, H_N \Psi \rangle.$$

In the Hartree–Fock theory, one restricts the consideration to **Slater determinants** and consider

$$E^{\text{HF}} = \inf_{\Psi \text{ a Slater determinant}} \langle \Psi, H_N \Psi \rangle.$$

Thus by the variational principle, we have the obvious **upper bound**

$$E_N \leq E^{\text{HF}}.$$

For a comparison, the Thomas–Fermi energy is neither an upper bound nor a lower bound to the full quantum energy. As we will see, in many situations, it is possible to obtain a good lower bound for  $E_N - E^{\text{HF}}$ , making the Hartree–Fock theory significantly more precise than the Thomas–Fermi theory.

In the same spirit of the density functional theory, an important advantage of the Hartree–Fock theory is that the energy expectation can be expressed purely in terms of the one–body density matrix of the Slater determinants. For a typical Hamiltonian  $H_N$  on  $L_a^2(\mathbb{R}^{dN})$  of the form

$$H_N = \sum_{i=1}^N \left( -h^2 \Delta_{x_i} + V(x_i) \right) + \lambda \sum_{1 \leq i < j \leq N} w(x_i - x_j)$$

and for a Slater determinant

$$\Psi_N = (u_1 \wedge u_2 \wedge \dots \wedge u_N)(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \text{sign}(\sigma) u_1(x_{\sigma(1)}) \dots u_N(x_{\sigma(N)}),$$

where  $\{u_i\}_{i=1}^N$  are orthonormal functions in  $L^2(\mathbb{R}^d)$ , we have

$$\langle \Psi_N, H_N \Psi_N \rangle = \mathcal{E}^{\text{HF}}(\gamma_{\Psi_N}^{(1)}), \quad \gamma_{\Psi_N}^{(1)} = \sum_{i=1}^N |u_i\rangle\langle u_i|$$

where

$$\mathcal{E}^{\text{HF}}(\gamma) := \text{Tr}((-h^2\Delta + V)\gamma) + \frac{\lambda}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x, y)|^2] w(x - y) dx dy.$$

The two terms in the interaction are called **direct** and **exchange** energy. Note that we always have

$$\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x, y)|^2 \geq 0, \quad \forall x, y \in \mathbb{R}^d.$$

Thus the Hartree–Fock energy can be rewritten as

$$E^{\text{HF}} = \inf_{\substack{0 \leq \gamma = \gamma^2 \leq 1 \\ \text{Tr } \gamma = N}} \mathcal{E}^{\text{HF}}(\gamma)$$

Here the condition  $\gamma = \gamma^2$  is to ensure that  $\gamma$  is a projection. For some computation, it is more convenient to ignore this condition since the set  $\{0 \leq \gamma \leq 1, \text{Tr } \gamma = N\}$  is convex. Actually, it is possible to do that without losing anything, provided that the interaction potential is non-negative.

**Theorem** (Lieb’s variational principle). *If  $w \geq 0$ , then*

$$E^{\text{HF}} = \inf_{\substack{0 \leq \gamma \leq 1 \\ \text{Tr } \gamma = N}} \mathcal{E}^{\text{HF}}(\gamma).$$

*Here  $V$  is assumed to be “regular enough” such that  $-h^2\Delta + V$  is bounded from below.*

This result was first proved by **Lieb** (1981). In the following we represent a simplified proof of **Bach** (1994).

*Proof.* By a density argument, it suffices to take the infimum on the right side only on finite-rank operators. We will prove that if  $0 \leq \gamma \leq 1$ ,  $\text{Tr } \gamma = N$  and  $\mathcal{E}^{\text{HF}}(\gamma) < \infty$ , then there



exists a projection  $0 \leq \tilde{\gamma} = \tilde{\gamma}^2 \leq 1$ ,  $\text{Tr } \tilde{\gamma} = N$  such that

$$\mathcal{E}^{\text{HF}}(\gamma) \geq \mathcal{E}^{\text{HF}}(\tilde{\gamma}).$$

To construct  $\tilde{\gamma}$ , let us write

$$\gamma = \sum_{i=1}^M \lambda_i |u_i\rangle\langle u_i|, \quad 0 < \lambda_i \leq 1, \quad \sum_{i=1}^M \lambda_i = N, \quad \{u_i\}_{i=1}^M \text{ orthonormal functions in } L^2(\mathbb{R}^d).$$

Then we have

$$\mathcal{E}^{\text{HF}}(\gamma) = \sum_{i=1}^M \lambda_i A_i + \frac{1}{2} \sum_{i,j=1}^M \lambda_i \lambda_j B_{ij}$$

where

$$A_i = \langle u_i, (-h^2 \Delta + V) u_i \rangle, \quad B_{ij} = \Re \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left( |u_i(x)|^2 |u_j(y)|^2 - \overline{u_i(x) u_j(y)} u_j(x) u_i(y) \right) w(x-y) \geq 0.$$

Note that  $B_{ij} = B_{ji}$  and  $B_{ii} = 0$ . Define

$$n(\gamma) := |\{\lambda_i : 0 < \lambda_i < 1\}|.$$

We assume that  $n(\gamma) \geq 2$ ; otherwise  $\gamma$  is already a projection. Then there are at least two eigenvalues  $\lambda_k, \lambda_\ell$  in  $(0, 1)$  and we can assume that

$$A_k + \sum_{i=1}^M \lambda_i B_{ik} \geq A_\ell + \sum_{i=1}^M \lambda_i B_{i\ell}.$$

Let  $\delta = \min\{\lambda_k, 1 - \lambda_\ell\} > 0$ , then either  $\lambda_k - \delta = 0$  or  $\lambda_\ell + \delta = 1$ . Define

$$\tilde{\gamma} = \left( \sum_{i \neq k, \ell} \lambda_i |u_i\rangle\langle u_i| \right) + (\lambda_k - \delta) |u_k\rangle\langle u_k| + (\lambda_\ell + \delta) |u_\ell\rangle\langle u_\ell|.$$

We have

$$n(\tilde{\gamma}) \leq n(\gamma) - 1$$

and

$$\mathcal{E}^{\text{HF}}(\gamma) - \mathcal{E}^{\text{HF}}(\tilde{\gamma}) = \lambda_k A_k + \lambda_\ell A_\ell + \sum_{i \neq k, \ell} \lambda_i \lambda_k B_{ik} + \sum_{i \neq k, \ell} \lambda_i \lambda_\ell B_{i\ell} + \lambda_k \lambda_\ell B_{k\ell}$$

$$\begin{aligned}
& - \left( (\lambda_k - \delta)A_k + (\lambda_\ell + \delta)A_\ell + \sum_{i \neq k, \ell}^M \lambda_i(\lambda_k - \delta)B_{ik} + \sum_{i \neq k, \ell}^M \lambda_i(\lambda_\ell + \delta)B_{i\ell} + (\lambda_k - \delta)(\lambda_\ell + \delta)B_{k\ell} \right) \\
& = \delta \left( A_k - A_\ell + \sum_{i=1}^M \lambda_i B_{ik} - \sum_{i=1}^M \lambda_i B_{i\ell} \right) + \delta^2 B_{k,\ell} \geq 0.
\end{aligned}$$

Thus we can replace  $\gamma$  by  $\tilde{\gamma}$  without increasing the energy. Here  $\tilde{\gamma}$  may be not yet a projection, but since  $n(\tilde{\gamma}) \leq n(\gamma) - 1$ , we can iterate the procedure and eventually obtain a projection after finitely many steps.  $\square$

In some situation, people are also interested in the **reduced Hartree–Fock energy**

$$E^{\text{rHF}} = \inf_{\substack{0 \leq \gamma \leq 1 \\ \text{Tr } \gamma = N}} \mathcal{E}^{\text{rHF}}(\gamma)$$

where

$$\mathcal{E}^{\text{rHF}}(\gamma) := \text{Tr}((-h^2\Delta + V)\gamma) + \frac{\lambda}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_\gamma(x)\rho_\gamma(y)w(x-y)dx dy.$$

Thus here we keep only the direct interaction energy and ignore the exchange energy. Mathematically, the reduced Hartree–Fock theory is easier to analyze since, for example, if  $\hat{w} \geq 0$  then  $\gamma \mapsto \mathcal{E}^{\text{rHF}}(\gamma)$  is convex. If  $w \geq 0$ , then the exchange energy is non-negative, and hence

$$E_N \leq E^{\text{HF}} \leq E^{\text{rHF}}.$$

In this chapter, we will see how good the Hartree–Fock and reduced Hartree–Fock approximations are. The main question is the lower bound for  $E_N$ . We will focus on the atomic case where

$$H_N = \sum_{i=1}^N \left( -N^{-2/3}\Delta_{x_i} - \frac{1}{|x_i|} \right) + N^{-1} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \quad \text{on } L_a^2(\mathbb{R}^{3N}).$$

A big achievement in mathematical physics during 1970-1990 is the following rigorous formula for the full quantum ground state energy

$$E_N = -c_1 N + c_2 N^{2/3} - c_3 N^{1/3} + o(N^{1/3})_{N \rightarrow \infty}.$$

Here the first term is determined by the Thomas–Fermi theory, as we discussed in the previous chapter. This was first proved by **Lieb** and **Simon** in 1970s. The second term is called Scott’s correction, corresponding for the electrons moving very close to the nucleus (at a distance

$N^{-2/3}$  which is much shorter than the semiclassical distance  $O(1)$ ). This was proved partially by **Hughes** (lower bound), and fully by **Siedentop** and **Weikard** in 1987. The third term is called the Dirac-Schwinger correction, corresponding to further and subtle correction in semiclassical approximation. This was proved by **Fefferman** and **Seco** in 1990s.

In this chapter, we will prove rigorously that

$$E_N = E^{\text{rHF}} + O(N^{1/3})_{N \rightarrow \infty}, \quad E_N = E^{\text{HF}} + o(N^{1/3})_{N \rightarrow \infty}.$$

## 9.1 Lieb–Oxford inequality

In the previous chapter, we have proved that if  $f_{\Psi_N} = N^{-1}\rho_{\Psi_N} \rightarrow f$ , then

$$\liminf_{N \rightarrow \infty} N^{-2} \left\langle \Psi_N, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi_N \right\rangle \geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f(x)f(y)}{|x - y|} dx dy.$$

A quantitative version of this mean–field approximation is

**Theorem** (Lieb–Oxford inequality). *For every normalized wave function  $\Psi \in L_a^2(\mathbb{R}^{3N})$  we have*

$$\left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle \geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_{\Psi}(x)\rho_{\Psi}(y)}{|x - y|} dx dy - C_{\text{LO}} \int_{\mathbb{R}^3} \rho_{\Psi}(x)^{4/3} dx.$$

We can choose  $C_{\text{LO}} = 1.68$ .

This result was first proved by **Lieb** (1979) with  $C_{\text{LO}} = 8.52$ . The constant  $C_{\text{LO}} = 1.68$  was obtained later by **Lieb** and Oxford (1981). Currently, the best known result is

$$1.64 \geq C_{\text{LO}} \geq 1.44.$$

In the following we will prove the Lieb–Oxford inequality with a worse constant, using the strategy of **Lieb**, **Solovej** and **Yngvason** based on the **Fefferman-de la Llave formula**

$$\frac{1}{|x|} = \frac{1}{\pi} \int_0^\infty (\mathbf{1}_{B_r} * \mathbf{1}_{B_r})(x) \frac{dr}{r^5}, \quad \forall x \in \mathbb{R}^3 \setminus \{0\}$$

and the **Hardy–Littlewood maximal function**.

**Definition** (Maximal function). For every  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ , the maximal function  $M_f : \mathbb{R}^d \rightarrow [0, \infty]$  is defined by

$$M_f(x) = \sup_{r>0} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y)| dy.$$

Note that

$$|f(x)| = \lim_{r \rightarrow 0} \frac{1}{|B(x, r)|} \int_{B(x, r)} |f(y)| dy \leq M_f(x), \quad \text{a.e. } x \in \mathbb{R}^d.$$

On the other hand, the following result is very helpful.

**Theorem** (Strong-type estimate). For every  $1 < p \leq \infty$ , we have

$$\|M_f\|_{L^p(\mathbb{R}^d)} \leq C_p \|f\|_{L^p(\mathbb{R}^d)}$$

with a constant  $C_p \in (0, \infty)$  independent of  $f$ .

The original Hardy–Littlewood maximal inequality gives a constant  $C_{d,p}$ . The fact that  $C_p$  can be chosen independently of the dimension is due to **Stein**. Note that the condition  $p > 1$  is crucial. For  $p = 1$ , we only have a **weak-type estimate** (see the proof below).

Let us postpone the proof of the strong-type estimate and provide

*Proof of the Lieb–Oxford inequality.* Let  $\Psi$  be a normalized wave function in  $L^2_a(\mathbb{R}^{3N})$  and denote  $\rho = \rho_\Psi$ . Recall that by the Fefferman-de la Llave representation

$$\frac{1}{|x - y|} = \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \mathbf{1}_{B_r}(x - z) \mathbf{1}_{B_r}(y - z),$$

we can write

$$\left\langle \Psi, \sum_{1 \leq i < j \leq N} w(x_i - x_j) \Psi \right\rangle = \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \left\langle \Psi, \sum_{1 \leq i < j \leq N} \mathbf{1}_{B_r}(x_i - z) \mathbf{1}_{B_r}(x_j - z) \Psi \right\rangle.$$

For every  $r > 0$  and  $z \in \mathbb{R}^3$ , by the Cauchy–Schwarz inequality we get

$$\begin{aligned} & \left\langle \Psi, \sum_{1 \leq i < j \leq N} \mathbf{1}_{B_r}(x_i - z) \mathbf{1}_{B_r}(x_j - z) \Psi \right\rangle \\ &= \frac{1}{2} \left[ \left\langle \Psi, \left( \sum_{i=1}^N \mathbf{1}_{B_r}(x_i - z) \right)^2 \Psi \right\rangle - \left\langle \Psi, \sum_{i=1}^N \mathbf{1}_{B_r}^2(x_i - z) \Psi \right\rangle \right]_+ \end{aligned}$$

$$\begin{aligned}
 &\geq \frac{1}{2} \left[ \left\langle \Psi, \sum_{i=1}^N \mathbf{1}_{B_r}(x_i - z) \Psi \right\rangle^2 - \left\langle \Psi, \sum_{i=1}^N \mathbf{1}_{B_r}(x_i - z) \Psi \right\rangle \right]_+ \\
 &= \frac{1}{2} [(\rho * \mathbf{1}_{B_r})^2(z) - (\rho * \mathbf{1}_{B_r})(z)]_+ \\
 &= \frac{1}{2} (\rho * \mathbf{1}_{B_r})^2(z) - \frac{1}{2} \min \left\{ (\rho * \mathbf{1}_{B_r})(z), (\rho * \mathbf{1}_{B_r})^2(z) \right\}.
 \end{aligned}$$

Integrating over  $z$  and  $dr/r^5$  we obtain

$$\begin{aligned}
 \left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle &\geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x - y|} dx dy \\
 &\quad - C \int_{\mathbb{R}^3} dz \int_0^\infty \frac{dr}{r^5} \min \left\{ (\rho * \mathbf{1}_{B_r})(z), (\rho * \mathbf{1}_{B_r})^2(z) \right\}.
 \end{aligned}$$

It remains to bound the error term. By the definition of the maximal function, we have

$$(\rho * \mathbf{1}_{B_r})(z) = \int_{B(z,r)} \rho(y) dy \leq |B(z,r)| M_\rho(z) = Cr^3 M_\rho(z).$$

Hence, for every  $R > 0$  we can bound

$$\begin{aligned}
 \int_0^\infty \frac{dr}{r^5} \min \left\{ (\rho * \mathbf{1}_{B_r})(z), (\rho * \mathbf{1}_{B_r})^2(z) \right\} &\leq \int_0^R \frac{dr}{r^5} (\rho * \mathbf{1}_{B_r})^2(z) + \int_R^\infty \frac{dr}{r^5} (\rho * \mathbf{1}_{B_r})(z) \\
 &\leq C \int_0^R \frac{dr}{r^5} r^6 M_\rho(z)^2 + \int_R^\infty \frac{dr}{r^5} r^3 M_\rho(z) \\
 &\leq C(R^2 M_\rho(z)^2 + R^{-1} M_\rho(z)).
 \end{aligned}$$

We can optimize the right side over  $R > 0$ . It is easy to see that the optimizer  $R$  satisfies

$$R^2 M_\rho(z)^2 \sim R^{-1} M_\rho(z) \sim \left( R^2 M_\rho(z)^2 (R^{-1} M_\rho(z))^2 \right)^{1/3} = M_\rho(z)^{4/3}.$$

Thus we obtain

$$\int_0^\infty \frac{dr}{r^5} \min \left\{ (\rho * \mathbf{1}_{B_r})(z), (\rho * \mathbf{1}_{B_r})^2(z) \right\} \leq CM_\rho(z)^{4/3}.$$

Finally, by the strong-type estimate we conclude that

$$\int_{\mathbb{R}^3} dz \int_0^\infty \frac{dr}{r^5} \min \left\{ (\rho * \mathbf{1}_{B_r})(z), (\rho * \mathbf{1}_{B_r})^2(z) \right\} \leq C \int_{\mathbb{R}^3} dz M_\rho(z)^{4/3} \leq C \int_{\mathbb{R}^3} dz \rho(z)^{4/3}.$$

This completes the proof of the Lieb–Oxford inequality (with a worse constant).  $\square$

We end this section by quickly recalling the proof of the strong-type estimate, which is one of the most useful results from harmonic analysis. We will need the following geometric result.

**Lemma** (Vitali covering lemma). *Let  $\{B_j\}_J$  be a family of balls in  $\mathbb{R}^d$  such that*

$$\sup_{j \in J} \text{diam}(B_j) < \infty.$$

*Then there exists a subfamily of **disjoint balls**  $\{B_j\}_{J'}$  such that*

$$\bigcup_{j \in J} B_j \subset \bigcup_{j \in J'} 5B_j.$$

*Here if  $B_j = B(x_j, r_j)$ , then  $5B_j = B(x_j, 5r_j)$ .*

Remarks:

- The set  $J$  can be finite, countable or uncountable. The subset  $J'$  is always at most countable since the balls  $\{B_j\}_{J'}$  are disjoint.
- The condition  $\sup_{j \in J} \text{diam}(B_j) < \infty$  is crucial. Without it, a counter example is  $B_j = B(0, j)$  with  $j = 1, 2, \dots$
- The constant 5 is not optimal. It can be replaced by  $3 + \varepsilon$  (and 3 if  $J$  is finite).

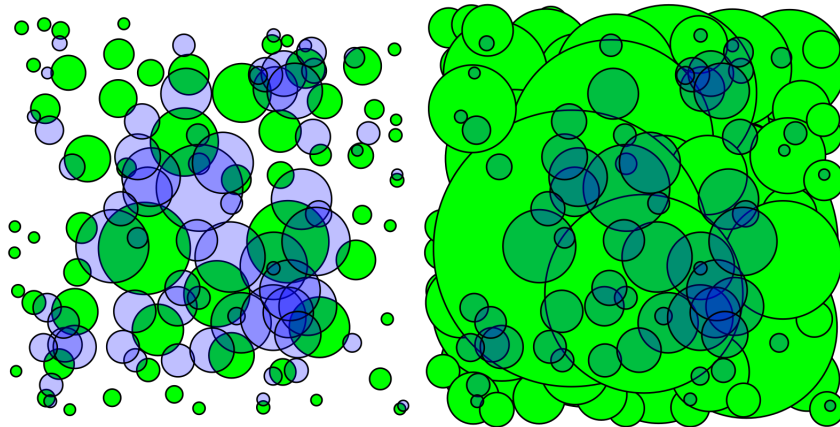


Figure from Wikipedia: The balls  $\{B_j\}_{j \in J}$  (left) and  $\{3B_j\}_{j \in J'}$  (right)

*Proof.* Here let us consider the simple version when  $J$  is finite (the infinite case is harder and left as an exercise). We choose  $J'$  by induction.

- First, we take a ball  $B_{j_1}$  of largest radius and put  $j_1 \in J'$ .
- Second, we ignore all balls with intersecting with  $B_{j_1}$ . If there is nothing left, then we stop. Otherwise, among all balls disjoint with  $B_{j_1}$ , we take a ball  $B_{j_2}$  of largest radius and put  $j_2 \in J'$ .
- Assume that we have chosen  $j_1, \dots, j_k \in J'$ . If every ball intersect with  $B_{j_1} \cup \dots \cup B_{j_k}$ , then we stop. Otherwise, among all balls disjoint with  $B_{j_1} \cup \dots \cup B_{j_k}$ , we take a ball  $B_{j_{k+1}}$  of largest radius and put  $j_{k+1} \in J'$ .

This procedure must stop after finitely many times. Then the resulting balls  $\{B_j\}_{J'} = \{B_{j_1}, \dots, B_{j_M}\}$  are clearly disjoint. Moreover, any ball  $B_i$  with  $i \in J$ , must intersect with a ball  $B_j$  with  $j \in J'$  such that the radius of  $B_i$  is  $\leq$  the radius of  $B_j$ . Then by the triangle inequality,

$$B_i \subset 3B_j \subset \bigcup_{j \in J'} 3B_j.$$

Consequently,

$$\bigcup_{i \in J} B_i \subset \bigcup_{j \in J'} 3B_j.$$

□

Now we are ready to provide

*Proof of the Hardy–Littlewood maximal inequality. Step 1.* We prove the **weak-type estimate**

$$\sup_{\lambda > 0} \lambda |\{M_f > \lambda\}| \leq C_d \|f\|_{L^1(\mathbb{R}^d)}, \quad \forall f \in L^1(\mathbb{R}^d).$$

Assume  $M_f(x) > \lambda$  for some  $x \in \mathbb{R}^d$ . Then by the definition of  $M_f(x)$ , we can find a ball  $B(x, r_x)$  such that

$$\frac{1}{|B(x, r_x)|} \int_{B(x, r_x)} |f(y)| dy > \lambda \iff \frac{1}{\lambda} \int_{B(x, r_x)} |f(y)| dy > |B(x, r_x)|.$$

By the Vitali covering lemma, from the collection  $\{B(x, r_x) : x \in J\}$  with  $J = \{x : M_f(x) > \lambda\}$  we can find a sub-collection of disjoint balls  $\{B(x, r_x) : x \in J'\}$  such that

$$\bigcup_{x \in J} B(x, r_x) \subset \bigcup_{x \in J'} 5B(x, r_x).$$

Consequently,

$$\begin{aligned} |\{M_f > \lambda\}| &\leq \left| \bigcup_{x \in J} B(x, r_x) \right| \leq \left| \bigcup_{x \in J'} 5B(x, r_x) \right| \leq \sum_{x \in J'} |5B(x, r_x)| \\ &= 5^d \sum_{x \in J'} |B(x, r_x)| \leq \frac{5^d}{\lambda} \sum_{x \in J'} \int_{B(x, r_x)} |f(y)| dy \leq \frac{5^d}{\lambda} \int_{\mathbb{R}^d} |f(y)| dy. \end{aligned}$$

This completes the proof of the weak-type estimate.

**Step 2.** Now we prove the strong-type estimate. For every  $f \in L^p(\mathbb{R}^d)$  we use the layer-cake representation

$$\int_{\mathbb{R}^d} M_f(x)^p dx = \int_0^\infty p \lambda^{p-1} |\{M_f > \lambda\}| d\lambda.$$

If we simply insert the weak-type estimate

$$|\{M_f > \lambda\}| \leq \frac{C_d}{\lambda} \|f\|_{L^1}$$

in the layer-cake representation, then we get  $\infty$  since  $\lambda^{p-2}$  is not integrable. However, we can split

$$|f| = |f| \mathbf{1}(|f| > \lambda/2) + |f| \mathbf{1}(|f| \leq \lambda/2) \leq g + \lambda/2, \quad g = |f| \mathbf{1}(|f| > \lambda/2)$$

which implies that

$$M_f \leq M_g + \lambda/2.$$

Therefore,

$$|\{M_f > \lambda\}| \leq |\{M_g > \lambda/2\}| \leq \frac{C_d}{\lambda/2} \|g\|_{L^1(\mathbb{R}^d)} = \frac{2C_d}{\lambda} \int_{\mathbb{R}^d} |f(y)| \mathbf{1}(|f(y)| > \lambda/2) dy$$

where we have applied the weak-type estimate for  $g$ . Inserting the latter bound in the layer-cake representation and using Fubini's theorem, we conclude that

$$\begin{aligned} \int_{\mathbb{R}^d} M_f(x)^p dx &\leq C_{d,p} \int_0^\infty d\lambda \lambda^{p-2} \int_{\mathbb{R}^d} dy |f(y)| \mathbf{1}(|f(y)| > \lambda/2) \\ &= C_{d,p} \int_{\mathbb{R}^d} dy |f(y)| \int_0^\infty d\lambda \lambda^{p-2} \mathbf{1}(|f(y)| > \lambda/2) \leq C_{d,p} \int_{\mathbb{R}^d} dy |f(y)|^p. \end{aligned}$$

This completes the proof of the strong-type estimate.  $\square$



## 9.2 Error bound for atomic reduced Hartree–Fock energy

Now we come back to the ground state energy  $E_N$  of the atomic Hamiltonian

$$H_N = \sum_{i=1}^N \left( -N^{-2/3} \Delta_{x_i} - \frac{1}{|x_i|} \right) + N^{-1} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \quad \text{on } L_a^2(\mathbb{R}^{3N}).$$

We will compare with the reduced Hartree–Fock energy

$$E^{\text{rHF}} = \inf_{\substack{0 \leq \gamma \leq 1 \\ \text{Tr } \gamma = N}} \mathcal{E}^{\text{rHF}}(\gamma)$$

where

$$\mathcal{E}^{\text{rHF}}(\gamma) := \text{Tr}((-N^{-2/3} \Delta - |x|^{-1})\gamma) + \frac{1}{2N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y)}{|x-y|} dx dy.$$

Recall that

$$E_N \leq E^{\text{HF}} \leq E^{\text{rHF}}.$$

**Theorem.** *We have  $E^{\text{rHF}} = E_N + O(N^{1/3})_{N \rightarrow \infty}$ .*

*Proof.* We know that  $E_N \leq E^{\text{rHF}}$ , and it suffices to consider the lower bound of  $E_N - E^{\text{rHF}}$ . Let a normalized wave function  $\Psi_N \in L_a^s(\mathbb{R}^{3N})$  satisfy

$$E_N \geq \langle \Psi_N, H_N \Psi_N \rangle + O(N^{-1}).$$

By the Lieb–Oxford inequality

$$\begin{aligned} \langle \Psi_N, H_N \Psi_N \rangle &\geq \text{Tr}((-N^{-2/3} \Delta + V)\gamma_{\Psi_N}^{(1)}) + \frac{1}{2N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_{\Psi_N}(x)\rho_{\Psi_N}(y)}{|x-y|} dx dy - CN^{-1} \int_{\mathbb{R}^3} \rho_{\Psi_N}^{4/3} \\ &= \mathcal{E}^{\text{rHF}}(\gamma_{\Psi_N}^{(1)}) - CN^{1/3} \int_{\mathbb{R}^3} f_{\Psi_N}^{4/3}. \end{aligned}$$

As in the proof of the validity of Thomas–Fermi theory, we know that

$$\langle \Psi_N, H_N \Psi_N \rangle \geq N \left( K \int_{\mathbb{R}^3} f_{\Psi_N}(x)^{5/3} dx - \int_{\mathbb{R}^3} \frac{f_{\Psi_N}(x)}{|x|} dx \right) \geq N \left( \frac{K}{2} \int_{\mathbb{R}^3} f_{\Psi_N}(x)^{5/3} dx - C \right)$$

for a constant  $K > 0$ . Since  $\langle \Psi_N, H_N \Psi_N \rangle \leq E_N + O(N^{-1}) \leq CN$ , we find that

$$\int_{\mathbb{R}^3} f_{\Psi_N}(x)^{5/3} dx \leq C.$$

We also have  $\int_{\mathbb{R}^3} f_{\Psi_N} = 1$ . Hence by Hölder's inequality,

$$\int_{\mathbb{R}^3} f_{\Psi_N}^{4/3} \leq C.$$

In conclusion, we have

$$\langle \Psi_N, H_N \Psi_N \rangle \geq \mathcal{E}^{\text{rHF}}(\gamma_{\Psi_N}^{(1)}) - CN^{1/3} \int_{\mathbb{R}^3} f_{\Psi_N}^{4/3} \geq E^{\text{rHF}} - CN^{1/3}.$$

This completes the proof of the desired lower bound for  $E_N$ . Thus in summary,

$$E^{\text{rHF}} \geq E_N \geq E^{\text{rHF}} - CN^{1/3}.$$

□

Next, we compare the reduced Hartree–Fock energy  $E^{\text{rHF}}$  with the Thomas–Fermi energy

$$E^{\text{TF}} := \inf_{0 \leq f \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3)} \left( K_3^{\text{cl}} \int_{\mathbb{R}^3} f^{5/3} - \int_{\mathbb{R}^3} \frac{f(x)}{|x|} dx + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{f(x)f(y)}{|x-y|} dx dy \right).$$

Recall that  $E^{\text{TF}}$  has a unique minimizer  $f^{\text{TF}}$  which satisfies  $\int_{\mathbb{R}^3} f^{\text{TF}} = 1$  and the TF equation

$$\frac{5}{3} K_3^{\text{cl}} (f^{\text{TF}})^{2/3} = \Phi^{\text{TF}}, \quad \Phi^{\text{TF}}(x) = |x|^{-1} - f^{\text{TF}} * |x|^{-1} \geq 0.$$

**Theorem.** *We have  $E^{\text{rHF}} = NE^{\text{TF}} + o(N)_{N \rightarrow \infty}$ .*

This result is not new as we already proved  $E_N = NE^{\text{TF}} + o(N)_{N \rightarrow \infty}$  and  $E_N = E^{\text{rHF}} + O(N^{1/3})$ . Nevertheless, the proof below gives another approach to the validity of the Thomas–Fermi theory  $E_N = NE^{\text{TF}} + o(N)_{N \rightarrow \infty}$ .

*Proof. Lower bound.* Take

$$0 \leq \gamma \leq 1 \text{ on } L^2(\mathbb{R}^3), \quad \text{Tr } \gamma = N$$

and consider

$$\mathcal{E}^{\text{rHF}}(\gamma) = \text{Tr}((-N^{-2/3}\Delta - |x|^{-1})\gamma) + \frac{1}{N}D(\rho_\gamma, \rho_\gamma)$$

where

$$D(f, g) := \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\overline{f(x)}g(y)}{|x-y|} dx dy.$$

Since the Coulomb potential is positive-type, namely  $D(f, f) \geq 0$ , we have

$$\begin{aligned} D(\rho_\gamma, \rho_\gamma) &= D(\rho_\gamma - Nf^{\text{TF}}, \rho_\gamma - Nf^{\text{TF}}) - N^2D(f^{\text{TF}}, f^{\text{TF}}) + 2ND(\rho_\gamma, f^{\text{TF}}) \\ &\geq -N^2D(f^{\text{TF}}, f^{\text{TF}}) + 2ND(\rho_\gamma, f^{\text{TF}}) \\ &= -N^2D(f^{\text{TF}}, f^{\text{TF}}) + N \text{Tr}((f^{\text{TF}} * |x|^{-1})\gamma). \end{aligned}$$

Hence,

$$\begin{aligned} \mathcal{E}^{\text{rHF}}(\gamma) &\geq \text{Tr}((-N^{-2/3}\Delta - |x|^{-1})\gamma) - ND(f^{\text{TF}}, f^{\text{TF}}) + \text{Tr}((f^{\text{TF}} * |x|^{-1})\gamma) \\ &= \text{Tr}((-N^{-2/3}\Delta - \Phi^{\text{TF}})\gamma) - ND(f^{\text{TF}}, f^{\text{TF}}) \\ &\geq \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- - ND(f^{\text{TF}}, f^{\text{TF}}). \end{aligned}$$

Optimizing over  $\gamma$  we obtain the lower bound

$$E^{\text{rHF}} \geq \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- - ND(f^{\text{TF}}, f^{\text{TF}}).$$

Note that

$$\begin{aligned} E^{\text{TF}} + D(f^{\text{TF}}, f^{\text{TF}}) &= K_3^{\text{cl}} \int_{\mathbb{R}^3} (f^{\text{TF}})^{5/3} - \int_{\mathbb{R}^3} \frac{f^{\text{TF}}}{|x|} dx + \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{f^{\text{TF}}(x)f^{\text{TF}}(y)}{|x-y|} dx dy \\ &= K_3^{\text{cl}} \int_{\mathbb{R}^3} (f^{\text{TF}})^{5/3} - \int_{\mathbb{R}^3} f^{\text{TF}} \Phi^{\text{TF}} = -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2} \end{aligned}$$

where we have used the pointwise equality

$$K_3^{\text{cl}}(f^{\text{TF}})^{5/3} - f^{\text{TF}}\Phi^{\text{TF}} + L_{1,3}^{\text{cl}}(\Phi^{\text{TF}})^{5/2} = 0$$

Recall that by the definition of  $K_3^{\text{cl}}$  and  $L_{1,3}^{\text{cl}}$  we have

$$K_3^{\text{cl}}a^{5/3} + L_{1,3}^{\text{cl}}b^{5/2} \geq ab, \quad \forall a, b \geq 0$$

and the equality occurs when

$$b = \frac{5}{3} K_3^{\text{cl}} a^{2/3}$$

which is exactly verified by the TF equation when  $a = f^{\text{TF}}$  and  $b = \Phi^{\text{TF}}$ .

Thus we have proved that

$$E^{\text{rHF}} - NE^{\text{TF}} \geq \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2}.$$

Since  $0 \leq \Phi^{\text{TF}} \in L^{5/2}(\mathbb{R}^3)$ , by Weyl's law we have

$$\lim_{N \rightarrow \infty} N^{-1} \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- = \lim_{\kappa = N^{2/3} \rightarrow \infty} \kappa^{-5/2} \text{Tr}(\Delta - \kappa\Phi^{\text{TF}})_- = -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2}.$$

Thus

$$E^{\text{rHF}} - NE^{\text{TF}} \geq o(N)_{N \rightarrow \infty}.$$

**Upper bound.** As in the proof of Weyl's law, we choose the trial state, with  $\kappa = N^{2/3}$ ,

$$\gamma := \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 - \kappa\Phi^{\text{TF}}(y) < 0) dk dy.$$

Then  $0 \leq \gamma \leq 1$  on  $L^2(\mathbb{R}^3)$  by the resolution of identity and

$$\begin{aligned} \text{Tr} \gamma &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \mathbf{1}(|2\pi k|^2 - \kappa\Phi^{\text{TF}}(y) < 0) dk dy = L_{0,3}^{\text{cl}} \int_{\mathbb{R}^3} |\kappa\Phi^{\text{TF}}(y)|^{3/2} \\ &= NL_{0,3}^{\text{cl}} \int_{\mathbb{R}^3} \left( \frac{5}{3} K_3^{\text{cl}} f^{\text{TF}}(x)^{2/3} \right)^{3/2} = N. \end{aligned}$$

Here we have used  $\int_{\mathbb{R}^3} f^{\text{TF}} = 1$  and

$$L_{0,3}^{\text{cl}} \left( \frac{5}{3} K_3^{\text{cl}} \right)^{3/2} = \frac{|B_1|}{(2\pi)^3} \left( \frac{5}{3} \cdot \frac{3}{5} (6\pi^2)^{2/3} \right)^{3/2} = \frac{\frac{4}{3}\pi}{(2\pi)^3} \cdot 6\pi^2 = 1.$$

Thus proceeding as in the Weyl's law upper bound, we find that

$$\text{Tr}((-\Delta - \kappa\Phi^{\text{TF}})\gamma) = -L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\kappa\Phi^{\text{TF}})^{5/2} + o(\kappa^{5/2})$$

which is equivalent to

$$\mathrm{Tr}((-N^{-2/3}\Delta - \Phi^{\mathrm{TF}})\gamma) = -NL_{1,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} (\Phi^{\mathrm{TF}})^{5/2} + o(N).$$

It remains to bound the direct term  $D(\rho_\gamma - Nf^{\mathrm{TF}}, \rho_\gamma - Nf^{\mathrm{TF}})$ . From the definition

$$\gamma = \iint_{|2\pi k|^2 - \kappa\Phi^{\mathrm{TF}}(y) < 0} |F_{k,y}\rangle\langle F_{k,y}| dk dy.$$

and  $F_{k,y}(x) = e^{2\pi i k \cdot x} G(x - y)$ ,  $\|G\|_{L^2(\mathbb{R}^3)} = 1$ , we can compute explicitly

$$\begin{aligned} \rho_\gamma(x) &= \gamma(x, x) = \iint_{|2\pi k|^2 - \kappa\Phi^{\mathrm{TF}}(y) < 0} |F_{k,y}(x)|^2 dk dy \\ &= \iint_{|2\pi k|^2 - \kappa\Phi^{\mathrm{TF}}(y) < 0} |G(x - y)|^2 dk dy = L_{0,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} (\kappa\Phi^{\mathrm{TF}}(y))^{3/2} |G(x - y)|^2 dy \\ &= N \int_{\mathbb{R}^3} f^{\mathrm{TF}}(y) |G(x - y)|^2 dy = N(f^{\mathrm{TF}} * G^2)(x). \end{aligned}$$

Here we are going to choose  $G^2 = G_N^2 \rightarrow \delta$  slowly as  $N \rightarrow \infty$ . Therefore,

$$f^{\mathrm{TF}} * G_N^2 \rightarrow f^{\mathrm{TF}} \quad \text{strongly in } L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3).$$

Consequently,

$$N^{-2}D(\rho_\gamma - Nf^{\mathrm{TF}}, \rho_\gamma - Nf^{\mathrm{TF}}) = D(f^{\mathrm{TF}} * G_N - f^{\mathrm{TF}}, f^{\mathrm{TF}} * G_N - f^{\mathrm{TF}}) \rightarrow 0.$$

Thus we conclude that for the above choice of  $\gamma$ ,

$$\begin{aligned} E^{\mathrm{rHF}} &\leq \mathcal{E}^{\mathrm{rHF}}(\gamma) = \mathrm{Tr}((-N^{-2/3}\Delta - |x|^{-1})\gamma) + \frac{1}{N}D(\rho_\gamma, \rho_\gamma) \\ &= \mathrm{Tr}((-N^{-2/3}\Delta - \Phi^{\mathrm{TF}}(x))\gamma) - \frac{1}{N}D(Nf^{\mathrm{TF}}, Nf^{\mathrm{TF}}) + \frac{1}{N}D(\rho_\gamma - Nf^{\mathrm{TF}}, \rho_\gamma - Nf^{\mathrm{TF}}) \\ &= -NL_{1,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} (\Phi^{\mathrm{TF}})^{5/2} - ND(f^{\mathrm{TF}}, f^{\mathrm{TF}}) + o(N) \\ &= NE^{\mathrm{TF}} + o(N). \end{aligned}$$

This completes the proof of the upper bound. □

Actually we have the following deeper result

**Theorem** (Scott correction). *We have  $E^{\text{rHF}} = NE^{\text{TF}} + \frac{1}{8}N^{2/3} + o(N^{2/3})_{N \rightarrow \infty}$ .*

Since  $E_N = E^{\text{rHF}} + O(N^{1/3})$ , the Scott correction holds for the full energy  $E_N$  as well. The proof of the Scott correction is significantly more complicated than the derivation of the Thomas–Fermi energy; below we will give an outline of the main ideas.

*Sketch of the proof. Lower bound.* From the above analysis we already showed that

$$\begin{aligned} E^{\text{rHF}} &\geq \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- - ND(f^{\text{TF}}, f^{\text{TF}}) \\ &= NE^{\text{TF}} + \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2}. \end{aligned}$$

Thus we need to prove the following correction to the semiclassical approximation

$$\text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- = -NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2} + \frac{1}{8}N^{2/3} + o(N^{2/3}).$$

Actually the contribution  $\frac{1}{8}N^{2/3}$  comes from the particles moving very close to the nucleus (of a distance  $O(N^{-2/3})$  which is much smaller than the semiclassical distance  $O(1)$ ). The contribution of these particles is comparable to the non-interacting case, namely the hydrogen atom. Thus the key ingredient of the proof is the following

**Theorem** (Hydrogen comparison). *We have*

$$\begin{aligned} &\left| \text{Tr}(-N^{-2/3}\Delta - \Phi^{\text{TF}})_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2} \right. \\ &\quad \left. - \left( \text{Tr}(-N^{-2/3}\Delta - |x|^{-1})_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (|x|^{-1} - 1)_+^{5/2} dx \right) \right| \leq o(N^{2/3}). \end{aligned}$$

Here we replace the potential  $|x|^{-1}$  by  $|x|^{-1} - 1$  to ensure that  $\int_{\mathbb{R}^3} (|x|^{-1} - 1)_+^{5/2} dx < \infty$ . The proof of this result is rather complicated as we need some advanced tools beyond the coherent states discussed in the course. See e.g. **Solovej** and **Spitzer** (2002).

To get the desired conclusion, we use the exact calculation for hydrogen. Recall that the operator  $-h^2\Delta - |x|^{-1}$  on  $L^2(\mathbb{R}^3)$  has eigenvalues  $-1/(4h^2n^2)$  with multiplicity  $n^2$ . Hence,

$$\text{Tr}(-h^2\Delta - |x|^{-1})_- = \sum_{1 \leq n \leq (2h)^{-1}} \left( -\frac{1}{4h^2} + n^2 \right) = -\frac{1}{12h^3} + \frac{1}{8h^2} + O(h^{-1}).$$

On the other hand,

$$L_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (|x|^{-1} - 1)_+^{5/2} dx = \frac{1}{12}.$$

Thus

$$\begin{aligned} & \text{Tr}(-N^{-2/3} \Delta - |x|^{-1} + 1)_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (|x|^{-1} - 1)_+^{5/2} dx \\ &= -\frac{N}{12} + \frac{N^{2/3}}{8} + O(N^{1/3}) + \frac{N}{12} = \frac{N^{2/3}}{8} + O(N^{1/3}). \end{aligned}$$

Hence, we obtain the desired lower bound

$$\begin{aligned} E^{\text{rHF}} - NE^{\text{TF}} &\geq \text{Tr}(-N^{-2/3} \Delta - \Phi^{\text{TF}})_- + NL_{1,3}^{\text{cl}} \int_{\mathbb{R}^3} (\Phi^{\text{TF}})^{5/2} \\ &\geq \frac{N^{2/3}}{8} + o(N^{2/3}). \end{aligned}$$

**Upper bound.** We can show that there exists an operator  $0 \leq \gamma \leq 1$  on  $L^2(\mathbb{R}^3)$  such that  $\text{Tr} \gamma = N$  and

$$\text{Tr}((-N^{-2/3} \Delta - \Phi^{\text{TF}})\gamma) = \text{Tr}(-N^{-2/3} \Delta - \Phi^{\text{TF}})_- + o(N^{2/3}),$$

$$N^{-1}D(\rho_\gamma - Nf^{\text{TF}}, \rho_\gamma - Nf^{\text{TF}}) \leq o(N^{2/3}).$$

The choice of  $\gamma$  is more complicated than just

$$\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |F_{k,y}\rangle \langle F_{k,y}| \mathbf{1}(|2\pi k|^2 - \kappa \Phi^{\text{TF}}(y) < 0) dk dy$$

since we have to do something more precise in the domain  $\{|x| \leq O(N^{2/3})\}$ . Actually the construction of  $\gamma$  follows from the proof of the hydrogen comparison we mentioned above.  $\square$

### 9.3 Bach's correlation inequality

In this section we discuss an improvement of the Lieb–Oxford inequality where the exchange term is taken into account.

**Theorem** (Bach's correlation inequality). *Take an arbitrary normalized wave function*

$\Psi \in L_a^2(\mathbb{R}^{3N})$  and denote  $\gamma = \gamma_\Psi^{(1)}$  and  $\gamma_T = \gamma - \gamma^2$ . Then

$$\left\langle \Psi, \sum_{i < j}^N \frac{1}{|x_i - x_j|} \Psi \right\rangle \geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y) - |\gamma^{(1)}(x, y)|^2}{|x - y|} dx dy - C \left( \int_{\mathbb{R}^3} \rho_\gamma^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} \rho_{\gamma_T}^{4/3} \right)^{1/4}.$$

This inequality was proved by **Bach** (1992) in his proof of the accuracy of the Hartree–Fock energy (we will come to this later). Note that by Pauli’s exclusion principle, the **truncated one-body density matrix**  $\gamma_T = \gamma - \gamma^2$  satisfies

$$0 \leq \gamma_T \leq \gamma \leq 1.$$

Consequently,  $0 \leq \rho_{\gamma_T} \leq \rho_\gamma$ . Moreover, if  $\Psi$  is a Slater determinant, then  $\gamma$  is a projection and hence  $\gamma_T = 0$ ; in this case we have the equality

$$\left\langle \Psi, \sum_{i < j}^N \frac{1}{|x_i - x_j|} \Psi \right\rangle = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y) - |\gamma^{(1)}(x, y)|^2}{|x - y|} dx dy.$$

In applications, when  $\gamma$  is close to a projection, then  $\gamma_T$  is close to 0 and the error term in Bach’s inequality is much smaller than that in the Lieb–Oxford inequality (with the price that we have included the exchange energy on the right side).

Recall that the Lieb–Oxford inequality was proved using the Fefferman-de la Llave decomposition and Hardy–Littlewood maximal function together with the following bound for  $X = \mathbf{1}_{B(z, r)}$ ,

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2} (\text{Tr}[X\gamma])^2 - \frac{1}{2} \text{Tr}(X\gamma)$$

Actually, this inequality easily follows from the identity

$$\sum_{i < j}^N X_i X_j = \frac{1}{2} \left( \sum_{i=1}^N X_i \right)^2 - \frac{1}{2} \sum_{i=1}^N X_i$$

and the Cauchy–Schwarz inequality. For Bach’s inequality, we will need the following improvement:



**Lemma.** Take an arbitrary normalized wave function  $\Psi \in L_a^2(\mathbb{R}^{3N})$  and denote  $\gamma = \gamma_\Psi^{(1)}$  and  $\gamma_T = \gamma - \gamma^2$ . Then we have

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2} \left( (\text{Tr}[X\gamma])^2 - \text{Tr}(X\gamma X\gamma) \right) - C \text{Tr}(X\gamma) \min\{1, (\text{Tr}[X\gamma_T])^{1/2}\}.$$

*Proof.* We follow the representation of **Graf** and **Solovej** (1994). Let  $P, Q$  be operators on  $L^2(\mathbb{R}^3)$  such that

$$1 = P + Q, \quad 1 \geq P, Q \geq 0.$$

Then we can decompose

$$\sum_{i < j}^N X_i X_j = \sum_{i < j}^N (P_i + Q_i)(P_j + Q_j) X_i X_j (P_j + Q_j)(P_i + Q_i) = A_1 + A_2 + A_3$$

where

$$\begin{aligned} A_1 &= \sum_{i < j}^N (P_i Q_j + Q_i P_j + Q_i Q_j) X_i X_j (P_i Q_j + Q_i P_j + Q_i Q_j), \\ A_2 &= \frac{1}{2} \sum_{i \neq j} (P_i P_j X_i X_j P_i P_j + 2P_i P_j X_i X_j P_i Q_j + 2Q_i P_j X_i X_j P_i P_j), \\ A_3 &= \sum_{i < j}^N (P_i P_j X_i X_j Q_i Q_j + Q_i Q_j X_i X_j P_i P_j). \end{aligned}$$

Note that  $A_1 \geq 0$  since  $X \geq 0$ . For  $A_2$ , we take an arbitrary constant  $\alpha \geq 0$  and complete the square

$$\frac{1}{2} \left| \alpha - \sum_i (PXP + 2PXQ)_i \right|^2 = \frac{1}{2} \left| \alpha - \sum_i (PX(1+Q))_i \right|^2 \geq 0$$

which is equivalent to

$$\begin{aligned} A_2 &= \frac{1}{2} \sum_{i \neq j} ((PXP)_i (PXP)_j + 2(PXP)_i (PXQ)_j + 2(QXP)_i (PXP)_j) \\ &\geq -\frac{1}{2} \alpha^2 + \frac{1}{2} \sum_i \left( 2\alpha(X - QXQ) - (1+Q)XP^2X(1+Q) \right)_i - 2 \sum_{i \neq j} (QXP)_i (PXQ)_j \end{aligned}$$

For  $A_3$ , we take an arbitrary constant  $\beta > 0$  and complete the square

$$\frac{1}{2} \left| \beta^{-1/2} \sum_i (QXP)_i - \beta^{1/2} (PXQ)_i \right|^2$$

which is equivalent to

$$\begin{aligned} A_3 &= \sum_{i < j} ((PXQ)_i (PXQ)_j + (QXP)_i (QXP)_j) \\ &\geq -\frac{1}{2} \sum_i \left( (PXQ)^2 + (QXP)^2 + \beta QXP^2XQ + \beta^{-1} PXQ^2XP \right)_i \\ &\quad - \frac{1}{2} (\beta + \beta^{-1}) \sum_{i \neq j} (QXP)_i (PXQ)_j. \end{aligned}$$

Thus in summary, for every  $\alpha \geq 0$  and  $\beta > 0$

$$\begin{aligned} \sum_{i < j}^N X_i X_j &\geq -\frac{1}{2} \alpha^2 - \frac{1}{2} (4 + \beta + \beta^{-1}) \sum_{i \neq j} (QXP)_i (PXQ)_j \\ &\quad + \frac{1}{2} \sum_i \left( 2\alpha (X - QXQ) - (1 + Q)XP^2X(1 + Q) \right. \\ &\quad \left. - (PXQ)^2 - (QXP)^2 - \beta QXP^2XQ - \beta^{-1} PXQ^2XP \right)_i. \end{aligned}$$

The one-body part can be estimated further as follows. We write

$$\begin{aligned} (1 + Q)XP^2X(1 + Q) &= QXP^2X + XP^2XQ + QXP^2XQ + XPX - XPQX \\ &\leq (1 + \beta)QXP^2XQ + \beta^{-1}XP^2X + XPX - XPQX. \end{aligned}$$

For the last term we have

$$\begin{aligned} XPQX - (PXQ)^2 - (QXP)^2 &= (P + Q)XPQX(P + Q) - PXPQXQ - QXPQXP \\ &= PXPQXP + QXPQXQ \geq 0 \end{aligned}$$

Thus

$$\begin{aligned} 2\alpha(X - QXQ) - (1 + Q)XP^2X(1 + Q) - (PXQ)^2 - (QXP)^2 - \beta QXP^2XQ - \beta^{-1} PXQ^2XP \\ &\geq 2\alpha(X - QXQ) - XPX - (1 + 2\beta)QXP^2XQ - \beta^{-1}(XP^2X + PXQ^2XP) \\ &\geq 2\alpha(X - QXQ) - XPX - (1 + 2\beta)\|XP^2X\|QXQ - \beta^{-1}(\|P\|^2X + \|Q\|^2PXP) \end{aligned}$$

$$\begin{aligned} &\geq 2\alpha(X - QXQ) - XPX - (1 + 2\beta) \operatorname{Tr}(XP)QXQ - \beta^{-1}(X + PXP) \\ &= 2\alpha X - XPX - (2\alpha + (1 + 2\beta) \operatorname{Tr}(XP))QXQ - \beta^{-1}(X + PXP). \end{aligned}$$

For the two-body part we have

$$\begin{aligned} \sum_{i \neq j} (QXP)_i (PXQ)_j &= \sum_{i \neq j} (PXP)_i (QXQ)_j - \frac{1}{2} (P_i Q_j - Q_i P_j) X_i X_j (P_i Q_j - Q_i P_j) \\ &\leq \sum_{i \neq j} (PXP)_i (QXQ)_j \leq \left( \sum_i (PXP)_i \right) \left( \sum_j (QXQ)_j \right) \leq \operatorname{Tr}(PXP) \left( \sum_j (QXQ)_j \right). \end{aligned}$$

Here in the last estimate we have used the Pauli's exclusion principle.

**Exercise.** Let  $A$  be a nonnegative trace class operator on  $L^2(\mathbb{R}^d)$ . Prove that

$$\sum_{i=1}^N A_i \leq \operatorname{Tr}(A) \quad \text{on } L_a^2(\mathbb{R}^{dN}).$$

Thus in summary, we have the operator inequality

$$\begin{aligned} \sum_{i < j}^N X_i X_j &\geq -\frac{1}{2} \alpha^2 \\ &+ \frac{1}{2} \sum_{i=1}^N \left( 2\alpha X - XPX - (2\alpha + (5 + 3\beta + \beta^{-1}) \operatorname{Tr}(XP))QXQ - \beta^{-1}(X + PXP) \right)_i. \end{aligned}$$

Therefore, for every normalized wave function  $\Psi \in L_a^2(\mathbb{R}^{3N})$  with  $\gamma = \gamma_\Psi^{(1)}$  we have

$$\begin{aligned} \left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle &\geq -\frac{1}{2} \alpha^2 \\ &+ \frac{1}{2} \operatorname{Tr} \left[ \left( 2\alpha X - XPX - (2\alpha + (5 + 3\beta + \beta^{-1}) \operatorname{Tr}(XP))QXQ - \beta^{-1}(X + PXP) \right) \gamma \right] \end{aligned}$$

Choosing  $\alpha = \operatorname{Tr}[X\gamma]$  we obtain

$$\begin{aligned} \left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle &\geq \frac{1}{2} (\operatorname{Tr}[X\gamma])^2 - \frac{1}{2} \operatorname{Tr}[XPX\gamma] \\ &- \frac{1}{2} \left( \left( 2 \operatorname{Tr}[X\gamma] + (5 + 3\beta + \beta^{-1}) \operatorname{Tr}[XP] \right) \operatorname{Tr}[QXQ\gamma] + \beta^{-1} \operatorname{Tr}[(X + PXP)\gamma] \right) \end{aligned}$$

Now we choose two different projections  $P$  to conclude.

**Choosing**  $P = 1$ , then  $Q = 0$  and we have

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq -\frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma] - \beta^{-1}\text{Tr}[X\gamma].$$

Then we can take  $\beta \rightarrow \infty$  to get

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma].$$

This is the easy bound we have used in the proof of the Lieb–Oxford inequality. If  $\text{Tr}[X\gamma_{\text{T}}] \geq 1$ , then we get the desired conclusion

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma] \min\{1, (\text{Tr}[X\gamma_{\text{T}}])^{1/2}\}.$$

**Choosing**  $P = \gamma$ , then  $Q = 1 - \gamma$ . Using

$$\text{Tr}(QXQ\gamma) = \text{Tr}((1 - \gamma)X(1 - \gamma)\gamma) \leq \text{Tr}(X(1 - \gamma)\gamma) = \text{Tr}(X\gamma_{\text{T}})$$

we have

$$\begin{aligned} \left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle &\geq \frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma X\gamma] \\ &\quad - \frac{1}{2}\left((7 + 3\beta + \beta^{-1})\text{Tr}[X\gamma]\text{Tr}[X\gamma_{\text{T}}] + 2\beta^{-1}\text{Tr}[X\gamma] + \text{Tr}(X\gamma X\gamma)\right) \\ &\geq \frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma X\gamma] \\ &\quad - \frac{1}{2}\text{Tr}[X\gamma]\left((7 + 3\beta + \beta^{-1})\text{Tr}[X\gamma_{\text{T}}] + 2\beta^{-1}\right). \end{aligned}$$

Now we only need to consider when  $\text{Tr}[X\gamma_{\text{T}}] \leq 1$ . In this case,

$$\inf_{\beta > 0} \left( (7 + 3\beta + \beta^{-1})\text{Tr}[X\gamma_{\text{T}}] + 2\beta^{-1} \right) \leq C(\text{Tr}[X\gamma_{\text{T}}])^{1/2}$$

(if  $\text{Tr}[X\gamma_{\text{T}}] = 0$  it is obvious; otherwise we can take  $\beta = (\text{Tr}[X\gamma_{\text{T}}])^{-1/2}$ ). Thus

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2}(\text{Tr}[X\gamma])^2 - \frac{1}{2}\text{Tr}[X\gamma X\gamma] - C\text{Tr}[X\gamma]\text{Tr}[X\gamma_{\text{T}}]^{1/2}.$$

In summary, in all cases we have

$$\left\langle \Psi, \sum_{i < j}^N X_i X_j \Psi \right\rangle \geq \frac{1}{2} (\text{Tr}[X\gamma])^2 - \frac{1}{2} \text{Tr}[X\gamma X\gamma] - C \text{Tr}[X\gamma] \min \left\{ 1, \text{Tr}[X\gamma_{\text{T}}]^{1/2} \right\}.$$

This completes the proof of the lemma. □

Now we are ready to provide

*Proof of Bach's correlation inequality.* Let  $\Psi$  be a normalized wave function in  $L_a^2(\mathbb{R}^{3N})$  and denote  $\gamma = \gamma_{\Psi}^{(1)}$ ,  $\gamma_{\text{T}} = \gamma - \gamma^2$ . Thanks to the Fefferman-de la Llave representation

$$\frac{1}{|x - y|} = \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \mathbf{1}_{B(z,r)}(x) \mathbf{1}_{B(z,r)}(y)$$

we can write

$$\left\langle \Psi, \sum_{i < j}^N \frac{1}{|x_i - x_j|} \Psi \right\rangle = \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \sum_{i < j}^N \left\langle \Psi, \mathbf{1}_{B(z,r)}(x_i) \mathbf{1}_{B(z,r)}(x_j) \Psi \right\rangle.$$

For every  $r > 0$  and  $z \in \mathbb{R}^3$ , applying the previous lemma to  $X = \mathbf{1}_{B(z,r)}$  (this multiplication operator is a projection) we have

$$\begin{aligned} \sum_{i < j}^N \left\langle \Psi, \mathbf{1}_{B(z,r)}(x_i) \mathbf{1}_{B(z,r)}(x_j) \Psi \right\rangle &\geq \frac{1}{2} \left( (\text{Tr}[\mathbf{1}_{B(z,r)}\gamma])^2 - \text{Tr}(\mathbf{1}_{B(z,r)}\gamma \mathbf{1}_{B(z,r)}\gamma) \right) \\ &\quad - C \text{Tr}(\mathbf{1}_{B(z,r)}\gamma) \min \{ 1, (\text{Tr}[\mathbf{1}_{B(z,r)}\gamma_{\text{T}}])^{1/2} \}. \end{aligned}$$

Using the Fefferman-de la Llave representation again we obtain

$$\begin{aligned} \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz (\text{Tr}[\mathbf{1}_{B(z,r)}\gamma])^2 &= \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \int_{\mathbb{R}^3} dx \mathbf{1}_{B(z,r)}(x) \rho_\gamma(x) \int_{\mathbb{R}^3} dy \mathbf{1}_{B(z,r)}(y) \rho_\gamma(y) \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x) \rho_\gamma(y)}{|x - y|} dx dy. \end{aligned}$$

Similarly, using

$$(\mathbf{1}_{B(z,r)}\gamma \mathbf{1}_{B(z,r)}\gamma)(x, x) = \int_{\mathbb{R}^3} dy \mathbf{1}_{B(z,r)}(x) \gamma(x, y) \mathbf{1}_{B(z,r)}(y) \gamma(y, x) = \int_{\mathbb{R}^3} dy \mathbf{1}_{B(z,r)}(x) \mathbf{1}_{B(z,r)}(y) |\gamma(x, y)|^2$$

we can write

$$\begin{aligned} \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \operatorname{Tr}[\mathbf{1}_{B(z,r)} \gamma \mathbf{1}_{B(z,r)} \gamma] &= \frac{1}{\pi} \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \int_{\mathbb{R}^3} dx \int_{\mathbb{R}^3} dy \mathbf{1}_{B(z,r)}(x) \mathbf{1}_{B(z,r)}(y) |\gamma(x,y)|^2 \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|\gamma(x,y)|^2}{|x-y|} dx dy. \end{aligned}$$

Thus

$$\begin{aligned} A &:= \sum_{i < j}^N \left\langle \Psi, \mathbf{1}_{B(z,r)}(x_i) \mathbf{1}_{B(z,r)}(x_j) \Psi \right\rangle - \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x) \rho_\gamma(y) - |\gamma(x,y)|^2}{|x-y|} dx dy \\ &\geq -C \int_0^\infty \frac{dr}{r^5} \int_{\mathbb{R}^3} dz \operatorname{Tr}(\mathbf{1}_{B(z,r)} \gamma) \min\{1, (\operatorname{Tr}[\mathbf{1}_{B(z,r)} \gamma_{\text{T}}])^{1/2}\}. \end{aligned}$$

We control this error term by the maximal functions similarly the proof of the Lieb–Oxford inequality:

$$\operatorname{Tr}(\mathbf{1}_{B(z,r)} \gamma) = \int_{B(z,r)} dy \rho_\gamma(y) \leq Cr^3 M_\rho(z), \quad M_\rho(z) := \sup_{r>0} \frac{1}{|B(z,r)|} \int_{B(z,r)} \rho_\gamma(y) dy$$

and

$$\operatorname{Tr}(\mathbf{1}_{B(z,r)} \gamma_{\text{T}}) \leq Cr^3 M_{\rho_{\text{T}}}(z), \quad M_{\rho_{\text{T}}}(z) = \sup_{r>0} \frac{1}{|B(z,r)|} \int_{B(z,r)} \rho_{\gamma_{\text{T}}}(y) dy.$$

Thus

$$A \geq -C \int_0^\infty \frac{dr}{r^2} \int_{\mathbb{R}^3} dz M_\rho(z) \min\{1, r^{3/2} M_{\rho_{\text{T}}}(z)^{1/2}\}.$$

For every  $z \in \mathbb{R}^3$ , for every  $R = R(z) > 0$  we can bound

$$\begin{aligned} \int_0^\infty \frac{dr}{r^2} \min\{1, r^{3/2} M_{\rho_{\text{T}}}(z)^{1/2}\} &\leq \int_0^R \frac{dr}{r^2} r^{3/2} M_{\rho_{\text{T}}}(z)^{1/2} + \int_R^\infty \frac{dr}{r^2} \\ &\leq C(M_{\rho_{\text{T}}}(z)^{1/2} R^{1/2} + R^{-1}). \end{aligned}$$

Then optimizing over  $R > 0$  we get

$$M_{\rho_{\text{T}}}(z)^{1/2} R^{1/2} \sim R^{-1} \sim \left( (M_{\rho_{\text{T}}}(z)^{1/2} R^{1/2})^2 R^{-1} \right)^{1/3} = M_{\rho_{\text{T}}}(z)^{1/3}$$

we get

$$\int_0^\infty \frac{dr}{r^2} \min\{1, r^{3/2} M_{\rho_{\text{T}}}(z)^{1/2}\} \leq CM_{\rho_{\text{T}}}(z)^{1/3}.$$

Thus

$$\begin{aligned}
A &\geq -C \int_{\mathbb{R}^3} dz M_\rho(z) M_{\rho_T}(z)^{1/3} \\
&\geq -C \left( \int_{\mathbb{R}^3} dz M_\rho(z)^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} dz M_{\rho_T}(z)^{4/3} \right)^{1/4} \\
&\geq -C \left( \int_{\mathbb{R}^3} dz \rho_\gamma(z)^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} dz \rho_{\gamma_T}(z)^{4/3} \right)^{1/4}.
\end{aligned}$$

Here we have used Hölder’s and the Hardy–Littlewood maximal inequalities. This completes the proof of Bach’s correlation inequality.  $\square$

## 9.4 Error bound for atomic Hartree–Fock energy

In this section we compare the ground state energy  $E_N$  of the atomic Hamiltonian

$$H_N = \sum_{i=1}^N \left( -N^{-2/3} \Delta_{x_i} - \frac{1}{|x_i|} \right) + N^{-1} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \quad \text{on } L_a^2(\mathbb{R}^{3N}).$$

with the Hartree–Fock energy

$$E^{\text{HF}} = \inf_{\substack{0 \leq \gamma \leq 1 \\ \text{Tr } \gamma = N}} \mathcal{E}^{\text{HF}}(\gamma)$$

where

$$\mathcal{E}^{\text{HF}}(\gamma) := \text{Tr}((-N^{-2/3} \Delta - |x|^{-1})\gamma) + \frac{1}{2N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x, y)|^2}{|x - y|} dx dy.$$

Since  $E_N \leq E^{\text{HF}} \leq E^{\text{rHF}}$ , from the analysis for  $E^{\text{rHF}}$  in the previous section we know that

$$0 \leq E^{\text{HF}} - E_N \leq E^{\text{rHF}} - E_N = O(N^{1/3}).$$

Now we prove that

**Theorem.** *We have  $E^{\text{HF}} = E_N + o(N^{1/3})_{N \rightarrow \infty}$ .*

This estimate was contained implicitly in the work of **Fefferman** and **Seco** (1990) when they derived the expansion

$$E_N = -c_1 N + c_2 N^{2/3} - c_3 N^{1/3} + o(N^{1/3})_{N \rightarrow \infty}.$$

This result was proved explicitly and elegantly by **Bach** (1992), using his correlation inequality.

In order to put Bach’s correlation inequality in a good use, we need to show that the one–body density matrix of an (approximate) ground state is close to a projection.

**Lemma.** *Let  $\Psi$  be a normalized wave function in  $L_a^2(\mathbb{R}^{3N})$  such that*

$$\langle \Psi, H_N \Psi \rangle = E_N + o(N).$$

*Then the truncated one–body density matrix  $\gamma_T = \gamma - \gamma^2$ ,  $\gamma = \gamma_\Psi^{(1)}$ , satisfies*

$$\mathrm{Tr}(\gamma_T) = o(N)_{N \rightarrow \infty}.$$

*Proof. Step 1.* Let

$$h = -N^{-2/3}\Delta - \Phi^{\mathrm{TF}}, \quad \Phi^{\mathrm{TF}} = |x|^{-1} - f^{\mathrm{TF}} * |x|^{-1}.$$

From the analysis of the reduced Hartree–Fock theory we know that

$$\langle \Psi, H_N \Psi \rangle \geq \mathrm{Tr}(h\gamma) - ND(f^{\mathrm{TF}}, f^{\mathrm{TF}}) + O(N^{1/3}) \geq \mathrm{Tr}(h_- \gamma) - \mathrm{Tr}(h_-) + E^{\mathrm{TF}} + o(N)$$

and

$$E_N \leq E^{\mathrm{rHF}} = E^{\mathrm{TF}} + o(N).$$

Therefore, since  $\Psi$  is an appropriate ground state we find that

$$0 \leq \mathrm{Tr}(h_- \gamma) - \mathrm{Tr}(h_-) \leq o(N).$$

This implicitly implies that  $\gamma$  is close to  $\mathbf{1}(h < 0)$  which is a projection. We will make this quantitative below.

**Step 2.** For every  $\varepsilon > 0$ , define

$$P_\varepsilon = \mathbf{1}(h < -\varepsilon).$$

Using  $0 \leq \gamma_T = \gamma(1 - \gamma) \leq \gamma$  and  $\gamma_T \leq 1 - \gamma$  we have

$$\begin{aligned} \mathrm{Tr}(\gamma_T) &= \mathrm{Tr}((1 - P_\varepsilon)\gamma_T) + \mathrm{Tr}(P_\varepsilon\gamma_T) \leq \mathrm{Tr}((1 - P_\varepsilon)\gamma) + \mathrm{Tr}(P_\varepsilon(1 - \gamma)) \\ &= \mathrm{Tr}(\gamma) - \mathrm{Tr}(P_\varepsilon) + 2\mathrm{Tr}(P_\varepsilon(1 - \gamma)). \end{aligned}$$



Combining with  $P_\varepsilon \leq -h_-/\varepsilon$  we obtain

$$\mathrm{Tr}(\gamma_{\mathrm{T}}) \leq N - \mathrm{Tr}(P_\varepsilon) + 2\varepsilon^{-1} \mathrm{Tr}(h_-(1 - \gamma)).$$

Recall the estimate  $\mathrm{Tr}(h_-(1 - \gamma)) = o(N)$  in Step 1. Moreover, for every  $\varepsilon > 0$ , by Weyl's law for the number of negative eigenvalues we have

$$\begin{aligned} \mathrm{Tr}(P_\varepsilon) &= \mathrm{Tr} \mathbf{1}(h + \varepsilon < 0) = \mathrm{Tr} \mathbf{1}(-\Delta - N^{2/3}(\Phi^{\mathrm{TF}} - \varepsilon) < 0) \\ &= NL_{0,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} [\Phi^{\mathrm{TF}} - \varepsilon]_+^{3/2} + o(N)_{N \rightarrow \infty}. \end{aligned}$$

Thus

$$\limsup_{N \rightarrow \infty} N^{-1} \mathrm{Tr}(\gamma_{\mathrm{T}}) \leq 1 - L_{0,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} [\Phi^{\mathrm{TF}} - \varepsilon]_+^{3/2}.$$

Sending  $\varepsilon \rightarrow 0^+$  on the right side, we conclude that

$$\limsup_{N \rightarrow \infty} N^{-1} \mathrm{Tr}(\gamma_{\mathrm{T}}) \leq 1 - L_{0,3}^{\mathrm{cl}} \int_{\mathbb{R}^3} (\Phi^{\mathrm{TF}})^{3/2} = 0.$$

□

**Exercise.** Let  $0 \leq V \in L^{d/2}(\mathbb{R}^d) \cap L^{1+d/2}(\mathbb{R}^d)$ . Prove that

$$\mathrm{Tr} \mathbf{1}(-\Delta + \lambda V < 0) = L_{0,d}^{\mathrm{cl}} \int_{\mathbb{R}^d} (\lambda V)^{d/2} + o(\lambda^{d/2})_{\lambda \rightarrow \infty}$$

*Hint: You can use Weyl's law for sum of eigenvalues and a Tauberian argument.*

Now we are ready to give

*Proof of the estimate  $E_N - E^{\mathrm{HF}} = o(N^{1/3})$ .* We only need to consider the lower bound for  $E_N - E^{\mathrm{HF}}$ . Take  $\Psi$  an approximate ground state for  $E_N$  such that

$$\langle \Psi, H_N \Psi \rangle = E_N + O(N^{-1}).$$

Let  $\gamma = \gamma_\Psi^{(1)}$  and  $\gamma_{\mathrm{T}} = \gamma - \gamma^2$ . By Bach's correlation inequality we have

$$\begin{aligned} \langle \Psi, H_N \Psi \rangle &\geq \mathrm{Tr}((-\Delta - |x|^{-1})\gamma) + \frac{1}{2N} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y) - |\gamma^{(1)}(x,y)|^2}{|x-y|} dx dy \\ &\quad - CN^{-1} \left( \int_{\mathbb{R}^3} \rho_\gamma^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} \rho_{\gamma_{\mathrm{T}}}^{4/3} \right)^{1/4} \end{aligned}$$

$$\geq E^{\text{HF}} - CN^{-1} \left( \int_{\mathbb{R}^3} \rho_\gamma^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} \rho_{\gamma_T}^{4/3} \right)^{1/4}.$$

It remains to bound the error term on the right side. By Hölder's and the Lieb–Thirring inequalities

$$\int_{\mathbb{R}^3} \rho_\gamma^{4/3} \leq \left( \int_{\mathbb{R}^3} \rho_\gamma^{5/3} \right)^{1/2} \left( \int_{\mathbb{R}^3} \rho_\gamma \right)^{1/2} \leq C(N^{5/3})^{1/2} N^{1/2} = CN^{4/3}.$$

For the truncated one-body density, we use  $\rho_{\gamma_T}^{5/3} \leq \rho_\gamma^{5/3}$  for the kinetic term and the bound  $\text{Tr } \gamma_T = o(N)$  in the previous lemma

$$\int_{\mathbb{R}^3} \rho_{\gamma_T}^{4/3} \leq \left( \int_{\mathbb{R}^3} \rho_{\gamma_T}^{5/3} \right)^{1/2} \left( \int_{\mathbb{R}^3} \rho_{\gamma_T} \right)^{1/2} \leq C(N^{5/3})^{1/2} o(N)^{1/2} = o(N^{4/3}).$$

Thus in summary,

$$N^{-1} \left( \int_{\mathbb{R}^3} \rho_\gamma^{4/3} \right)^{3/4} \left( \int_{\mathbb{R}^3} \rho_{\gamma_T}^{4/3} \right)^{1/4} \leq CN^{-1} o(N^{4/3}) = o(N^{1/3}).$$

This completes the proof of the theorem. □

## 9.5 Hartree–Fock energy of a homogeneous Fermi gas

So far we have focus on the atomic Hamiltonian, which was the original motivation of the development of Hartree–Fock theory. In practice, the Hartree–Fock approximation can be applied to a wide class of fermionic systems. In this section, we consider a simpler situation where the particles are confined in the torus  $\mathbb{T}^3 = [0, 2\pi]^3$  (with the periodic boundary condition), described by the Hamiltonian

$$H_N = \sum_{i=1}^N -N^{-2/3} \Delta_{x_i} + N^{-1} \sum_{1 \leq i < j \leq N} V(x_i - x_j) \quad \text{on } L^2(\mathbb{T}^{3N}).$$

Here the potential  $V : \mathbb{R}^3 \rightarrow \mathbb{R}$  is periodic, even, bounded and of positive type

$$V(x) = \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) e^{ik \cdot x}, \quad \widehat{V}(k) \geq 0, \quad \forall k \in \mathbb{Z}^3.$$

The ground state energy of  $H_N$  is

$$E_N = \inf_{\|\Psi\|_{L^2_a(\mathbb{R}^{3N})}} \langle \Psi, H_N \Psi \rangle.$$

The corresponding Hartree–Fock energy is

$$E^{\text{HF}} = \inf_{\substack{0 \leq \gamma = \gamma^2 \leq 1 \\ \text{Tr } \gamma = N}} \left( \text{Tr}(-N^{-2/3} \Delta \gamma) + \frac{1}{2N} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \left[ \rho_\gamma(x) \rho_\gamma(y) - |\gamma(x, y)|^2 \right] V(x - y) dx dy \right).$$

Without the interaction (i.e.  $V \equiv 0$ ), the Hartree–Fock minimizer is given by the **plane waves**

$$\gamma^{\text{pw}} = \sum_{p \in B_F} |f_p\rangle \langle f_p|, \quad f_p(x) = (2\pi)^{-3/2} e^{ip \cdot x}, \quad \forall p \in \mathbb{Z}^3.$$

Here for simplicity we assume that the Fermi ball  $B_F := B(0, k_F) \cap \mathbb{Z}^3$  contains exactly  $N$  integer points, such that  $\text{Tr } \gamma^{\text{pw}} = N$  (put differently,  $N$  is defined via  $k_F$ ). Thus

$$k_F \sim \left(\frac{3}{4\pi}\right)^{1/3} N^{1/3} + O(1)_{N \rightarrow \infty}.$$

Now we consider the interacting case. Given the regular interaction potential, it turns out that

- The HF theory is good:  $E^{\text{HF}} - E_N$  is  $O(N^{-1/3})$  (instead of  $o(N^{1/3})$  as in the atomic case).
- The HF theory is trivial:  $\gamma^{\text{pw}}$  remains the unique HF minimizer.

**Theorem.** Assume that  $\widehat{V} \geq 0$  and  $\sum_{k \in \mathbb{Z}^3} (1 + |k|) \widehat{V}(k) < \infty$ . Then

$$E_N \geq \mathcal{E}^{\text{HF}}(\gamma^{\text{pw}}) + O(N^{-1/3}).$$

Consequently,  $E^{\text{HF}} = E_N + O(N^{-1/3})$ .

*Proof.* Let  $\Psi_N \in L^2_a(\mathbb{R}^{3N})$  be an arbitrary normalized wave function. For the kinetic energy, we have

$$\left\langle \Psi_N, \sum_{j=1}^N (-\Delta_{x_j}) \Psi_N \right\rangle = \text{Tr}(-\Delta \gamma_{\Psi_N}^{(1)}) \geq \text{Tr}(-\Delta \gamma^{\text{pw}}).$$

Hence, it suffices to show that

$$\left\langle \Psi_N, \sum_{1 \leq \ell < j \leq N} V(x_\ell - x_j) \Psi_N \right\rangle \geq \frac{1}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \left[ \rho_{\gamma^{\text{pw}}}(x) \rho_{\gamma^{\text{pw}}}(y) - |\gamma^{\text{pw}}(x, y)|^2 \right] w(x-y) dx dy + O(N^{2/3}).$$

Since  $\widehat{V} \geq 0$  we can use Onsager's lemma:

$$\begin{aligned} \sum_{1 \leq \ell < j \leq N} V(x_\ell - x_j) + \frac{N}{2} V(0) &= \frac{1}{2} \sum_{\ell, j=1}^N V(x_\ell - x_j) = \frac{1}{2} \sum_{p \in \mathbb{Z}^3} \sum_{\ell, j=1}^N \widehat{V}(p) e^{ip \cdot (x_\ell - x_j)} \\ &= \frac{1}{2} \sum_{p \in \mathbb{Z}^3} \widehat{V}(p) \left| \sum_{\ell=1}^N e^{ip \cdot x_\ell} \right|^2 \geq \frac{1}{2} \widehat{V}(0) N^2. \end{aligned}$$

Note that  $V(0)$  is finite as  $\widehat{V} \in \ell^1$ . Thus we have the pointwise estimate

$$\sum_{1 \leq \ell < j \leq N} V(x_\ell - x_j) \geq \frac{1}{2} \widehat{V}(0) N^2 - \frac{N}{2} V(0)$$

which in particular implies that

$$\left\langle \Psi_N, \sum_{1 \leq \ell < j \leq N} V(x_\ell - x_j) \Psi_N \right\rangle \geq \frac{1}{2} \widehat{V}(0) N^2 - \frac{N}{2} V(0).$$

Next, consider the interaction energy of the plane waves. Note that

$$\gamma^{\text{pw}}(x, y) = \sum_{p \in B_F} f_p(x) \overline{f_p(y)} = (2\pi)^{-3} \sum_{p \in B_F} e^{ip \cdot (x-y)}, \quad \rho^{\text{pw}}(x) = \gamma^{\text{pw}}(x, x) = \frac{N}{(2\pi)^3}.$$

Hence,

$$\int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \rho_{\gamma^{\text{pw}}}(x) \rho_{\gamma^{\text{pw}}}(y) V(x-y) dx dy = N^2 (2\pi)^{-3} \int V = N^2 \widehat{V}(0)$$

and

$$\begin{aligned} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma^{\text{pw}}(x, y)|^2 V(x-y) dx dy &= (2\pi)^{-6} \iint_{\mathbb{T}^3 \times \mathbb{T}^3} \left| \sum_{p \in B_F} e^{ip \cdot (x-y)} \right|^2 V(x-y) dx dy \\ &= (2\pi)^{-3} \int_{\mathbb{T}^3} \left| \sum_{p \in B_F} e^{ip \cdot x} \right|^2 V(x) dx \\ &= (2\pi)^{-3} \int_{\mathbb{T}^3} \sum_{p \in B_F} e^{ip \cdot x} \sum_{q \in B_F} e^{-iq \cdot x} \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) e^{ik \cdot x} dx \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \sum_{p, q \in B_F} (2\pi)^{-3} \int_{\mathbb{T}^3} e^{i(p-q+k) \cdot x} \\
&= \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \sum_{p, q \in B_F} \delta_{p-q+k=0} \\
&= \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) |B_F \cap (B_F + k)|.
\end{aligned}$$

Here  $|B_F \cap (B_F + k)|$  is the number of integer points in  $B_F \cap (B_F + k)$ . We have

$$|B_F \cap (B_F + k)| = |B_F| - |B_F \setminus (B_F + k)| \geq N - C|k|N^{2/3}$$

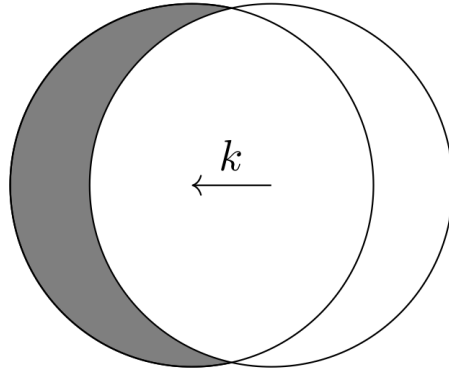


Figure:  $B_F \setminus (B_F + k)$  is the set of integer points in the grey area

Thus the exchange term can be bounded from below by

$$\begin{aligned}
\int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma^{\text{pw}}(x, y)|^2 V(x - y) dx dy &\geq \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) (N - C|k|N^{2/3}) \\
&= N \sum_{k \in \mathbb{Z}^3} V(0) - CN^{2/3} \sum_{k \in \mathbb{Z}^3} |k| \widehat{V}(k) = NV(0) + O(N^{2/3}).
\end{aligned}$$

In summary,

$$\left\langle \Psi_N, \sum_{1 \leq \ell < j \leq N} V(x_\ell - x_j) \Psi_N \right\rangle \geq \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [\rho_{\gamma^{\text{pw}}}(x) \rho_{\gamma^{\text{pw}}}(y) - |\gamma^{\text{pw}}(x - y)|^2] V(x - y) dx dy + O(N^{2/3}).$$

Hence, we conclude that

$$E_N \geq \mathcal{E}^{\text{HF}}(\gamma^{\text{pw}}) + O(N^{-1/3}) \geq E^{\text{HF}} + O(N^{-1/3}) \geq E_N + O(N^{-1/3}).$$

In particular, we have  $E_N = E^{\text{HF}} + O(N^{-1/3})$ . □

**Theorem.** Assume that  $\widehat{V} \geq 0$  and  $N^{-1/3} \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \leq 1/3$ . Then  $\gamma^{\text{pw}}$  is the unique Hartree–Fock minimizer.

*Proof.* We use the argument of **Gontier–Hainzl–Lewin** (2018) where they proved that  $E^{\text{HF}} - \mathcal{E}^{\text{HF}}(\gamma^{\text{pw}})$  is exponentially small for the electron gas in an infinite volume. Here since we have a bounded potential on finite box, the spectral gap of the Laplacian dominates the effect of the mean–field interaction potential.

Recall that  $h = N^{-1/3}$  and  $\lambda = N^{-1}$ . The condition on the potential reads

$$\lambda \|\widehat{V}\|_{\ell^1} \leq \frac{1}{3} h^2.$$

**Step 1.** Take  $0 \leq \gamma = \gamma^2 \leq 1$  on  $L^2(\mathbb{T}^3)$  with  $\text{Tr } \gamma = N$ . Let us show that

$$\begin{aligned} \mathcal{E}(\gamma) - \mathcal{E}(\gamma^{\text{pw}}) &\geq \text{Tr}((-h^2\Delta - \lambda G)(\gamma - \gamma^{\text{pw}})) \\ &\quad - \frac{\lambda}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma(x, y) - \gamma^{\text{pw}}(x, y)|^2 V(x - y) dx dy. \end{aligned}$$

By the definition of the Hartree–Fock functional,

$$\begin{aligned} \mathcal{E}(\gamma) - \mathcal{E}(\gamma^{\text{pw}}) &= \text{Tr}(-h^2\Delta(\gamma - \gamma^{\text{pw}})) - \frac{\lambda}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [|\gamma(x, y)|^2 - |\gamma^{\text{pw}}(x, y)|^2] V(x - y) dx dy \\ &\quad + \frac{\lambda}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [\rho_\gamma(x)\rho_\gamma(y) - \rho_{\gamma^{\text{pw}}}(x)\rho_{\gamma^{\text{pw}}}(x)] V(x - y) dx dy. \end{aligned}$$

Note that  $\rho_{\gamma^{\text{pw}}}$  is a constant since the kernel of  $\gamma^{\text{pw}}$  is  $\gamma^{\text{pw}}(x - y)$  with

$$\gamma^{\text{pw}}(x) = (2\pi)^{-3} \sum_{p \in B_F} e^{ip \cdot x}.$$

Hence, from the condition  $\widehat{V} \geq 0$  and  $\int_{\mathbb{T}^3} \rho_\gamma = \int_{\mathbb{T}^3} \rho_{\gamma^{\text{pw}}} = N = (2\pi)^{-3} \rho_{\gamma^{\text{pw}}}(x)$  we find that

$$\int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [\rho_\gamma(x)\rho_\gamma(y) - \rho_{\gamma^{\text{pw}}}(x)\rho_{\gamma^{\text{pw}}}(x)] V(x - y) dx dy \geq 0.$$

Indeed,

$$\int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \rho_\gamma(x)\rho_\gamma(y) V(x - y) dx dy = \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \rho_\gamma(x)\rho_\gamma(y) \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) e^{-ik \cdot (x-y)} dx dy$$

$$\begin{aligned}
&= (2\pi)^6 \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) |\widehat{\rho}_\gamma(k)|^2 \geq (2\pi)^6 \widehat{V}(0) |\widehat{\rho}_\gamma(0)|^2 \\
&= \widehat{V}(0) N^2 = (2\pi)^6 \widehat{V}(0) |\widehat{\rho}_{\gamma^{\text{pw}}}(k)|^2 \\
&= \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} \rho_{\gamma^{\text{pw}}}(x) \rho_{\gamma^{\text{pw}}}(x) V(x-y) dx dy \geq 0.
\end{aligned}$$

We can rewrite the exchange term using

$$|\gamma(x, y)|^2 - |\gamma^{\text{pw}}(x-y)|^2 = |\gamma(x, y) - \gamma^{\text{pw}}(x-y)|^2 + 2\Re(\gamma(x, y) - \gamma^{\text{pw}}(x-y))\gamma^{\text{pw}}(y-x)$$

and

$$\int_{\mathbb{T}^3} \int_{\mathbb{T}^3} (\gamma(x, y) - \gamma^{\text{pw}}(x-y))\gamma^{\text{pw}}(y-x)V(y-x) dx dy = \text{Tr}((\gamma - \gamma^{\text{pw}})G)$$

where  $G$  is an operator on  $L^2(\mathbb{T}^3)$  with kernel  $\gamma^{\text{pw}}(x-y)V(x-y)$ . Equivalently, in Fourier space  $G$  is the multiplication operator  $G(k)$  with

$$G(k) = \sum_{p \in B_F} \widehat{V}(k-p)$$

because

$$(Gf)(x) = \int_{\mathbb{T}^3} \gamma^{\text{pw}}(x-y)V(x-y)f(y) dy = \int_{\mathbb{T}^3} (2\pi)^{-3} \sum_{p \in B_F} e^{ip \cdot (x-y)} V(x-y) f(y) dy$$

and hence

$$\begin{aligned}
\widehat{Gf}(k) &= (2\pi)^{-3} \int_{\mathbb{T}^3} (Gf)(x) e^{-ik \cdot x} dx \\
&= (2\pi)^{-6} \sum_{p \in B_F} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} e^{-ik \cdot x} e^{ip \cdot (x-y)} V(x-y) f(y) dx dy \\
&= (2\pi)^{-6} \sum_{p \in B_F} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} e^{-ik \cdot (y+z)} e^{ip \cdot z} V(z) f(y) dx dy \\
&= (2\pi)^{-6} \sum_{p \in B_F} \left( \int_{\mathbb{T}^3} e^{-i(k-p) \cdot z} V(z) dz \right) \left( \int_{\mathbb{T}^3} e^{-ik \cdot y} f(y) dy \right) \\
&= \sum_{p \in B_F} \widehat{V}(k-p) \widehat{f}(k).
\end{aligned}$$

In particular  $G \geq 0$  and hence  $\text{Tr}(G(\gamma - \gamma^{\text{pw}})) \in \mathbb{R}$ . Thus

$$\begin{aligned} & \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [|\gamma(x, y)|^2 - |\gamma^{\text{pw}}(x - y)|^2] V(x - y) dx dy \\ &= \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma(x, y) - \gamma^{\text{pw}}(x - y)|^2 V(x - y) dx dy + 2 \text{Tr}(G(\gamma - \gamma^{\text{pw}})). \end{aligned}$$

This completes the desired equality.

**Step 2.** Next, consider the kinetic term. We find an operator  $A \geq 0$  on  $L^2(\mathbb{T}^3)$  such that

$$\text{Tr}[(-h^2\Delta - \lambda G)(\gamma - \gamma^{\text{pw}})] = \text{Tr}[A(\gamma - \gamma^{\text{pw}})^2].$$

The point here is that we can represent the kinetic term as a quadratic expression of  $\gamma - \gamma^{\text{pw}}$ , similarly to the interaction energy. Since  $\gamma$  and  $\gamma^{\text{pw}}$  are projection we have

$$(\gamma - \gamma^{\text{pw}})^2 = (\gamma^{\text{pw}})^\perp (\gamma - \gamma^{\text{pw}}) (\gamma^{\text{pw}})^\perp - \gamma^{\text{pw}} (\gamma - \gamma^{\text{pw}}) \gamma^{\text{pw}}, \quad (\gamma^{\text{pw}})^\perp = 1 - \gamma^{\text{pw}}.$$

Hence,

$$\text{Tr}[A(\gamma - \gamma^{\text{pw}})^2] = \text{Tr}[(\gamma^{\text{pw}})^\perp A (\gamma^{\text{pw}})^\perp - \gamma^{\text{pw}} A \gamma^{\text{pw}}] (\gamma - \gamma^{\text{pw}}).$$

Since  $\text{Tr}(\gamma) = \text{Tr}(\gamma^{\text{pw}}) = N$ , it suffices to find  $A$  such that

$$(\gamma^{\text{pw}})^\perp A (\gamma^{\text{pw}})^\perp - \gamma^{\text{pw}} A \gamma^{\text{pw}} = -h^2\Delta - \lambda G - c_0$$

for a constant  $c_0 \in \mathbb{R}$ . We will choose  $A$  to be the multiplication operator  $A(k)$  on the Fourier space, for which the latter identity becomes

$$A(k)\mathbf{1}(k \in B_F^c) - A(k)\mathbf{1}(k \in B_F) = h^2|k|^2 - \lambda G(k) - c_0,$$

or equivalently

$$A(k) = |h^2|k|^2 - \lambda G(k) - c_0| = \begin{cases} h^2|k|^2 - \lambda G(k) - c_0, & k \in B_F^c \\ -(h^2|k|^2 - \lambda G(k) - c_0), & k \in B_F. \end{cases}$$

This choice is only possible if  $c_0$  satisfies

$$\sup_{k \in B_F} \left( h^2|k|^2 - \lambda G(k) \right) \leq c_0 \leq \inf_{k \in B_F^c} \left( h^2|k|^2 - \lambda G(k) \right).$$



In fact, for every  $k_1 \in B_F$ ,  $k_2 \in B_F^c$  we have  $|k_2| > |k_1|$  (as we assume that the Fermi ball is completely filled) and hence  $|k_2|^2 - |k_1|^2 \geq 1$  (since the left side is a positive integer). On the other hand,

$$0 \leq G(k) = \sum_{p \in B_F} \widehat{V}(p+k) \leq \|\widehat{V}\|_{\ell^1} < \infty.$$

Thus

$$h^2|k_2|^2 - \lambda G(k_2) - \left( h^2|k_1|^2 - \lambda G(k_1) \right) \geq h^2 - \lambda \|V\|_{\ell^1}.$$

Hence, if  $\lambda \|\widehat{V}\|_{\ell^1} \leq h^2/3$ , we have

$$\inf_{k \in B_F^c} \left( h^2|k|^2 - \lambda G(k) \right) - \sup_{k \in B_F} \left( h^2|k|^2 - \lambda G(k) \right) \geq \frac{2}{3}h^2.$$

We can take

$$c_0 = \frac{1}{2} \inf_{k \in B_F^c} \left( h^2|k|^2 - \lambda G(k) \right) + \frac{1}{2} \sup_{k \in B_F} \left( h^2|k|^2 - \lambda G(k) \right)$$

which satisfies the above condition and moreover,

$$A(k) = |h^2|k|^2 - \lambda G(k) - c_0| \geq \frac{h^2}{3}, \quad \forall k \in \mathbb{Z}^3.$$

**Step 3.** Using

$$A \geq h^2/3 \geq \lambda \|\widehat{V}\|_{\ell^1} \geq \lambda \|V\|_{L^\infty}$$

we conclude that

$$\begin{aligned} \mathcal{E}(\gamma) - \mathcal{E}(\gamma^{\text{pw}}) &\geq \text{Tr}[A(\gamma - \gamma^{\text{pw}})^2] - \frac{\lambda}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma(x, y) - \gamma^{\text{pw}}(x, y)|^2 V(x-y) dx dy \\ &\geq \frac{h^2}{3} \text{Tr}[(\gamma - \gamma^{\text{pw}})^2] - \frac{\lambda \|V\|_{L^\infty}}{2} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} |\gamma(x, y) - \gamma^{\text{pw}}(x, y)|^2 \\ &= \left( \frac{h^2}{3} - \frac{\lambda \|V\|_{L^\infty}}{2} \right) \text{Tr}[(\gamma - \gamma^{\text{pw}})^2] \geq \frac{h^2}{6} \text{Tr}[(\gamma - \gamma^{\text{pw}})^2]. \end{aligned}$$

Thus  $\gamma^{\text{pw}}$  is the unique Hartree–Fock minimizer. □

The **correlation energy**  $E_N - E^{\text{HF}}$  will be studied in the next chapter.

# Chapter 10

## Correlation energy

Since Slater determinants are the least correlated fermionic states, the difference  $E_N - E^{\text{HF}}$  is called the **correlation energy**. Calculation the correlation energy is generally difficult. In this chapter, we will formulate a general framework to discuss the correlation energy, and then focus on the homogeneous gas described by the Hamiltonian

$$H_N = \sum_{i=1}^N h^2(-\Delta_{x_i}) + N^{-1} \sum_{1 \leq i < j \leq N} V(x_i - x_j) \quad \text{on } L^2(\mathbb{T}^{3N}), \quad h = N^{-1/3}.$$

We will prove the following result of **Benedikter–Nam–Porta–Schlein–Seiringer** (2020).

**Theorem.** *If  $\widehat{V} \geq 0$ , compactly supported and small enough, then*

$$E_N = E^{\text{HF}} + E^{\text{corr}} + o(N^{-1/3}).$$

*Here the correlation energy is given by  $E^{\text{corr}} = N^{-1/3} \sum_k g(k)$  with*

$$g(k) = -\frac{\pi \kappa^2}{2} \widehat{V}(k) + \frac{\kappa}{\pi} \int_0^\infty \log \left[ 1 + 2\pi \kappa \widehat{V}(k) (1 - \arctan^{-1}(\lambda^{-1})) \right] d\lambda, \quad \kappa = \left( \frac{3}{4\pi} \right)^{1/3}.$$

It is convenient to use the Fock space formalism where the number of particles are not fixed. The reason is that we have to perturb the Hartree–Fock minimizer and the correlation energy will be described by the **excited particles** which live in a Fock space rather than in a fixed  $n$ -body space.

## 10.1 Fock space formalism

**Definition.** Let  $L^2(\Omega)$  be a one-particle Hilbert space (with  $\Omega \subset \mathbb{R}^d$ ). The corresponding **fermionic Fock space** is the Hilbert space

$$\mathcal{F} = \mathcal{F}(L^2(\Omega)) = \bigoplus_{n=0}^{\infty} L_a^2(\Omega^n) = \mathbb{C} \oplus L^2(\Omega) \oplus L_a^2(\Omega^2) \oplus \dots$$

- Any vector in  $\mathcal{F}$  has the form  $\Psi = (\Psi_n)_{n=0}^{\infty}$  where  $\Psi_n \in L_a^2(\Omega^n)$  and

$$\|\Psi\|_{\mathcal{F}}^2 = \sum_{n=0}^{\infty} \|\Psi_n\|_{L_a^2(\Omega^n)}^2$$

- The vector  $|0\rangle = (1, 0, 0, \dots)$  is called the **vacuum**.

On Fock space, we will use the second quantization method, which goes back to **Dirac** (1927). A key concept is **creation and annihilation operators**.

**Definition.** For any  $f \in L^2(\Omega)$ , we can define the **creation operator**  $a^*(f)$  and the **annihilation operator**  $a(f)$  on the fermionic Fock space  $\mathcal{F}(L_a^2(\Omega))$  as follows:

- $a^*(f) : L_a^2(\Omega^n) \rightarrow L_a^2(\Omega^{n+1})$  for all  $n = 0, 1, 2, \dots$

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

- $a(f) : L_a^2(\Omega^{n+1}) \rightarrow L_a^2(\Omega^n)$  for all  $n = 0, 1, 2, \dots$

$$(a(f)\Psi_n)(x_1, \dots, x_{n-1}) = \sqrt{n} \int_{\Omega} \overline{f(x)} \Psi_n(x, x_1, \dots, x_{n-1}) dx.$$

Remarks:

- $f \mapsto a^*(f)$  is **linear**, but  $f \mapsto a(f)$  is **anti-linear**.
- $a(f)|0\rangle = 0$  and  $a^*(f)|0\rangle = f$ . More generally, if  $\{u_i\}_{i=1}^N$  are orthonormal functions in  $L^2(\Omega)$ , then the corresponding Slater determinant can be written as

$$u_1 \wedge u_2 \wedge \dots \wedge u_N = a^*(u_1) \dots a^*(u_N) |0\rangle.$$

For example,

$$a^*(u_1)a^*(u_2)|0\rangle = a^*(u_1)u_2 = \frac{1}{\sqrt{2}}(u_1(x_1)u_2(x_2) - u_1(x_2)u_2(x_1)) = u_1 \wedge u_2.$$

The following exercise shows that we cannot put two particles in the same quantum state, which is consistent with Pauli's exclusion principle (we will come back to Pauli's exclusion principle later).

**Exercise.** Prove that for every  $f \in L^2(\Omega)$ , we have  $(a^*(f))^2 = 0$  on the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$ .

The following exercise shows that  $a^*(f)$  is adjoint of  $a(f)$ .

**Exercise.** Prove that for all  $f \in L^2(\Omega)$ , we have

$$\langle a^*(f)\Psi, \Phi \rangle_{\mathcal{F}} = \langle \Psi, a(f)\Phi \rangle_{\mathcal{F}}, \quad \forall \Psi, \Phi \in \mathcal{F}(L^2(\Omega)).$$

It turns out that the creation and annihilation operators satisfy the following nice algebraic relations.

**Theorem** (Canonical Anti-commutation Relations - CAR). Consider the Fock space  $\mathcal{F}(L^2(\Omega))$ . For all  $f, g \in L^2(\Omega)$ , we have

$$\{a(f), a(g)\} = 0, \quad \{a^*(f), a^*(g)\} = 0, \quad \{a(f), a^*(g)\} = \langle f, g \rangle.$$

Here  $\{A, B\} := AB + BA$ .

A consequence of the third relation is that

$$a^*(f)a(f) + a(f)a^*(f) = \|f\|^2.$$

Thus both  $a(f)$  and  $a^*(f)$  are bounded and  $\|a(f)\|_{\text{op}} \leq \|f\|$ ,  $\|a^*(f)\|_{\text{op}} \leq \|f\|$ .

*Proof.* **Step 1.** First, let us prove that  $\{a(f), a(g)\} = 0$ , namely

$$a(f)a(g) + a(g)a(f) = 0.$$

It suffices to show that

$$a(f)a(g)\Psi_n = -a(g)a(f)\Psi_n$$

for any function  $\Psi_n \in L_a^2(\Omega^n)$  and for any  $n \geq 2$ . By the definition of the annihilation operator, we have

$$\begin{aligned} (a(f)a(g)\Psi_n)(x_1, \dots, x_{n-2}) &= \sqrt{n-1} \int \overline{f(y)}(a(g)\Psi_n)(y, x_1, \dots, x_{n-2}) dy \\ &= \sqrt{n-1} \int \overline{f(y)} \left( \sqrt{n} \int \overline{g(x)}\Psi_n(x, y, x_1, \dots, x_{n-2}) dx \right) dy \\ &= \sqrt{n(n-1)} \iint \overline{f(y)g(x)}\Psi_n(x, y, x_1, \dots, x_{n-2}) dx dy. \end{aligned}$$

Similarly,

$$\begin{aligned} (a(g)a(f)\Psi_n)(x_1, \dots, x_{n-2}) &= \sqrt{n(n-1)} \iint \overline{g(y)f(x)}\Psi_n(x, y, x_1, \dots, x_{n-2}) dx dy \\ &= \sqrt{n(n-1)} \iint \overline{f(y)g(x)}\Psi_n(y, x, x_1, \dots, x_{n-2}) dx dy. \end{aligned}$$

The equality  $a(f)a(g)\Psi_n = -a(g)a(f)\Psi_n$  follows from the anti-symmetry

$$\Psi_n(x, y, x_1, \dots, x_{n-2}) = -\Psi_n(y, x, x_1, \dots, x_{n-2}).$$

**Step 2.** Since  $a^*(f)$  is the adjoint of  $a(f)$ , by taking the adjoint of

$$\{a(f), a(g)\} = a(f)a(g) + a(g)a(f) = 0$$

we find that  $\{a^*(f), a^*(g)\} = 0$ .

**Step 3.** Finally, we prove that

$$\{a(f), a^*(g)\} = a(f)a^*(g) + a^*(g)a(f) = \langle f, g \rangle.$$

When testing with the vacuum, we have

$$a(f)a^*(g)|0\rangle + a^*(g)a(f)|0\rangle = a(f)g - 0 = \langle f, g \rangle.$$

Now consider any function  $\Psi_n \in L_a^2(\Omega^n)$  with any  $n \geq 1$ . We have

$$\begin{aligned}
(a(f)a^*(g)\Psi_n)(x_1, \dots, x_n) &= a(f)(a^*(g)\Psi_n)(x_1, \dots, x_n) \\
&= \sqrt{n+1} \int \overline{f(x_{n+1})} (a^*(g)\Psi_n)(x_{n+1}, x_1, \dots, x_n) dx_{n+1} \\
&= \sqrt{n+1} \int \overline{f(x_{n+1})} (-1)^n (a^*(g)\Psi_n)(x_1, \dots, x_n, x_{n+1}) dx_{n+1} \\
&= \sqrt{n+1} \int \overline{f(x_{n+1})} (-1)^n \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} (-1)^{i-1} g(x_i) \Psi_n(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) dx_{n+1} \\
&= \int \overline{f(x_{n+1})} \sum_{i=1}^{n+1} (-1)^{n+i-1} g(x_i) \Psi_n(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) dx_{n+1} \\
&= \langle f, g \rangle \Psi_n(x_1, \dots, x_n) + \sum_{i=1}^n (-1)^{n+1-i} g(x_i) \int \overline{f(x_{n+1})} \Psi_n(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}) dx_{n+1}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
(a^*(g)a(f)\Psi_n)(x_1, \dots, x_n) &= a^*(g)(a(f)\Psi_n)(x_1, \dots, x_n) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n (-1)^{i-1} g(x_i) (a(f)\Psi_n)(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n (-1)^{i-1} g(x_i) \sqrt{n} \int \overline{f(x_{n+1})} \Psi_n(x_{n+1}, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) dx_{n+1} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n (-1)^{i-1} g(x_i) \sqrt{n} \int \overline{f(x_{n+1})} (-1)^{n-1} \Psi_n(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_{n+1}) dx_{n+1} \\
&= \sum_{i=1}^n (-1)^{n-i} g(x_i) \int \overline{f(x_{n+1})} \Psi_n(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_{n+1}) dx_{n+1}.
\end{aligned}$$

Using Since  $(-1)^{n+1-i} + (-1)^{n-i} = 0$ , we obtain

$$a(f)a^*(g)\Psi_n + a^*(g)a(f)\Psi_n = \langle f, g \rangle \Psi_n$$

for all  $\Psi_n \in L_a^2(\Omega^n)$ . Thus  $\{a(f), a^*(g)\} = \langle f, g \rangle$ .  $\square$

Using the creation and annihilation operators, we can write many operators on Fock space in a convenient way. For example, consider the **number operator**

$$\mathcal{N} := \sum_{n=0}^{\infty} n \mathbb{1}_{L_a^2(\Omega^n)}.$$

For any vector  $\Psi = (\Psi_n)_{n=0}^\infty \in \mathcal{F}$ , the expectation of the number of particles is

$$\langle \Psi, \mathcal{N}\Psi \rangle = \sum_{n=0}^{\infty} n \|\Psi_n\|^2.$$

Let  $\{u_i\}_{i=1}^\infty$  be an orthonormal basis for  $L^2(\Omega)$ . Then we can write

$$\mathcal{N} = \sum_{i=1}^{\infty} a^*(u_i)a(u_i).$$

More generally, we have

**Theorem** (Second quantization of one-body operators). *Let  $h$  be a self-adjoint operator on the one-body Hilbert space  $L^2(\Omega)$ . Then the operator on the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$*

$$d\Gamma(h) := \bigoplus_{n=0}^{\infty} \left( \sum_{i=1}^n h_i \right) = 0 \oplus h \oplus (h \otimes \mathbb{1} + \mathbb{1} \otimes h) \oplus \dots$$

*is called the second quantization of  $h$ . It can be rewritten as*

$$d\Gamma(h) = \sum_{m,n \geq 1} \langle u_m, hu_n \rangle a^*(u_m)a(u_n).$$

*Here  $\{u_n\}_{n \geq 1}$  is an orthonormal basis for  $L^2(\Omega)$ . The representation is **independent of the choice of the basis** (provided that all  $\langle u_m, hu_n \rangle$  are finite).*

*Proof.* Let us write  $a_n = a(u_n)$  for short. It suffices to prove that

$$\sum_{i=1}^N h_i \Psi_N = \sum_{m,n \geq 1} \langle u_m, hu_n \rangle a_m^* a_n \Psi_N$$

for all  $\Psi_N \in \mathcal{H}^{\otimes_s N}$  and for all  $N$ . Recall from a previous computation

$$(a_m^* a_n \Psi_N)(x_1, \dots, x_N) = \sum_{i=1}^N (-1)^{N-i} u_m(x_i) \int \overline{u_n(y)} \Psi_N(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, y) dy.$$

Therefore,

$$\sum_{m,n} \langle u_m, hu_n \rangle (a_m^* a_n \Psi_N)(x_1, \dots, x_N)$$

$$\begin{aligned}
&= \sum_{m,n} \langle u_m, hu_n \rangle \sum_{i=1}^N (-1)^{N-i} u_m(x_i) \int \overline{u_n(y)} \Psi_N(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, y) dy \\
&= \sum_{i=1}^N (-1)^{N-i} \sum_n \left( \sum_m \langle u_m, hu_n \rangle u_m(x_i) \right) \int \overline{u_n(y)} \Psi_N(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, y) dy \\
&= \sum_{i=1}^N (-1)^{N-i} \sum_n (hu_n)(x_i) \int \overline{u_n(y)} \Psi_N(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, y) dy \\
&= \sum_{i=1}^N (-1)^{N-i} \sum_n \left[ (|hu_n\rangle \langle u_n|)_N \Psi_N \right] (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, x_i) \\
&= \sum_{i=1}^N (-1)^{N-i} \left[ \left( h \sum_n |u_n\rangle \langle u_n| \right)_N \Psi_N \right] (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, x_i) \\
&= \sum_{i=1}^N (-1)^{N-i} \left[ h_N \Psi_N \right] (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, x_i) = \sum_{i=1}^N \left[ h_i \Psi_N \right] (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_N).
\end{aligned}$$

Here we have used the Parseval's identity

$$\sum_m \langle u_m, hu_n \rangle u_m = hu_n,$$

the resolution of the identity operator

$$\sum_n |u_n\rangle \langle u_n| = \mathbf{1},$$

and the anti-symmetry

$$(-1)^{N-i} \Psi(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N, x_i) = \Psi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_N).$$

This completes the proof. □

For the two-body interaction operators, we have

**Theorem** (Second quantization of two-body operators). *Let  $W$  be a self-adjoint operator on  $L^2(\Omega^2)$  such that  $W_{12} = W_{21}$ . Then the operator on the fermionic Fock space  $\mathcal{F}(\mathcal{H})$*

$$\bigoplus_{n=0}^{\infty} \left( \sum_{1 \leq i < j \leq n} W_{ij} \right) = 0 \oplus 0 \oplus W_{12} \oplus (W_{12} + W_{23} + W_{13}) \oplus \dots$$



is called the second quantization of  $W$ . It can be rewritten as

$$\bigoplus_{n=0}^{\infty} \left( \sum_{1 \leq i < j \leq n} W_{ij} \right) = \frac{1}{2} \sum_{m,n,p,q \geq 1} \langle u_m \otimes u_n, W u_p \otimes u_q \rangle a^*(u_m) a^*(u_n) a(u_q) a(u_p)$$

Here  $\{u_n\}_{n \geq 1}$  is an orthonormal basis for  $L^2(\Omega)$ . The representation is **independent of the choice of the basis**.

The proof of this result is left as an exercise.

Remarks:

- From the method of second quantization, the typical Hamiltonian

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

on  $L_a^2(\Omega^N)$  can be extended to be an operator on the fermionic Fock space  $\mathcal{F}$  as

$$\bigoplus_{N=0}^{\infty} H_N = \sum_{m,n} h_{mn} a_m^* a_n + \frac{1}{2} \sum_{m,n,p,q} W_{mnpq} a_m^* a_n^* a_q a_p$$

where  $a_n = a(u_n)$  with an orthonormal basis  $\{u_n\}$  for  $\mathcal{H}$  and

$$h_{mn} = \langle u_m, h u_n \rangle, \quad W_{mnpq} = \langle u_m \otimes u_n, W u_p \otimes u_q \rangle.$$

- In the littérature, people also use the creation and annihilation operators  $a_x^*$  and  $a_x$ ,  $x \in \Omega$ , defined by

$$a^*(f) = \int f(x) a_x^* dx, \quad a(f) = \int \overline{f(x)} a_x dx, \quad \forall f \in \mathcal{H}.$$

These operator-valued distributions satisfy the CAR

$$\{a_x, a_y\} = 0, \quad \{a_x^*, a_y^*\} = 0, \quad \{a_x, a_y^*\} = \delta_0(x - y).$$

The advantage of these notations is that we can use the second quantization without specifying an orthonormal basis for  $\mathcal{H}$ . For example, the typical Hamiltonian

$$H_N = \sum_{i=1}^N (-\Delta_{x_i} + V(x_i)) + \sum_{1 \leq i < j \leq N} W(x_i - x_j)$$

on  $L_a^2(\Omega^N)$  can be extended to be an operator on Fock space as

$$\bigoplus_{N=0}^{\infty} H_N = \int_{\Omega} a_x^*(-\Delta_x + V(x))a_x dx + \frac{1}{2} \iint_{\Omega \times \Omega} W(x-y)a_x^*a_y^*a_y a_x dx dy.$$

**Definition.** Let  $\Psi$  be a normalized vector in the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$ . Its one-body density matrix  $\gamma_{\Psi}^{(1)}$  is a trace class operator on  $L^2(\Omega)$  defined by

$$\langle g, \gamma_{\Psi}^{(1)} f \rangle = \langle \Psi, a^*(f)a(g)\Psi \rangle, \quad \forall f, g \in L^2(\Omega)$$

**Exercise.** Let  $\Psi$  be a normalized vector in the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$  with  $\langle \Psi, \mathcal{N}\Psi \rangle < \infty$ . Prove that its one-body density matrix satisfies

$$0 \leq \gamma_{\Psi}^{(1)} \leq 1, \quad \text{Tr} \gamma_{\Psi}^{(1)} = \langle \Psi, \mathcal{N}\Psi \rangle.$$

## 10.2 Particle-hole transformation

**Theorem.** Let  $\{u_i\}_{i=1}^{\infty}$  be an orthonormal basis for  $L^2(\Omega)$ . For every  $N \geq 1$ , there exists a unitary operator  $R$  on the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$  such that

$$R|0\rangle = u_1 \wedge u_2 \wedge \dots \wedge u_N = a^*(u_1)\dots a^*(u_N)|0\rangle$$

and

$$R^* a^*(u_i) R = \begin{cases} a(u_i) & \text{if } i \leq N, \\ a^*(u_i) & \text{if } i > N. \end{cases}$$

Moreover,  $R = R^* = R^{-1}$ .

*Proof.* The fermionic Fock space has an orthonormal basis of Slater determinants

$$\left\{ a^*(u_{i_1})\dots a^*(u_{i_\ell})|0\rangle : 1 \leq i_1 < i_2 < \dots < i_\ell, \quad \ell = 0, 1, 2, \dots \right\}.$$

The operator  $R$  is defined by

$$R a^*(u_{i_1})\dots a^*(u_{i_\ell})|0\rangle = a^*(u_{j_1})\dots a^*(u_{j_k})|0\rangle$$

where  $1 \leq j_1 < j_2 < \dots < j_k$  is determined from  $1 \leq i_1 < i_2 < \dots < i_\ell$  such that

$$\{j_1, \dots, j_k\} = \bigcup \{i_s : i_s > N\} \bigcup \{1 \leq r \leq N : r \notin \{i_1, \dots, i_\ell\}\}.$$

Then clearly  $R$  is a unitary operator since it maps an orthonormal basis to an orthonormal basis. Moreover,  $R = R^* = R^{-1}$ . Since the identity

$$R^* a^*(u_i) R = \begin{cases} a(u_i) & \text{if } i \leq N, \\ a^*(u_i) & \text{if } i > N \end{cases}$$

holds for all Slater determinants, it holds for any vector on Fock space by the linearity.  $\square$

Note that the transformation  $R$  is a special example of a **Bogoliubov transformation**. We will come back to Bogoliubov theory later in connection to the **bosonic picture**.

Now let us focus on the homogeneous Fermi gas where  $N$  fermions are confined in the torus  $\mathbb{T}^3 = [0, 2\pi]^3$  (with the periodic boundary condition), described by the Hamiltonian

$$H_N = \sum_{i=1}^N h^2(-\Delta_{x_i}) + N^{-1} \sum_{1 \leq i < j \leq N} V(x_i - x_j) \quad \text{on } L^2(\mathbb{T}^{3N}), \quad h = N^{-1/3}.$$

For simplicity we assume that  $\widehat{V} \geq 0$  and is compactly supported. Using the annihilation operators

$$a_p = a(u_p), \quad u_p = (2\pi)^{-3/2} e^{ip \cdot x}$$

we can write in the second quantization formalism (**why?**)

$$H_N = \sum_{k \in \mathbb{Z}^3} h^2 p^2 a_p^* a_p + \frac{1}{2N} \sum_{k, p, q \in \mathbb{Z}^3} \widehat{V}(k) a_{p+k}^* a_{q-k}^* a_q a_p.$$

Also for simplicity we assume that

$$N = |B_F|, \quad B_F = B(0, k_F) \cap \mathbb{Z}^3.$$

The Fermi ball  $B_F$  corresponding to the Slater determinant  $\bigwedge_{p \in B_F} u_p$  which is the unique Hartree–Fock minimizer. To find a correction to the Hartree–Fock theory, we apply the particle–hole transformation

$$R|0\rangle = \bigwedge_{p \in B_F} u_p$$

and

$$R^* a_p^* R = \begin{cases} a_p & \text{if } p \in B_F, \\ a_p^* & \text{if } p \in B_F^c. \end{cases}$$

This operator really makes a hole in a Fermi ball and create a particle outside.

For a warm-up, let us consider the kinetic term. We have

$$\begin{aligned} R^* \left( \sum_{p \in \mathbb{Z}^3} h^2 p^2 a_p^* a_p \right) R &= \sum_{p \in \mathbb{Z}^3} h^2 p^2 R^* (a_p^* a_p) R \\ &= \sum_{p \in B_F^c} h^2 p^2 (a_p^* a_p) + \sum_{p \in B_F} h^2 p^2 (a_p a_p^*) \\ &= \sum_{p \in B_F^c} h^2 p^2 (a_p^* a_p) + \sum_{p \in B_F} h^2 p^2 (1 - a_p^* a_p) \\ &= \sum_{p \in B_F^c} h^2 p^2 (a_p^* a_p) - \sum_{p \in B_F} h^2 p^2 a_p^* a_p + h^2 \sum_{p \in B_F} p^2. \end{aligned}$$

We define

$$\mathbb{H}_0 := R^* \left( \sum_{p \in \mathbb{Z}^3} h^2 p^2 a_p^* a_p \right) R - h^2 \sum_{p \in B_F} p^2 = \sum_{p \in B_F^c} h^2 p^2 (a_p^* a_p) - \sum_{p \in B_F} h^2 p^2 a_p^* a_p.$$

Note that the operator  $\mathbb{H}_0$  does not seem positive at first sight, but it is, at least for the relevant class of wave functions.

**Lemma.** *There exists a constant  $c_0 \in \frac{1}{2}\mathbb{N}$  such that for every  $\psi = R\Psi$  with  $\Psi$  a normalized function in  $L_a^2(\mathbb{T}^{3N})$ , we have*

$$\mathbb{H}_0 \psi = h^2 \sum_{k \in \mathbb{Z}^3} |k^2 - c_0| a_k^* a_k \psi \quad \text{and} \quad \inf_{k \in \mathbb{Z}^3} |k^2 - c_0| \geq \frac{1}{2}.$$

*Proof.* Note that

$$\begin{aligned} R^* \mathcal{N} R &= R^* \left( \sum_{p \in \mathbb{Z}^3} a_p^* a_p \right) R = \sum_{p \in B_F^c} a_p^* a_p + \sum_{p \in B_F} a_p a_p^* \\ &= \sum_{p \in B_F^c} a_p^* a_p + \sum_{p \in B_F} (1 - a_p^* a_p) = N + \sum_{p \in B_F^c} a_p^* a_p - \sum_{p \in B_F} a_p^* a_p \\ &=: N + \mathcal{N}^p - \mathcal{N}^h. \end{aligned}$$

Hence, the condition  $\psi = R\Psi$  with  $\Psi \in L_a^2(\mathbb{T}^{3N})$  implies that

$$(\mathcal{N}^p - \mathcal{N}^h)\psi = 0.$$

Using the fact that  $|k_2|^2 - |k_1|^2 \geq 1$  if  $k_2 \in B_F^c$ ,  $k_1 \in B_F$  (since the Fermi ball is completely filled) we obtain

$$\inf_{k \in B_F^c} |k|^2 - \sup_{k \in B_F} |k|^2 \geq 1.$$

Define

$$c_0 := \frac{1}{2} \inf_{k \in B_F^c} |k|^2 + \frac{1}{2} \sup_{k \in B_F} |k|^2.$$

Then

$$\sup_{k \in B_F} |k|^2 \leq c_0 \leq \inf_{k \in B_F^c} |k|^2, \quad \inf_{k \in \mathbb{Z}^3} |k^2 - c_0| \geq \frac{1}{2}.$$

Moreover, using  $(\mathcal{N}^p - \mathcal{N}^h)\psi = 0$  we find that

$$\begin{aligned} \mathbb{H}_0\psi &= \sum_{p \in B_F^c} h^2 p^2 a_p^* a_p \psi - \sum_{p \in B_F} h^2 p^2 a_p^* a_p \psi \\ &= \sum_{p \in B_F^c} h^2 (p^2 - c_0) a_p^* a_p \psi + \sum_{p \in B_F} h^2 (c_0 - p^2) a_p^* a_p \psi + c_0 (\mathcal{N}^p - \mathcal{N}^h)\psi \\ &= \sum_{p \in \mathbb{Z}^3} h^2 |p^2 - c_0| a_p^* a_p \psi. \end{aligned}$$

□

Now we turn to the interaction part. We introduce the set  $\Gamma^{\text{nor}}$  of all momenta  $k = (k_1, k_2, k_3)$  in  $\mathbb{Z}^3 \cap \text{supp } \hat{V}$  satisfying

$$k_3 > 0 \text{ or } (k_3 = 0 \text{ and } k_2 > 0) \text{ or } (k_2 = k_3 = 0 \text{ and } k_1 > 0).$$

This set is chosen such that

$$\Gamma^{\text{nor}} \cap (-\Gamma^{\text{nor}}) = \emptyset, \quad \Gamma^{\text{nor}} \cup (-\Gamma^{\text{nor}}) = (\mathbb{Z}^3 \cap \text{supp } \hat{V}) \setminus \{0\}.$$

A length but straightforward computation shows that

$$R^* \left( \frac{1}{2N} \sum_{k,p,q \in \mathbb{Z}^3} \hat{V}(k) a_{p+k}^* a_{q-k}^* a_q a_p \right) R = \frac{N}{2} \hat{V}(0) - \frac{1}{2N} \sum_{p,q \in B_F} \hat{V}(p-q) + \mathbb{Q} + \mathbb{X}$$

where

$$\begin{aligned} \mathbb{Q} &= N^{-1} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( b^*(k)b(k) + b^*(-k)b(-k) + b^*(k)b^*(-k) + b(-k)b(k) \right), \\ \mathbb{X} &= \frac{1}{2N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( \mathfrak{D}^*(k)\mathfrak{D}(k) + \mathfrak{D}^*(-k)\mathfrak{D}(-k) \right) \\ &\quad + \frac{1}{2N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( \mathfrak{D}^*(-k)b(k) + \mathfrak{D}^*(k)b(-k) + h.c. \right) \\ &\quad + \frac{1}{2N} \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \left( 3 \sum_{h \in B_F \cap (B_F+k)} a_h^* a_h - 2 \sum_{p \in B_F^c \cap (B_F+k)} a_p^* a_p - \sum_{p \in B_F^c \cap (B_F^c+k)} a_p^* a_p \right) \end{aligned}$$

and

$$\begin{aligned} b^*(k) &:= \sum_{p \in B_F^c \cap (B_F+k)} a_p^* a_{p-k}, \\ \mathfrak{D}^*(k) &:= \sum_{p \in B_F^c \cap (B_F^c+k)} a_p^* a_{p-k} - \sum_{p \in B_F \cap (B_F+k)} a_p^* a_{p-k}. \end{aligned}$$

In summary, we have

**Lemma.**

$$R^* H_N R = E^{\text{HF}} + \mathbb{H}_0 + \mathbb{Q} + \mathbb{X}.$$

This follows from the above computations and the following expression of the Hartree–Fock energy

$$\begin{aligned} E^{\text{HF}} &= h^2 \text{Tr}(-\Delta \gamma^{\text{pw}}) + \frac{1}{2N} \int_{\mathbb{T}^3} \int_{\mathbb{T}^3} [\rho_{\gamma^{\text{pw}}}(x)\rho_{\gamma^{\text{pw}}}(y) - |\gamma^{\text{pw}}(x-y)|^2] V(x-y) dx dy \\ &= h^2 \sum_{p \in B_F} p^2 + \frac{N}{2} \widehat{V}(0) - \frac{1}{2N} \sum_{p, q \in B_F} \widehat{V}(p-q). \end{aligned}$$

since

$$\gamma^{\text{pw}}(x-y) = (2\pi)^{-3} \sum_{p \in B_F} e^{ip \cdot x} e^{-ip \cdot y} = \sum_{p \in B_F} e^{ip \cdot (x-y)}, \quad \rho_{\gamma^{\text{pw}}}(x) = (2\pi)^{-3} N.$$

As we will see, the terms  $\mathbb{H}_0$  and  $\mathbb{Q}$  contribute to the leading order of the correlation energy (which is  $\sim N^{-1/3}$ ) and the term  $\mathbb{X}$  can be treated as a small error. More precisely,  $\mathbb{H}_0$  and  $\mathbb{Q}$  are **bosonizable terms**, namely they can be compared with certain **quasi-bosonic operators**, while  $\mathbb{X}$  is **non-bosonizable** but can be removed.

## 10.3 Estimates for kinetic and number operators

In this section we derive some useful estimates for  $\mathbb{H}_0$  and  $\mathcal{N}$ . We start with

**Lemma.** *Let  $\Psi \in L_a^2(\mathbb{T}^3)$  be a normalized function such that*

$$\langle \Psi, H_N \Psi \rangle \leq E^{\text{HF}} + O(N^{-1/3}).$$

*Then the state  $\psi = R\Psi$  satisfies*

$$\langle \psi, \mathbb{H}_0 \psi \rangle \leq CN^{-1/3}, \quad \langle \psi, \mathcal{N} \psi \rangle \leq CN^{1/3}.$$

*Proof.* In the previous chapter we have proved that

$$\frac{1}{N} \sum_{i < j}^N V(x_i - x_j) \geq \frac{N}{2} \widehat{V}(0) - \frac{1}{2} V(0) = \frac{N}{2} \widehat{V}(0) - \frac{1}{2N} \sum_{p, q \in B_F} \widehat{V}(p - q) + O(N^{-1/3}).$$

Moreover, for the kinetic part we can write

$$h^2 \left\langle \Psi, \sum_{p \in \mathbb{Z}^3} p^2 a_p^* a_p \Psi \right\rangle = h^2 \left\langle R\psi, \left( \sum_{p \in \mathbb{Z}^3} p^2 a_p^* a_p \right) R\psi \right\rangle = h^2 \sum_{p \in B_F} p^2 + \langle \psi, \mathbb{H}_0 \psi \rangle.$$

Thus

$$\langle \Psi, H_N \Psi \rangle \geq E^{\text{HF}} + O(N^{-1/3}) + \langle \psi, \mathbb{H}_0 \psi \rangle.$$

Consequently, if  $\langle \Psi, H_N \Psi \rangle \leq E^{\text{HF}} + O(N^{-1/3})$ , then

$$\langle \psi, \mathbb{H}_0 \psi \rangle \leq CN^{-1/3}.$$

Moreover, we can write

$$\mathbb{H}_0 \psi = h^2 \sum_{k \in \mathbb{Z}^3} |k^2 - c_0| a_k^* a_k \psi \quad \text{and} \quad \inf_{k \in \mathbb{Z}^3} |k^2 - c_0| \geq \frac{1}{2}.$$

and hence

$$\langle \psi, \mathbb{H}_0 \psi \rangle = h^2 \sum_{k \in \mathbb{Z}^3} |k^2 - c_0| \langle \psi, a_k^* a_k \psi \rangle \geq \frac{h^2}{2} \sum_{k \in \mathbb{Z}^3} \langle \psi, a_k^* a_k \psi \rangle = \frac{h^2}{2} \sum_{k \in \mathbb{Z}^3} \langle \psi, \mathcal{N} \psi \rangle.$$

Since  $\langle \psi, \mathbb{H}_0 \psi \rangle \leq CN^{-1/3}$  and  $h = N^{-1/3}$ , we find that  $\langle \psi, \mathcal{N} \psi \rangle \leq CN^{1/3}$ .  $\square$

By a technical reason, we will focus on a **well-prepared approximate ground state**.

**Lemma.** *There exists a normalized function  $\Psi \in L_a^2(\mathbb{T}^{3N})$  such that*

$$\langle \Psi, H_N \Psi \rangle \leq E_N + O(N^{-2/3})$$

*and that the state  $\psi = R\Psi$  satisfies*

$$\psi = \mathbf{1}(\mathcal{N} \leq CN^{1/3})\psi, \quad \langle \psi, \mathbb{H}_0 \psi \rangle \leq CN^{-1/3}.$$

We will need the following localization technique on Fock space. The idea goes back to **Lieb** and **Solovej** (2001). The formulation below is taken from a paper of **Lewin–Nam–Serfaty–Solovej** (2013).

**Lemma** (IMS formula on Fock space). *Let  $A$  be a non-negative operator on the fermionic Fock space  $\mathcal{F}(L^2(\Omega))$  such that  $P_i D(A) \subset D(A)$  and  $P_i A P_j = 0$  if  $|i - j| > \ell$ , where  $P_i = \mathbf{1}(\mathcal{N} = i)$ . Let  $f, g : \mathbb{R} \rightarrow [0, 1]$  be smooth functions such that  $f^2 + g^2 = 1$ ,  $f(x) = 1$  for  $x \leq 1/2$  and  $f(x) = 0$  for  $x \geq 1$ . For any  $M \geq 1$  define*

$$f_M := f(\mathcal{N}/M), \quad g_M := g(\mathcal{N}/M).$$

*Then*

$$\pm \left( A - f_M A f_M - g_M A g_M \right) \leq \frac{C_f \ell^3}{M^2} [A]_{\text{diag}} \mathbf{1}(\mathcal{N} \leq M + \ell)$$

*where  $C_f = \|f'\|_{L^\infty}^2 + \|g'\|_{L^\infty}^2$  and  $[A]_{\text{diag}} := \sum_{i=0}^{\infty} P_i A P_i$ .*

*Proof.* Using the “double commutator identities”

$$[[A, f_M], f_M] = f_M^2 A + A f_M^2 - 2f_M A f_M,$$

$$[[A, g_M], g_M] = g_M^2 A + A g_M^2 - 2g_M A g_M.$$

we have the “IMS-identity”

$$A - f_M A f_M - g_M A g_M = \frac{1}{2} ([[A, f_M], f_M] + [[A, g_M], g_M]).$$

This is an analogue of the standard formula for the Laplacian

$$(-\Delta) - f(-\Delta)f - g(-\Delta)g = -\frac{1}{2}(|\nabla f|^2 + |\nabla g|^2), \quad f^2 + g^2 = 1$$



which was named after Ismagilov, Morgan, Simon and Israel Michael Sigal.

Next, by decomposing further

$$\mathbf{1}_{\mathcal{F}} = \sum_{i=0}^{\infty} P_i$$

we find that

$$\begin{aligned} [[A, f_M], f_M] &= \sum_{i,j=0}^{\infty} P_i [[A, f_M], f_M] P_j = \sum_{i,j=0}^{\infty} \left( f_M^2(i) + f_M^2(j) - 2f_M(i)f_M(j) \right) P_i A P_j \\ &= \sum_{i,j=0}^{\infty} \left( f_M(i) - f_M(j) \right)^2 P_i A P_j = \sum_{1 \leq |i-j| \leq \ell} \left( f(i/M) - f(j/M) \right)^2 P_i A P_j. \end{aligned}$$

In the last equality we have used the assumption that  $P_i A P_j = 0$  if  $|i - j| > \ell$ . Combining with a similar formula for  $g_M$ , we arrive at

$$A - f_M A f_M - g_M A g_M = \frac{1}{2} \sum_{1 \leq |i-j| \leq \ell} \left[ \left( f(i/M) - f(j/M) \right)^2 + \left( g(i/M) - g(j/M) \right)^2 \right] P_i A P_j.$$

Since  $f, g$  are smooth, we have the uniform bound for all  $|i - j| \leq \ell$ :

$$\left( f(i/M) - f(j/M) \right)^2 + \left( g(i/M) - g(j/M) \right)^2 \leq C_f \frac{\ell^2}{M^2} \mathbf{1}(i, j \leq M + \ell), \quad C_f = \|f'\|_{L^\infty}^2 + \|g'\|_{L^\infty}^2.$$

On the other hand, since  $A \geq 0$  we have the Cauchy-Schwarz inequality

$$\pm (P_i A P_j + h.c.) \leq P_i A P_i + P_j A P_j.$$

Thus we conclude that

$$\begin{aligned} &\pm \left( A - f_M A f_M - g_M A g_M \right) \\ &= \pm \frac{1}{4} \sum_{1 \leq |i-j| \leq \ell} \left[ \left( f(i/M) - f(j/M) \right)^2 + \left( g(i/M) - g(j/M) \right)^2 \right] (P_i A P_j + P_j A P_i) \\ &\leq \frac{1}{4} \frac{C_f \ell^2}{M^2} \sum_{1 \leq |i-j| \leq \ell} \mathbf{1}(i, j \leq M + \ell) (P_i A P_i + P_j A P_j) \\ &\leq \frac{C_f \ell^3}{M^2} \sum_{i=0}^{\infty} \mathbf{1}(i \leq M + \ell) P_i A P_i = \frac{C_f \ell^3}{M^2} A_{\text{diag}} \mathbf{1}(\mathcal{N} \leq M + \ell). \end{aligned}$$

This completes the proof of the lemma. □

*Application to the homogeneous gas. Step 1.* We will apply the above abstract result to

$$A = R^* H_N R - E^{\text{HF}} + C_0 N^{-1/3}$$

We can choose  $C_0$  sufficiently large such that  $A \geq 0$ . We can take  $\ell = 4$  as the Hamiltonian  $A$  changes particle number by at most  $\pm 4$ . Moreover, from the explicit formula

$$A = \mathbb{H}_0 + \mathbb{Q} + \mathbb{X} + C_0 N^{-1/3}$$

we find that

$$\begin{aligned} A_{\text{diag}} &= \mathbb{H}_0 + N^{-1} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) b^*(k) b(k) + \frac{1}{2N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( \mathfrak{D}^*(k) \mathfrak{D}(k) + \mathfrak{D}^*(-k) \mathfrak{D}(-k) \right) \\ &\quad + \frac{1}{2N} \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \left( 3 \sum_{h \in B_F \cap (B_F + k)} a_h^* a_h - 2 \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_p - \sum_{p \in B_F^c \cap (B_F^c + k)} a_p^* a_p \right) + C_0 N^{-1/3} \end{aligned}$$

with

$$b^*(k) := \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}, \quad \mathfrak{D}^*(k) := \sum_{p \in B_F^c \cap (B_F^c + k)} a_p^* a_{p-k} - \sum_{p \in B_F \cap (B_F + k)} a_p^* a_{p-k}.$$

Let us estimate  $A_{\text{diag}}$  from above. Clearly

$$\pm \left( 3 \sum_{h \in B_F \cap (B_F + k)} a_h^* a_h - 2 \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_p - \sum_{p \in B_F^c \cap (B_F^c + k)} a_p^* a_p \right) \leq 3 \sum_{p \in \mathbb{Z}^3} a_p^* a_p = 3\mathcal{N}.$$

Moreover, for every  $|k| \sim O(1)$  by the Cauchy–Schwarz inequality we have

$$\begin{aligned} b^*(k) b(k) &= \sum_{p, q \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}^* a_{q-k} a_q \\ &\leq \sum_{p, q \in B_F^c \cap (B_F + k)} \frac{1}{2} \left( a_p^* a_{p-k}^* a_{p-k} a_p + a_q^* a_{q-k}^* a_{q-k} a_q \right) \\ &= |B_F^c \cap (B_F + k)| \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}^* a_{p-k} a_p \\ &\leq C N^{2/3} \sum_{p \in \mathbb{Z}^3} a_p^* a_p = C N^{2/3} \mathcal{N} \end{aligned}$$

and

$$\begin{aligned}
\mathfrak{D}^*(k)\mathfrak{D}(k) &\leq 2 \sum_{p,q \in B_F^c \cap (B_F^c + k)} a_p^* a_{p-k} a_{q-k}^* a_q + 2 \sum_{p,q \in B_F \cap (B_F + k)} a_p^* a_{p-k} a_{q-k}^* a_q \\
&\leq 4 \sum_{p,q \in \mathbb{Z}^3} a_p^* a_{p-k} a_{q-k}^* a_q = 4 \sum_{p,q \in \mathbb{Z}^3} a_p^* (\delta_{pq} - a_{q-k}^* a_{p-k}) a_q \\
&= 4\mathcal{N} + 4 \sum_{p,q \in \mathbb{Z}^3} a_p^* a_{q-k}^* a_q a_{p-k} \\
&\leq 4\mathcal{N} + 2 \sum_{p,q \in \mathbb{Z}^3} (a_p^* a_{q-k}^* a_{q-k} a_p + a_{p-k}^* a_q^* a_q a_{p-k}) \\
&= 4\mathcal{N} + 2 \sum_{p \in \mathbb{Z}^3} (a_p^* \mathcal{N} a_p + a_{p-k}^* \mathcal{N} a_{p-k}) \\
&= 4\mathcal{N} + 2 \sum_{p \in \mathbb{Z}^3} (a_p^* a_p (\mathcal{N} - 1) + a_{p-k}^* a_{p-k} (\mathcal{N} - 1)) \\
&= 4\mathcal{N} + 4\mathcal{N}(\mathcal{N} - 1) = 4\mathcal{N}^2.
\end{aligned}$$

Thus in summary,

$$A_{\text{diag}} \leq \mathbb{H}_0 + 4N^{-1}\mathcal{N}^2 + CN^{-1/3}\mathcal{N} + CN^{-1/3}.$$

Hence, for every  $M \sim N^{1/3}$  the abstract localization lemma gives

$$\pm \left( A - f_M A f_M - g_M A g_M \right) \leq \frac{C}{M^2} (\mathbb{H}_0 + C).$$

which is equivalent to

$$\pm \left( R H_N R^* - f_M R H_N R^* f_M - g_M R H_N R^* g_M \right) \leq \frac{C}{M^2} (\mathbb{H}_0 + C).$$

**Step 2.** Now let  $\Psi_{\text{gs}} \in L_a^2(\mathbb{T}^{3N})$  be a ground state for  $H_N$  and denote  $\psi_{\text{gs}} = R\Psi_{\text{gs}}$ . Then we know that

$$\langle \psi_{\text{gs}}, \mathbb{H}_0 \psi_{\text{gs}} \rangle \leq CN^{-1/3}, \quad \langle \psi_{\text{gs}}, \mathcal{N} \psi_{\text{gs}} \rangle \leq CN^{1/3}.$$

Hence, with  $M \sim N^{1/3}$  we have

$$\left\langle \psi_{\text{gs}}, \left( R H_N R^* - f_M R H_N R^* f_M - g_M R H_N R^* g_M \right) \psi_{\text{gs}} \right\rangle \geq -CN^{-2/3}.$$

Note that

$$\langle \psi_{\text{gs}}, R H_N R^* \psi_{\text{gs}} \rangle = \langle \Psi_{\text{gs}}, H_N \Psi_{\text{gs}} \rangle = E_N$$

and

$$\langle \psi_{\text{gs}}, g_M R H_N R^* g_M \psi_{\text{gs}} \rangle \geq E_N \|g_M \psi_{\text{gs}}\|^2 = E_N (1 - \|f_M \psi_{\text{gs}}\|^2).$$

Thus we conclude that

$$\langle \psi_{\text{gs}}, f_M R H_N R^* f_M \psi_{\text{gs}} \rangle \leq E_N \|f_M \psi_{\text{gs}}\|^2 + C N^{-2/3}.$$

Note that we can choose  $M = C N^{1/3} \geq 4 \langle \psi_{\text{gs}}, \mathcal{N} \psi_{\text{gs}} \rangle$ , which ensures that

$$\|g_M \psi_{\text{gs}}\|^2 = \langle \psi_{\text{gs}}, g_M^2 \psi_{\text{gs}} \rangle \leq \langle \psi_{\text{gs}}, \frac{2\mathcal{N}}{M} \psi_{\text{gs}} \rangle \leq \frac{1}{2},$$

namely

$$\|f_M \psi_{\text{gs}}\|^2 = 1 - \|g_M \psi_{\text{gs}}\|^2 \geq \frac{1}{2}.$$

Finally, we define

$$\psi = \frac{f_M \psi_{\text{gs}}}{\|f_M \psi_{\text{gs}}\|}, \quad \Psi = R\psi.$$

Then we have

$$\langle \Psi, H_N \Psi \rangle = \langle \psi, R H_N R^* \psi \rangle \leq E_N + C N^{-2/3}.$$

Moreover,  $\psi = \mathbf{1}(\mathcal{N} \leq M) \psi_{\text{gs}} = \mathbf{1}(\mathcal{N} \leq C N^{1/3}) \psi_{\text{gs}}$  and

$$\langle \psi, \mathbb{H}_0 \psi \rangle \leq 2 \langle \psi_{\text{gs}}, f_M \mathbb{H}_0 f_M \psi_{\text{gs}} \rangle \leq 2 \langle \psi_{\text{gs}}, \mathbb{H}_0 \psi_{\text{gs}} \rangle \leq C N^{-1/3}.$$

□

## 10.4 Removing the non-bosonizable term

In this section we estimate the error term

$$\begin{aligned} \mathbb{X} &= \frac{1}{2N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( \mathfrak{D}^*(k) \mathfrak{D}(k) + \mathfrak{D}^*(-k) \mathfrak{D}(-k) \right) \\ &+ \frac{1}{2N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( \mathfrak{D}^*(-k) b(k) + \mathfrak{D}^*(k) b(-k) + h.c. \right) \\ &+ \frac{1}{2N} \sum_{k \in \mathbb{Z}^3} \widehat{V}(k) \left( 3 \sum_{h \in B_F \cap (B_F + k)} a_h^* a_h - 2 \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_p - \sum_{p \in B_F^c \cap (B_F^c + k)} a_p^* a_p \right) \end{aligned}$$

where

$$b^*(k) := \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}^*, \quad \mathfrak{D}^*(k) := \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k} - \sum_{p \in B_F \cap (B_F + k)} a_p^* a_{p-k}.$$

**Lemma.** *Let  $\psi = R\Psi$  with  $\Psi \in L_a^2(\mathbb{T}^3)$  and*

$$\psi = \mathbf{1}(\mathcal{N} \leq CN^{1/3})\psi, \quad \langle \psi, \mathbb{H}_0 \psi \rangle \leq CN^{-1/3}.$$

*Then we have*

$$\pm \langle \psi, \mathbb{X} \psi \rangle \leq o(N^{-1/3})$$

*Proof. Step 1.* The third sum in  $\mathbb{X}$  is easy since

$$\pm N^{-1} \left( 3 \sum_{h \in B_F \cap (B_F + k)} a_h^* a_h - 2 \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_p - \sum_{p \in B_F \cap (B_F + k)} a_p^* a_p \right) \leq 3N^{-1} \mathcal{N}$$

and

$$N^{-1} \langle \psi, \mathcal{N} \psi \rangle \leq O(N^{-2/3}).$$

**Step 2.** Now let us prove that for every  $0 \neq k \in \mathbb{Z}^3$ ,  $|k| \sim O(1)$ ,

$$N^{-1} \langle \psi, \mathfrak{D}^*(k) \mathfrak{D}(k) \psi \rangle \leq o(N^{-1/3}) \iff \langle \psi, \mathfrak{D}^*(k) \mathfrak{D}(k) \psi \rangle \leq o(N^{2/3}).$$

Note that we have proved  $\mathfrak{D}^*(k) \mathfrak{D}(k) \leq 4\mathcal{N}^2$ . However, from that bound and the a-priori estimate on  $\mathcal{N}$  we only have

$$\langle \psi, \mathfrak{D}^*(k) \mathfrak{D}(k) \psi \rangle \leq 4 \langle \psi, \mathcal{N}^2 \psi \rangle \leq CN^{2/3}.$$

To obtain the refinement  $o(N^{2/3})$  we need a better approach.

By the Cauchy–Schwarz inequality  $(A+B)^*(A+B) \leq 2(A^*A + B^*B)$  we have

$$\begin{aligned} \langle \psi, \mathfrak{D}^*(k) \mathfrak{D}(k) \psi \rangle &\leq 2 \sum_{p, q \in \mathbb{Z}^3} \langle \psi, a_p^* a_{p-k} a_{q-k}^* a_q \psi \rangle \\ &= 2 \sum_{p, q \in \mathbb{Z}^3} \langle \psi, a_p^* (\delta_{pq} - a_{q-k}^* a_{p-k}) a_q \psi \rangle \\ &= 2 \langle \psi, \mathcal{N} \psi \rangle - 2 \sum_{p, q \in \mathbb{Z}^3} \langle \psi, a_p^* a_{q-k}^* a_{p-k} a_q \psi \rangle. \end{aligned}$$

The first term is bounded by  $O(N^{1/3})$ . For the second term, by the Cauchy–Schwarz inequality again we can bound

$$\begin{aligned}
& \left| \sum_{p,q \in \mathbb{Z}^3} \langle \psi, a_p^* a_{q-k}^* a_{p-k} a_q \psi \rangle \right| \\
& \leq \sum_{p,q \in \mathbb{Z}^3} \|a_{q-k} a_p \psi\| \|a_{p-k} a_q \psi\| \\
& \leq \sum_{p,q \in \mathbb{Z}^3} (A_p \|a_{q-k} a_p \psi\|^2 + B_p \|a_{p-k} a_q \psi\|^2) \\
& \leq \sum_{p \in \mathbb{Z}^3} (A_p \|a_p (\mathcal{N} + 1)^{1/2} \psi\|^2 + B_p \|a_{p-k} (\mathcal{N} + 1)^{1/2} \psi\|^2)
\end{aligned}$$

provided that

$$A_p > 0, \quad B_p > 0, \quad 2\sqrt{A_p B_p} \geq 1.$$

Combining with the condition  $\psi = \mathbf{1}(\mathcal{N} \leq CN^{1/3})\psi$  we have

$$\left| \sum_{\mathbb{Z}^3} \langle \psi, a_p^* a_{q-k}^* a_{p-k} a_q \psi \rangle \right| \leq CN^{1/3} \sum_{p \in \mathbb{Z}^3} (A_p \|a_p \psi\|^2 + B_p \|a_{p-k} \psi\|^2)$$

To estimate further the right side, we split the sum into two parts:

$$\begin{aligned}
X_1 &= \{p \in \mathbb{Z}^3 : \max\{|p^2 - c_0|, |(p-k)^2 - c_0|\} \geq \ell\}, \\
X_2 &= \{p \in \mathbb{Z}^3 : \max\{|p^2 - c_0|, |(p-k)^2 - c_0|\} < \ell\}
\end{aligned}$$

with a large parameter  $\ell$  (eventually we can take  $\ell = N^{2/15}$ ). Here the constant  $c_0$  is taken from the representation of the kinetic operator

$$\mathbb{H}_0 \psi = \sum_{k \in \mathbb{Z}^3} |k^2 - c_0| a_k^* a_k \psi \quad \text{and} \quad \inf_{k \in \mathbb{Z}^3} |k^2 - c_0| \geq \frac{1}{2}.$$

**Part 1:** For  $p \in X_1$  we choose

$$A_p = \frac{|p^2 - c_0|}{\ell^{1/2}}, \quad B_p = \frac{|(p-k)^2 - c_0|}{\ell^{1/2}}$$

Then  $2\sqrt{A_p B_p} > 1$  because

$$\min\{|p^2 - c_0|, |(p-k)^2 - c_0|\} \geq 1/2$$

and

$$\max\{|p^2 - c_0|, |(p - k)^2 - c_0|\} \geq \ell.$$

Thus using the kinetic bound

$$\langle \psi, \mathbb{H}_0 \psi \rangle = h^2 \sum_{p \in \mathbb{Z}^3} |p^2 - c_0| \langle \psi a_p^* a_p \psi \rangle \leq CN^{-1/3}$$

we have

$$\sum_{p \in X_1} (A_p \|a_p \psi\|^2 + B_p \|a_{p-k} \psi\|^2) \leq \frac{1}{\ell^{1/2}} \sum_{p \in \mathbb{Z}^3} |p^2 - c_0| \langle \psi a_p^* a_p \psi \rangle \leq \frac{CN^{1/3}}{\ell^{1/2}}.$$

**Part 2:** For  $p \in X_2$  we simply choose  $A_p = B_p = 1$ . Using the fermionic property  $\|a_p\|_{\text{op}} \leq 1$  we obtain

$$\sum_{p \in X_2} (A_p \|a_p \psi\|^2 + B_p \|a_{p-k} \psi\|^2) \leq 2|X_2|$$

We can show that (see below)

$$|X_2| \leq C\ell^2 N^\varepsilon, \quad \forall \varepsilon > 0.$$

Thus in conclusion,

$$\left| \sum_{p, q \in \mathbb{Z}^3} \langle \psi, a_p^* a_{q-k}^* a_{p-k} a_q \psi \rangle \right| \leq CN^{1/3} \left( \frac{N^{1/3}}{\ell^{1/2}} + \ell^2 N^\varepsilon \right).$$

By optimizing over  $\ell$  (e.g. taking  $\ell \sim N^{2/15}$ ) we find that

$$\left| \sum_{p, q \in \mathbb{Z}^3} \langle \psi, a_p^* a_{q-k}^* a_{p-k} a_q \psi \rangle \right| \leq o(N^{2/3}).$$

Thus we get the desired estimate

$$\langle \psi, \mathfrak{D}^*(k) \mathfrak{D}(k) \psi \rangle \leq o(N^{2/3}).$$

**Counting problem for  $|X_2|$ .** We prove that

$$|X_2| \leq C\ell^2 N^\varepsilon, \quad \forall \varepsilon > 0.$$

*Proof.* Note that for  $p \in X_2$  by the triangle inequality

$$|2p \cdot k| = |p^2 - (p - k)^2 + k^2| \leq |p^2 - c_0| + |c_0 - (p - k)^2| + k^2 \leq 2\ell + k^2 \leq 4\ell.$$

Hence,

$$X_2 \subset \{p \in \mathbb{Z}^3 : |p^2 - c_0| \leq \ell, |p \cdot k| \leq 2\ell\}.$$

The desired inequality  $|X_2| \leq C\ell^2 N^\varepsilon$  follows from the fact that for every given  $r, s \in \mathbb{Z}$ ,  $|r - c_0| \leq \ell$ ,  $|s| \leq 2\ell$ , we have

$$\left| \{p \in \mathbb{Z}^3 : p^2 = r, p \cdot k = s\} \right| \leq CN^\varepsilon, \quad \forall \varepsilon > 0.$$

We will need a fundamental fact from number theory.

**Theorem** (Integer points on ellipses). *Let  $d_0 \in \mathbb{N}$ . Then when  $M \rightarrow \infty$ , the equation*

$$m^2 + d_0 n^2 = M$$

*has at most  $O(M^\varepsilon)$  solutions  $(m, n) \in \mathbb{Z}^2$ , for every  $\varepsilon > 0$ .*

This result was proved by **Cilleruelo** and **Córdoba** [Lattice points on ellipses, Duke Math. J., 1994]. Let us accept it and conclude the counting argument.

**Easy case:** Assume  $k = (k_1, 0, 0)$  with  $k_1 \neq 0$ . Then condition  $p \cdot k = s$  determines  $p_1$  uniquely and for every given  $p_1$  the equation

$$p_2^2 + p_3^2 = r - p_1^2$$

as at most  $O(N^\varepsilon)$  solutions  $(p_2, p_3) \in \mathbb{Z}^2$  for every  $\varepsilon > 0$ . Here we used  $|r - p_1^2| \leq CN^{2/3}$ .

**General case:** Now we turn to the general case. We can assume that  $k = (k_1, k_2, k_3)$  with  $k_1 \neq 0$  and  $(k_2, k_3) \neq (0, 0)$ . Then we use the fact that the following vectors are orthogonal in  $\mathbb{R}^3$ :

$$k = (k_1, k_2, k_3), \quad k_\perp = (0, -k_3, k_2), \quad k'_\perp = (-k_2^2 - k_3^2, k_1 k_2, k_1 k_3).$$



Using the orthogonality of  $(k, k_\perp, k'_\perp)$  we can write

$$|p|^2 = \left| p \cdot \frac{k}{|k|} \right|^2 + \left| p \cdot \frac{k_\perp}{|k_\perp|} \right|^2 + \left| p \cdot \frac{k'_\perp}{|k'_\perp|} \right|^2.$$

Combining with  $|k'_\perp| = |k||k_\perp|$ , we obtain

$$|k'_\perp|^2 |p|^2 = |k_\perp|^2 |p \cdot k|^2 + |k|^2 |p \cdot k_\perp|^2 + |p \cdot k'_\perp|^2.$$

Every  $p \in \mathbb{Z}^3$  is determined uniquely by  $(n_1, n_2, n_3) \in \mathbb{Z}^3$  with

$$n_1 = p \cdot k, \quad n_2 = p \cdot k_\perp, \quad n_3 = p \cdot k'_\perp.$$

In particular, the constraints

$$k \cdot p = s, \quad p^2 = r$$

can be rewritten as

$$n_1 = s, \quad |k|^2 n_2^2 + n_3^2 = |k'_\perp|^2 r - |k_\perp|^2 s^2.$$

For every given  $r, s \in O(k_F)$ , the second equation has at most  $O(N^\varepsilon)$  solutions  $(n_2, n_3) \in \mathbb{Z}^2$ . This completes the proof of  $|X_2| \leq C\ell^2 N^\varepsilon$ . Thus Step 2 is finished.  $\square$

**On the counting problem on ellipses.** Here is a proof of the counting problem on **circles**.

*Proof.* We need to prove that for  $M \rightarrow \infty$  the equation

$$x^2 + y^2 = M$$

has at most  $O(M^\varepsilon)$  solutions  $(x, y) \in \mathbb{Z}^2$ . In this case, the number of solutions, denoted by  $r_2(M)$ , can be computed explicitly. For every  $M \in \mathbb{N}$  we can write uniquely

$$M = 2^\alpha \cdot m \cdot \prod_p p^{\beta_p}$$

where the last product is taken over prime numbers  $p \equiv 1 \pmod{4}$ , and  $m$  is the product of (powers of) primes  $\equiv 3 \pmod{4}$ . Then we have Jacobi's formula

$$r_2(M) = \begin{cases} 0, & \text{if } m \text{ is not a square} \\ 4 \prod_p (1 + \beta_p), & \text{if } m \text{ is a square.} \end{cases}$$

Consequently, we get the upper bound

$$r_2(M) \leq 4 \prod_p (1 + \beta_p)$$

over prime numbers  $p \equiv 3 \pmod{4}$  which are factors of  $M$ . We can divide the product into two parts.

**Part 1.** If  $p \leq K$  for a large constant, then we simply bound  $\beta_p$  by  $\log M$  and get

$$\prod_{p \leq K} (1 + \beta_p) \leq (C \log M)^K.$$

**Part 2.** If  $p \geq K$ , then using

$$1 + \beta_p \leq e^{\beta_p} \leq K^{\beta_p / \log K} \leq p^{\beta_p / \log K}$$

we find that

$$\prod_{p > K} (1 + \beta_p) \leq \prod_{p > K} p^{\beta_p / \log K} \leq M^{1 / \log K}.$$

In summary, we obtain

$$r_2(M) \leq (C \log M)^K M^{1 / \log K}$$

for any  $K$  large. This implies that  $r_2(M) \leq O(M^\varepsilon)$  for any  $\varepsilon > 0$ .  $\square$

**Step 3.** For the second term of  $\mathbb{X}$ , we use the Cauchy–Schwarz inequality

$$\pm (\mathfrak{D}^*(k)b(-k) + h.c.) \leq \varepsilon^{-1} \mathfrak{D}^*(k)\mathfrak{D}(k) + \varepsilon b^*(k)b(k).$$

We need to prove that the expectation against  $\psi$  is  $o(N^{2/3})$ . We have proved that

$$\langle \psi, \mathfrak{D}^*(k)\mathfrak{D}(k)\psi \rangle \leq o(N^{2/3}).$$

Therefore, it suffices to show that

$$\langle \psi, b^*(k)b(k)\psi \rangle \leq CN^{2/3}$$

and optimize over  $\varepsilon > 0$ . The latter bound follows from the kinetic bound

$$\langle \psi, \mathbb{H}_0\psi \rangle \leq CN^{-1/3}$$

and the following lemma.

**Lemma.** *For every  $0 \neq k \in \mathbb{Z}^3$ ,  $|k| \sim O(1)$  we have*

$$N^{-1} \langle \psi, b^*(k)b(k)\psi \rangle \leq C \langle \psi, \mathbb{H}_0\psi \rangle.$$

This inequality was first proved by **Hainzl–Porta–Rexze** (2018). Below is a simplified proof. Recall that

$$b^*(k) = \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}^*, \quad \mathbb{H}_0\psi = \sum_{p \in \mathbb{Z}^3} h^2 |p^2 - c_0| a_p^* a_p \psi.$$

*Proof of the lemma.* As in [HPR], By the triangle and Cauchy–Schwarz inequalities we can write

$$\begin{aligned} \|b(k)\psi\|^2 &\leq \left( \sum_{p \in B_F^c \cap (B_F + k)} \|a_{p-k} a_p \psi\| \right)^2 \\ &\leq \left( \sum_{p \in B_F^c \cap (B_F + k)} |p^2 - (p-k)^2| \|a_{p-k} a_p \psi\|^2 \right) \left( \sum_{p \in B_F^c \cap (B_F + k)} \frac{1}{|p^2 - (p-k)^2|} \right). \end{aligned}$$

The first term can be bounded by the kinetic energy

$$\begin{aligned} &\sum_{p \in B_F^c \cap (B_F + k)} |p^2 - (p-k)^2| \|a_{p-k} a_p \psi\|^2 \\ &= \sum_{p \in B_F^c \cap (B_F + k)} |(p-k)^2 - c_0| \|a_{p-k} a_p \psi\|^2 + \sum_{p \in B_F^c \cap (B_F + k)} |c_0 - p^2| \|a_{p-k} a_p \psi\|^2 \\ &\leq \sum_{p \in B_F^c \cap (B_F + k)} |(p-k)^2 - c_0| \|a_{p-k} \psi\|^2 + \sum_{p \in B_F^c \cap (B_F + k)} |c_0 - p^2| \|a_p \psi\|^2 \\ &\leq \sum_{p \in \mathbb{Z}^3} |p^2 - c_0| \|a_p \psi\|^2 = N^{2/3} \langle \psi, \mathbb{H}_0\psi \rangle. \end{aligned}$$

Thus it remains to show that

$$\sum_{p \in B_F^c \cap (B_F + k)} \frac{1}{|p^2 - (p-k)^2|} \leq CN^{1/3}.$$

Here the sum is taken over  $O(N^{2/3})$  terms. Note that if  $p \in B_F^c \cap (B_F + k)$ , then  $|p| \sim N^{1/3}$ ,

and hence

$$|(p-k)^2 - p^2| = 2|p \cdot k| + O(1) \leq CN^{1/3}.$$

If this term were  $\geq N^{1/3}$  for all  $p \in B_F^c \cap (B_F + k)$ , then we are done. However,  $|p \cdot k|$  may be small (e.g.  $O(1)$ ) and we have to count these terms carefully. We write

$$\sum_{p \in B_F^c \cap (B_F + k)} \frac{1}{|p^2 - (p-k)^2|} \leq \sum_{|s| \leq CN^{1/3}} \frac{C|B_s|}{|s|+1}$$

where

$$B_s = \{p \in B_F^c \cap (B_F + k) : p \cdot k = s\}.$$

We count  $|B_s|$  using a similar idea of counting  $|X_2|$  in Step 2.

**First try.** We can write

$$B_s = \bigcup_r B_{s,r}, \quad B_{s,r} = \{p \in B_F^c \cap (B_F + k) : p \cdot k = s, p^2 = r\}.$$

Here the condition  $p \in B_F^c \cap (B_F + k)$  means

$$p^2 \geq k_F^2 > (p-k)^2.$$

When  $p \cdot k = s$ ,  $p^2 = r$ , it is equivalent to

$$k_F^2 + 2p \cdot k - k^2 > p^2 \geq k_F^2 \iff k_F^2 + 2s - k^2 > r \geq k_F^2.$$

Thus for any given  $s$ , we have at most  $C(|s|+1)$  choices of  $r$ . Moreover, for given  $(s, r)$ , we have  $|B_{s,r}| \leq O(N^\varepsilon)$  for every  $\varepsilon > 0$ , by the same argument as in counting  $|X_2|$ . Thus we conclude that

$$|B_s| \leq C_\varepsilon N^\varepsilon (|s|+1), \quad \forall \varepsilon > 0.$$

From this bound we can obtain

$$\sum_{p \in B_F^c \cap (B_F + k)} \frac{1}{|p^2 - (p-k)^2|} \leq \sum_{|s| \leq CN^{1/3}} \frac{C|B_s|}{|s|+1} \leq C_\varepsilon N^{1/3+\varepsilon}, \quad \forall \varepsilon > 0$$

which is close to the desired bound  $O(N^{1/3})$ , but not enough.

**Second try.** Let us proceed differently, using the following

**Theorem** (Integer points in ellipses). *Let  $d_0 \in \mathbb{N}$ . For every  $R > 0$  consider the ellipse*

$$E(R) = \{(x, y) \in \mathbb{R}^2 : x^2 + d_0 y^2 \leq R^2\}.$$

*Then the number of integer points in  $E(R)$  is*

$$S(R) = |E(R)| + O(R^{2/3})_{R \rightarrow \infty}.$$

When  $d_0 = 1$ , this is the bound from the Gauss circle problem that we discussed before. The result for ellipses is slightly more difficult, but could be obtained by a similar argument (in general, the bound  $O(R^{2/3})$  holds for any convex set, and it is optimal).

Now let us show that

$$|B_s| \leq C(|s| + N^{2/9}).$$

**Easy case:** Assume  $k = (k_1, 0, 0)$  with  $k_1 \neq 0$ . Then condition  $p \cdot k = s$  determines  $p_1$  uniquely. Recall that the condition  $p \in B_F^c \cap (B_F + k)$  means

$$k_F^2 + 2p \cdot k - k^2 > p^2 \geq k_F^2 \iff k_F^2 + 2s - k^2 - p_1^2 > p_2^2 + p_3^2 \geq k_F^2 - p_1^2.$$

The number of integer solutions  $(p_2, p_3) \in \mathbb{Z}^2$  is equal to the integer points in the annulus  $B(0, R_2) \setminus B(0, R_1)$  with

$$R_1 = \sqrt{k_F^2 - p_1^2}, \quad R_2 = \sqrt{k_F^2 + 2s - k^2 - p_1^2}.$$

Note that

$$R_1 < R_2 \leq CN^{1/3}, \quad R_2^2 - R_1^2 \leq C(|s| + 1).$$

By Gauss circle problem, the number of integer solutions in the annulus  $B(0, R_2) \setminus B(0, R_1)$  is

$$|B(0, R_2)| - |B(0, R_1)| + O(R_2^{2/3}) = \pi(R_2^2 - R_1^2) + O(R_2^{2/3}) \leq C(|s| + N^{2/9}).$$

**General case:** Assume that  $k = (k_1, k_2, k_3)$  with  $k_1 \neq 0$  and  $(k_2, k_3) \neq (0, 0)$ . Using the orthogonal vectors in  $\mathbb{R}^3$ :

$$k = (k_1, k_2, k_3), \quad k_\perp = (0, -k_3, k_2), \quad k'_\perp = (-k_2^2 - k_3^2, k_1 k_2, k_1 k_3).$$

we can write

$$|k'_\perp|^2 |p|^2 = |k_\perp|^2 |p \cdot k|^2 + |k|^2 |p \cdot k_\perp|^2 + |p \cdot k'_\perp|^2.$$

Moreover, every  $p \in \mathbb{Z}^3$  is determined uniquely by  $(n_1, n_2, n_3) \in \mathbb{Z}^3$  with

$$n_1 = p \cdot k, \quad n_2 = p \cdot k_\perp, \quad n_3 = p \cdot k'_\perp.$$

From the conditions  $p \in B_F^c \cap (B_F + k)$  and  $n_1 = k \cdot p = s$  we have

$$k_F^2 + 2p \cdot k - k^2 > p^2 \geq k_F^2 \iff |k'_\perp|^2(k_F^2 + 2s - k^2) > |k_\perp|^2 s^2 + |k|^2 n_2^2 + n_3^2 \geq |k'_\perp|^2 k_F^2.$$

Thus  $(n_2, n_3) \in E(R_2) \setminus E(R_1)$  where  $E(R)$  is the ellipse

$$E(R) = \{(x, y) \in \mathbb{R}^2 : |k|^2 x^2 + y^2 \leq R^2\}$$

and

$$R_1 = \sqrt{|k'_\perp|^2 k_F^2 - |k_\perp|^2 s^2}, \quad R_2 = \sqrt{|k'_\perp|^2 (k_F^2 + 2s - k^2) - |k_\perp|^2 s^2}.$$

Note that  $R_1 < R_2 \leq CN^{1/3}$  and  $R_2^2 - R_1^2 \leq C(|s| + 1)$ . Hence, by the Gauss counting problem on ellipses, the number of integer points  $(n_2, n_3) \in \mathbb{Z}^2$  in  $E(R_2) \setminus E(R_1)$  is

$$|E(R_2)| - |E(R_1)| + O(R_2^{2/3}) = \pi \frac{R_2^2}{|k|} - \pi \frac{R_1^2}{|k|} + O(N^{2/9}) \leq C(|s| + N^{2/9}).$$

Thus in conclusion, we have proved for all  $|k| \sim O(1)$ ,

$$|B_s| \leq C \min\{N^\varepsilon(|s| + 1), |s| + N^{2/9}\}.$$

Therefore,

$$\begin{aligned} \sum_{p \in B_F^c \cap (B_F + k)} \frac{1}{|p^2 - (p - k)^2|} &\leq \sum_{|s| \leq CN^{1/3}} \frac{C|B_s|}{|s| + 1} \\ &\leq \sum_{|s| \leq N^{2/9}} \frac{CN^\varepsilon(|s| + 1)}{|s| + 1} + \sum_{|s| = N^{2/9}}^{CN^{1/3}} \frac{C(|s| + 1)}{|s| + 1} \\ &= CN^{2/9+\varepsilon} + CN^{1/3}. \end{aligned}$$

This completes the proof of the bound

$$N^{-1} \langle \psi, b^*(k)b(k)\psi \rangle \leq C \langle \psi, \mathbb{H}_0 \psi \rangle.$$

□

Thus we have estimated the non-bosonizable part  $\pm\langle\psi, \mathbb{X}\psi\rangle \leq o(N^{-1/3})$ .  $\square$

## 10.5 Diagonalization of bosonizable term

Now we conclude the proof of

**Theorem.** *If  $\widehat{V} \geq 0$ , compactly supported and small enough, then*

$$E_N = E^{\text{HF}} + E^{\text{corr}} + o(N^{-1/3}).$$

Here the correlation energy is given by  $E^{\text{corr}} = N^{-1/3} \sum_k g(k)$  with

$$g(k) = -\frac{\pi\kappa^2}{2}\widehat{V}(k) + \frac{\kappa}{\pi} \int_0^\infty \log \left[ 1 + 2\pi\kappa\widehat{V}(k)(1 - \arctan^{-1}(\lambda^{-1})) \right] d\lambda, \quad \kappa = \left( \frac{3}{4\pi} \right)^{1/3}.$$

The discussion in this section is only heuristic. We consider the bosonizable term

$$\mathbb{H}_0 + \mathbb{Q} = \sum_{p \in \mathbb{Z}^3} h^2 |p^2 - k_F^2| a_p^* a_p + \frac{1}{N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( b^*(k)b(k) + b^*(-k)b(-k) + b^*(k)b^*(-k) + b(-k)b(k) \right).$$

Recall that

$$b^*(k) = \sum_{p \in B_F^c \cap (B_F + k)} a_p^* a_{p-k}^*.$$

For any given  $k$  and  $p \in B_F^c \cap (B_F + k)$ , we think of the operator

$$b_p^*(k) := a_p^* a_{p-k}^*$$

as a **bosonic creation operator**. The reason is that it satisfies

$$[b_p^*(k), b_q^*(k)] = a_p^* a_{p-k}^* a_q^* a_{q-k}^* - a_q^* a_{q-k}^* a_p^* a_{p-k}^* = 0, \quad [b_p(k), b_q(k)] = 0$$

and

$$\begin{aligned} [b_p(k), b_q^*(k)] &= a_{p-k} a_p a_q^* a_{q-k}^* - a_q^* a_{q-k}^* a_{p-k} a_p \\ &= a_{p-k} (\delta_{pq} - a_q^* a_p) a_{q-k}^* - a_q^* (\delta_{pq} - a_{p-k} a_{q-k}^*) a_p \\ &= \delta_{pq} (a_{p-k} a_{q-k}^* - a_q^* a_p) - a_{p-k} a_q^* a_p a_{q-k}^* + a_q^* a_{p-k} a_{q-k}^* a_p \\ &= \delta_{pq} (a_{p-k} a_{p-k}^* - a_p^* a_p) \end{aligned}$$

$$\begin{aligned}
&= \delta_{pq}(1 - a_{p-k}^* a_{p-k} - a_p^* a_p) \\
&= \delta_{pq} - \delta_{pq}(a_{p-k}^* a_{p-k} + a_p^* a_p).
\end{aligned}$$

These relations look similar to **approximate CCR** for bosonic operator. The error term  $\delta_{pq}(a_{p-k}^* a_{p-k} - a_p^* a_p)$  in  $[b_p(k), b_q^*(k)]$  is not identically equal to 0, but it is small in average since

$$\sum_{p \in B_F^c \cap (B_F + k)} \langle \psi, a_p^* a_p \psi \rangle \leq \langle \psi, \mathcal{N} \psi \rangle \leq CN^{1/3} \ll N^{2/3}.$$

Similarly, we can show that if  $k \neq \ell$ , then

$$[b_p^*(k), b_q^*(\ell)] = 0, \quad [b_p(k), b_q(\ell)] = 0, \quad [b_p(k), b_q^*(\ell)] \approx 0.$$

This means that the different momenta  $k$  correspond to different Fock spaces. This is consistent with the **random phase approximation** developed by **Bohm–Pines** (1960s).

Thus the interaction term

$$\begin{aligned}
\mathbb{Q} &= \frac{1}{N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \left( b^*(k)b(k) + b^*(-k)b(-k) + b^*(k)b^*(-k) + b(-k)b(k) \right) \\
&= \frac{1}{N} \sum_{k \in \Gamma^{\text{nor}}} \widehat{V}(k) \sum_{p, q \in B_F^c \cap (B_F + k)} \left( b_p^*(k)b_q(k) + b_p^*(-k)b_q(-k) + b_p^*(k)b_q^*(-k) + b_p(-k)b_q(k) \right)
\end{aligned}$$

looks like a quadratic Hamiltonian in a bosonic Fock space.

It is somewhat less obvious that the kinetic operator is also quadratic in terms of  $b_p^*(k)$  and  $b_p(k)$ . Heuristically,

$$\mathbb{H}_0 = \sum_{p \in \mathbb{Z}^3} h^2 |p^2 - k_F^2| a_p^* a_p \approx \sum_{k \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + k)} h^2 (p^2 - (p-k)^2) b_p^*(k) b_p(k) =: \widetilde{\mathbb{H}}_0.$$

Indeed, it does not hold in general, but it holds for a class of quantum state close to the ground state. Our key observation is that

$$\begin{aligned}
[\mathbb{H}_0, b_q^*(k)] &= \sum_{p \in \mathbb{Z}^3} \sum_{q \in B_F^c \cap (B_F + k)} h^2 |p^2 - k_F^2| [a_p^* a_p, a_q^* a_{q-k}^*] \\
&= \sum_{p \in \mathbb{Z}^3} \sum_{q \in B_F^c \cap (B_F + k)} h^2 |p^2 - k_F^2| \left( [a_p^* a_p, a_q^*] a_{q-k}^* + a_q^* [a_p^* a_p, a_{q-k}^*] \right) \\
&= \sum_{p \in \mathbb{Z}^3} \sum_{q \in B_F^c \cap (B_F + k)} h^2 |p^2 - k_F^2| \left( \delta_{pq} a_q^* a_{q-k}^* + \delta_{p, q-k} a_q^* a_{q-k}^* \right)
\end{aligned}$$



$$\begin{aligned}
&= \sum_{q \in B_F^c \cap (B_F + k)} h^2(|q^2 - k_F^2| + |(q - k)^2 - k_F^2|) a_q^* a_{q-k}^* \\
&= h^2(q^2 - (q - k)^2) b_q^*(k)
\end{aligned}$$

while

$$\begin{aligned}
[\widetilde{\mathbb{H}}_0, b_q^*(k)] &= \sum_{\ell \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + \ell)} \sum_{q \in B_F^c \cap (B_F + k)} h^2(p^2 - (p - \ell)^2) [b_p^*(\ell) b_p(\ell), b_q^*(k)] \\
&\approx \sum_{\ell \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + \ell)} \sum_{q \in B_F^c \cap (B_F + k)} h^2(p^2 - (p - \ell)^2) \delta_{k, \ell} \delta_{p, q} b_q^*(k) \\
&= h^2(q^2 - (q - k)^2) b_q^*(k).
\end{aligned}$$

Thus although  $\mathbb{H}_0$  and  $\widetilde{\mathbb{H}}_0$  look very different, the commutators  $[\mathbb{H}_0, b_q^*(k)]$  and  $[\widetilde{\mathbb{H}}_0, b_q^*(k)]$  are close. This is the so-called **linearization of the kinetic operator**.

Now we focus on

$$\begin{aligned}
\widetilde{\mathbb{H}}_0 + \mathbb{Q} &= \sum_{k \in \Gamma^{\text{nor}}} \left( \sum_{p \in B_F^c \cap (B_F + k)} h^2(p^2 - (p - k)^2) b_p^*(k) b_p(k) \right. \\
&\quad \left. + \frac{1}{N} \widehat{V}(k) \sum_{p, q \in B_F^c \cap (B_F + k)} \left( b_p^*(k) b_q(k) + b_p^*(-k) b_q(-k) + b_p^*(k) b_q^*(-k) + b_p(-k) b_q(k) \right) \right).
\end{aligned}$$

This operator can be treated similarly as a quadratic bosonic operator. If it were a truly bosonic, it could be diagonalized by a Bogoliubov transformation. Here it is only approximately bosonic, but we can still define a Bogoliubov transformation of the form  $e^B$  with

$$B = \sum_{k \in \Gamma^{\text{nor}}} \sum_{p, q \in B_F^c \cap (B_F + k)} K(k)_{p, q} b_p^*(k) b_q^*(k) - h.c.$$

The matrices  $K(k)$  are determined exactly as in the truly bosonic case (there is an algebraic formula for that). Thus using the approximate CCR we find that

$$e^{-B} (\widetilde{\mathbb{H}}_0 + \mathbb{Q}) e^B \approx E^{\text{corr}} + \sum_{k \in \Gamma^{\text{nor}}} A_p(k) b_p^*(k) b_p(k)$$

where  $E^{\text{corr}}$  is the desired correlation energy and  $A_p(k) \geq 0$ .

Thanks to the linearization of the kinetic operator, the difference  $\mathbb{H}_0 - \widetilde{\mathbb{H}}_0$  is mostly invariant

under the Bogoliubov transformation. Indeed,

$$\begin{aligned} e^{-B}(\mathbb{H}_0 - \tilde{\mathbb{H}}_0)e^B - (\mathbb{H}_0 - \tilde{\mathbb{H}}_0) &= \int_0^1 \frac{d}{dt} \left\{ e^{-tB}(\mathbb{H}_0 - \tilde{\mathbb{H}}_0)e^{tB} \right\} dt \\ &= \int_0^1 e^{-tB}[\mathbb{H}_0 - \tilde{\mathbb{H}}_0, B]e^{tB} dt \end{aligned}$$

This term is small because  $[\mathbb{H}_0 - \tilde{\mathbb{H}}_0, B] \approx 0$ , which follows from the fact that  $[\mathbb{H}_0 - \tilde{\mathbb{H}}_0, b_p^*(k)] \approx 0$ . Thus in summary we have

$$e^{-B}R^*H_NRe^B \approx E^{\text{HF}} + E^{\text{corr}} + \mathbb{H}_0 - \tilde{\mathbb{H}}_0 + \sum_{k \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + k)} A_p(k)b_p^*(k)b_p(k)$$

For an upper bound, we can apply the above operator inequality for the vacuum and find that

$$E_N \leq \langle 0 | e^{-B}R^*H_NRe^B | 0 \rangle \approx E^{\text{HF}} + E^{\text{corr}}.$$

The lower bound is more difficult as we have to estimate

$$\mathbb{H}_0 - \tilde{\mathbb{H}}_0 + \sum_{k \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + k)} A_p(k)b_p^*(k)b_p(k)$$

from below. At this point, we need the smallness condition on the interaction potential  $V$ . Note that

$$\sum_{k \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + k)} A_p(k)b_p^*(k)b_p(k) - \tilde{\mathbb{H}}_0 = \sum_{k \in \Gamma^{\text{nor}}} \sum_{p \in B_F^c \cap (B_F + k)} \left( A_p(k) - h^2(p^2 - (p-k)^2) \right) b_p^*(k)b_p(k).$$

When  $\hat{V}$  is small, then

$$\left| A_p(k) - h^2(p^2 - (p-k)^2) \right| \leq \varepsilon h^2(p^2 - (p-k)^2).$$

Hence, we can conclude using the operator inequality

$$\sum_{p \in B_F^c \cap (B_F + k)} h^2(p^2 - (p-k)^2)b_p^*(k)b_p(k) \leq C\mathbb{H}_0$$

which can be proved similarly to the kinetic inequality in the previous section.

# Chapter 11

## Stability of matter

We consider a sample of ordinary matter composed of  $N$  quantum electrons and  $M$  classical nuclei located at  $\{R_k\}_{k=1}^M \subset \mathbb{R}^3$ . The system is described by the Hamiltonian

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i}) - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{|x_i - R_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}$$

on  $L_a^2(\mathbb{R}^{3N})$ . Here the nuclear charges satisfy

$$0 < Z_k \leq Z, \quad \forall k = 1, 2, \dots, M.$$

Let  $E(M, N)$  be the ground state energy of the system, namely

$$E(M, N) = \inf_{\{R_k\}} \inf_{\|\Psi\|_{L_a^2(\mathbb{R}^{3N})} = 1} \langle \Psi, H_{M,N} \Psi \rangle.$$

Note that we minimize over both the electronic wave functions and the positions of nuclei. The nuclear repulsion  $\sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}$  is not important for the first infimum, but it is crucial for the second.

It is expected that this microscopic description of matter is consistent with what we observe everyday in the macroscopic level. In particular, we expect **the existence of thermodynamic limit**, namely

$$\lim_{\substack{M, N \rightarrow \infty \\ N/(M+N) \rightarrow \eta}} \frac{E(M+N)}{M+N} = e(\eta).$$

This implies, for example, the energy of two half-filled glasses of water is essentially the same with the energy of one fully-filled glass of water. Proving this existence is a nontrivial issue.

A crucial step of the proof is the uniform lower bound

**Theorem** (The stability of matter). *If  $0 < Z_k \leq Z$  for all  $k$ , then*

$$E(M + N) \geq -C_Z(M + N), \quad \forall M, N.$$

Once it is done, the existence of thermodynamic limit follows easily from a general argument based on the sub-additivity

$$E(M_1 + M_2, N_1 + N_2) \leq E(M_1, N_1) + E(M_2, N_2).$$

The stability of matter was first proved by **Dyson** and **Lenard** (1967). In 1975, **Lieb** and **Thirring** gave a very short proof, using their kinetic inequality

$$\left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i}) \Psi \right\rangle \geq K_3 \int_{\mathbb{R}^3} \rho_{\Psi}(x)^{5/3} dx$$

and **Teller's no-binding theorem** in the Thomas–Fermi theory. Another route to the stability due to **Solovej** is to use the Lieb–Thirring inequality together with **Baxter's electrostatic inequality**. These approaches will be discussed in this chapter.

## 11.1 Teller's no-binding theorem

Take  $M$  nuclei located at  $\{R_k\}_{k=1}^M \subset \mathbb{R}^3$  and with the nuclear charges  $\{Z_k\}_{k=1}^M$ ,  $Z_k > 0$ . Denote the nuclear potential

$$V(x) = - \sum_{k=1}^M \frac{Z_k}{|x - R_k|}.$$

Let us consider the Thomas–Fermi functional, with a constant  $c^{\text{TF}} > 0$ ,

$$\mathcal{E}_V^{\text{TF}}(\rho) = \int_{\mathbb{R}^3} (c^{\text{TF}} \rho(x)^{5/3} + V(x)\rho(x)) dx + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x - y|} dx dy + \sum_{1 \leq \ell < k \leq M} \frac{Z_{\ell} Z_k}{|R_{\ell} - R_k|}.$$

The Thomas–Fermi (absolute) ground state energy is

$$E^{\text{TF}}(\{R_k\}, \{Z_k\}) = \inf_{0 \leq \rho \in L^1 \cap L^{5/3}} \mathcal{E}_V^{\text{TF}}(\rho).$$

Similar to the atomic case, we have

**Theorem** (Existence in the Thomas–Fermi theory). *For every given  $\{R_k\}_{k=1}^M$  and  $\{Z_k\}_{k=1}^M$ , the infimum  $E^{\text{TF}}(\{R_k\}, \{Z_k\})$  is obtained by a unique density  $\rho^{\text{TF}}$  and it satisfies*

$$\int_{\mathbb{R}^3} \rho^{\text{TF}} = \sum_{k=1}^M Z_k.$$

The proof of this theorem is left as an exercise. Actually we do not really need this result for the proof below.

The main result in this section is that if we minimize  $E^{\text{TF}}(\{R_k\}, \{Z_k\})$  over  $\{R_k\}_{k=1}^M$ , then

$$E^{\text{TF}}(\{Z_k\}) = \inf_{\{R_k\}} E^{\text{TF}}(\{R_k\}, \{Z_k\})$$

is obtained when the nuclei are infinitely separated, namely **there is no molecular binding** in the Thomas–Fermi theory.

**Theorem** (Teller's no-binding theorem). *For every given  $\{Z_k\}_{k=1}^M$ , we have*

$$E^{\text{TF}}(\{Z_k\}) = \sum_{k=1}^M E_{\text{atom}}^{\text{TF}}(Z_k)$$

Here

$$E_{\text{atom}}^{\text{TF}}(Z_k) = \inf_{0 \leq \rho \in L^1 \cap L^{5/3}} \left( \int_{\mathbb{R}^3} \left( c^{\text{TF}} \rho(x)^{5/3} - Z_k \frac{\rho(x)}{|x|} \right) dx + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x-y|} dx dy \right).$$

This theorem was discovered by **Teller** (1962). A rigorous proof was found by **Lieb** and **Simon** (1977). The following proof is due to **Baxter** (1980), with some further simplification of **Loss**.

*Proof.* The inequality

$$E^{\text{TF}}(\{Z_k\}) \leq \sum_{k=1}^M E_{\text{atom}}^{\text{TF}}(Z_k)$$

is obvious since we can always put each nucleus infinitely far from the others. Here we focus on the lower bound. It suffices to show that for every given  $\{R_k\}_{k=1}^M$  and  $\{Z_k\}_{k=1}^M$ ,

$$E^{\text{TF}}(\{R_k\}_{k=1}^M, \{Z_k\}_{k=1}^M) \geq \sum_{k=1}^M E_{\text{atom}}^{\text{TF}}(Z_k).$$

By induction in  $M$ , we only need to show that

$$E^{\text{TF}}(\{R_k\}_{k=1}^M, \{Z_k\}_{k=1}^M) \geq E^{\text{TF}}(\{R_k\}_{k=1}^{M-1}, \{Z_k\}_{k=1}^{M-1}) + E_{\text{atom}}^{\text{TF}}(Z_M).$$

Denote

$$V(x) = - \sum_{k=1}^M \frac{Z_k}{|x - R_k|} = V_A(x) + V_B(x), \quad V_A(x) = - \sum_{k=1}^{M-1} \frac{Z_k}{|x - R_k|}, \quad V_B = - \frac{Z_M}{|x - R_M|}.$$

It suffices to show that for every  $0 \leq \rho \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3)$  (we can also assume  $\rho > 0$  everywhere), we can find two functions  $g, h \geq 0$ ,  $g + h = \rho$  such that

$$\mathcal{E}_V^{\text{TF}}(\rho) \geq \mathcal{E}_{V_A}^{\text{TF}}(g) + \mathcal{E}_{V_B}^{\text{TF}}(h).$$

For the kinetic energy, the condition  $g + h = \rho$  immediately implies the pointwise inequality  $(g + h)^{5/3} \geq g^{5/3} + h^{5/3}$ , and hence

$$\int_{\mathbb{R}^3} \rho^{5/3} \geq \int_{\mathbb{R}^3} g^{5/3} + \int_{\mathbb{R}^3} h^{5/3}.$$

Thus it remains to compare Coulomb potentials. We need to find  $g + h = \rho$  such that

$$\begin{aligned} & - \int_{\mathbb{R}^3} \rho(x)V(x)dx + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x-y|} dx dy + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|} \\ & \geq - \int_{\mathbb{R}^3} g(x)V_A(x)dx + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{g(x)g(y)}{|x-y|} dx dy + \sum_{1 \leq \ell < k \leq M-1} \frac{Z_\ell Z_k}{|R_\ell - R_k|} \\ & \quad - \int_{\mathbb{R}^3} h(x)V_B(x)dx + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{h(x)h(y)}{|x-y|} dx dy \end{aligned}$$

which is equivalent to

$$- \int_{\mathbb{R}^3} h(x)V_A(x)dx - \int_{\mathbb{R}^3} g(x)V_B(x)dx + \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{g(x)h(y)}{|x-y|} dx dy + \sum_{1 \leq \ell < M-1} \frac{Z_\ell Z_M}{|R_\ell - R_M|} \geq 0.$$

To make the notation more transparent, let us introduce

$$D(g, h) = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\overline{g(x)h(y)}}{|x-y|} dx dy = \frac{1}{2} \int_{\mathbb{R}^3} g(h * |x|^{-1})$$

(this notation extends naturally when  $g, h$  are measures) and

$$V_A(x) = - \sum_{k=1}^{M-1} \frac{Z_k}{|x - R_k|} = - \frac{1}{|x|} * m_A, \quad m_A = \sum_{k=1}^{M-1} Z_k \delta(x - R_k),$$

$$V_B(x) = - \frac{Z_M}{|x - R_M|} = - \frac{1}{|x|} * m_B, \quad m_B = Z_M \delta(x - R_M).$$

Then the desired inequality can be written as

$$-2D(h, m_A) - 2D(g, m_B) + 2D(g, h) + 2D(m_A, m_B) = 2D(g - m_A, h - m_B) \geq 0.$$

To construct  $g$  and  $h$ , we need the following special version of **Baxter's electrostatic inequality**.

**Exercise.** Let  $0 \leq \rho \in L^1(\mathbb{R}^3) \cap L^{5/3}(\mathbb{R}^3)$ . Then there exists  $0 \leq g \leq \rho$  such that

$$|x|^{-1} * g \leq |x|^{-1} * m_A, \quad \text{a.e. } x \in \mathbb{R}^3.$$

Moreover,  $|x|^{-1} * g = |x|^{-1} * m_A$  on  $\{x \in \mathbb{R}^3 : g(x) < \rho(x)\}$ .

(For our application, it suffices to prove the exercise with  $\rho > 0$  everywhere.)

Now we can conclude the proof of no-binding theorem. We choose  $g$  as in the above exercise and take  $h = \rho - g \geq 0$ . Then

$$\begin{aligned} 2D(g - m_A, h - m_B) &= \int_{\mathbb{R}^3} (|x|^{-1} * g - |x|^{-1} * m_A)(h - m_B) \\ &= \int_{g < \rho} + \int_{g = \rho} = 0 - \int_{g = \rho} (|x|^{-1} * g - |x|^{-1} * m_A)m_B \geq 0. \end{aligned}$$

□

## 11.2 First proof of the stability of matter

Now come back to the Hamiltonian on  $L_a^2(\mathbb{R}^{3N})$ :

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i} + V(x_i)) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}, \quad V(x) = - \sum_{k=1}^M \frac{Z_k}{|x - R_k|}.$$

We prove that the ground state energy

$$E(M, N) = \inf_{\{R_k\}} \inf_{\|\Psi\|_{L_a^2(\mathbb{R}^{3N})} = 1} \langle \Psi, H_{M,N} \Psi \rangle$$

satisfies the lower bound  $E(M, N) \geq -C(M + N)$ .

*First proof of the stability of matter.* Take a normalized wave function  $\Psi \in L_a^2(\mathbb{R}^{3N})$ . Then

$$\langle \Psi, H_{M,N} \Psi \rangle = \left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i}) \Psi \right\rangle + \int_{\mathbb{R}^3} V \rho_\Psi + \left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}.$$

By the Lieb–Thirring inequality we have

$$\left\langle \Psi, \sum_{i=1}^N (-\Delta_{x_i}) \Psi \right\rangle \geq K \int_{\mathbb{R}^3} \rho_\Psi^{5/3}$$

with a constant  $K > 0$ . Moreover, by the Lieb–Oxford inequality

$$\begin{aligned} \left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle &\geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\Psi(x) \rho_\Psi(y)}{|x - y|} dx dy - C \int_{\mathbb{R}^3} \rho_\Psi^{4/3} \\ &\geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\Psi(x) \rho_\Psi(y)}{|x - y|} dx dy - \frac{K}{2} \int_{\mathbb{R}^3} \rho_\Psi^{5/3} - C \int_{\mathbb{R}^3} \rho_\Psi \end{aligned}$$

Thus

$$\begin{aligned} \langle \Psi, H_{M,N} \Psi \rangle &\geq \frac{K}{2} \int_{\mathbb{R}^3} \rho_\Psi^{5/3} + \int_{\mathbb{R}^3} V \rho_\Psi + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\Psi(x) \rho_\Psi(y)}{|x - y|} dx dy \\ &\quad + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|} - CN \\ &= \mathcal{E}_V^{\text{TF}}(\rho_\Psi) - CN. \end{aligned}$$

By Teller’s no-binding theorem,

$$\mathcal{E}_V^{\text{TF}}(\rho_\Psi) \geq \sum_{k=1}^M E_{\text{atom}}^{\text{TF}}(Z_k) \geq -C \sum_{k=1}^M Z_k^{7/3} \geq -CMZ^{7/3}.$$

Thus we conclude that

$$\langle \Psi, H_{M,N} \Psi \rangle \geq -CMZ^{7/3} - CN.$$

Optimizing over  $\Psi \in L_a^2(\mathbb{R}^{3N})$ , the lower bound for  $E(M, N)$  follows.  $\square$



Remark: In the original proof of Lieb and Thirring (1975), they did not use the Lieb–Oxford inequality. They bound the indirect energy by Teller's no-binding theorem as follows. From

$$\frac{K}{2} \int_{\mathbb{R}^3} \rho^{5/3} - \int_{\mathbb{R}^3} \sum_{k=1}^M \frac{\rho(x)}{|x - R_k|} + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_{\Psi}(x)\rho_{\Psi}(y)}{|x - y|} dx dy + \sum_{1 \leq \ell < k \leq M} \frac{Z_{\ell} Z_k}{|R_{\ell} - R_k|} \geq -C \sum_{k=1}^M Z_k^{7/3},$$

taking  $Z_k = 1$ ,  $M = N$ ,  $R_k = x_k$ ,  $\rho = \rho_{\Psi}$  with  $\Psi \in L_a^2(\mathbb{R}^{3N})$  we have

$$\frac{K}{2} \int_{\mathbb{R}^3} \rho_{\Psi}^{5/3} - \sum_{i=1}^N (\rho_{\Psi} * |x|^{-1})(x_i) + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_{\Psi}(x)\rho_{\Psi}(y)}{|x - y|} dx dy + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \geq -CN.$$

This inequality holds pointwise for all  $\{x_i\}_{i=1}^N \subset \mathbb{R}^3$ . Taking the expectation against the normalized wave function  $\Psi \in L_a^2(\mathbb{R}^{3N})$ , we obtain

$$\frac{K}{2} \int_{\mathbb{R}^3} \rho_{\Psi}^{5/3} - \int_{\mathbb{R}^3} (\rho_{\Psi} * |x|^{-1})\rho_{\Psi} + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x - y|} dx dy + \left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle \geq -CN$$

which is equivalent to

$$\left\langle \Psi, \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} \Psi \right\rangle \geq \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x - y|} dx dy - \frac{K}{2} \int_{\mathbb{R}^3} \rho_{\Psi}^{5/3} - CN.$$

### 11.3 Baxter's electrostatic inequality

In 1980, **Baxter** proposed a simplified approach to the stability of matter. He proved the following inequality that quantifies electrostatic screening (effectively, any electron sees only the nearest nucleus).

**Theorem** (Baxter's electrostatic inequality). *For any  $\{x_i\}_{i=1}^N, \{R_k\}_{k=1}^M \subset \mathbb{R}^3$  and  $Z > 0$ . Then*

$$\sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=1}^N \sum_{k=1}^M \frac{Z}{|x_i - R_k|} + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_{\ell} - R_k|} \geq - \sum_{i=1}^N \frac{2Z + 1}{\mathfrak{D}(x_i)}$$

where  $\mathfrak{D}(x) = \min_{1 \leq k \leq M} |x - R_k|$  is the distance to the nearest nucleus.

We will derive Baxter's result from another basic electrostatic inequality. The idea is due to **Solovej**.

*Proof. Step 1.* Define

$$\Phi(x) = \sum_{k=1}^M \frac{Z}{|x - R_k|} - \frac{Z}{\mathfrak{D}(x)}.$$

The function  $\Phi(x)$  is the potential generated by all nuclei but the nearest one. In particular, we can write

$$\sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} = \frac{Z}{2} \sum_{k=1}^M \Phi(R_k).$$

A useful picture to keep in mind is the **Voronoi cells** where the nucleus  $R_k$  is associated with the cell

$$\Gamma_k = \{x \in \mathbb{R}^3 : |x - R_k| < |x - R_\ell| \text{ for all } \ell \neq k\}.$$

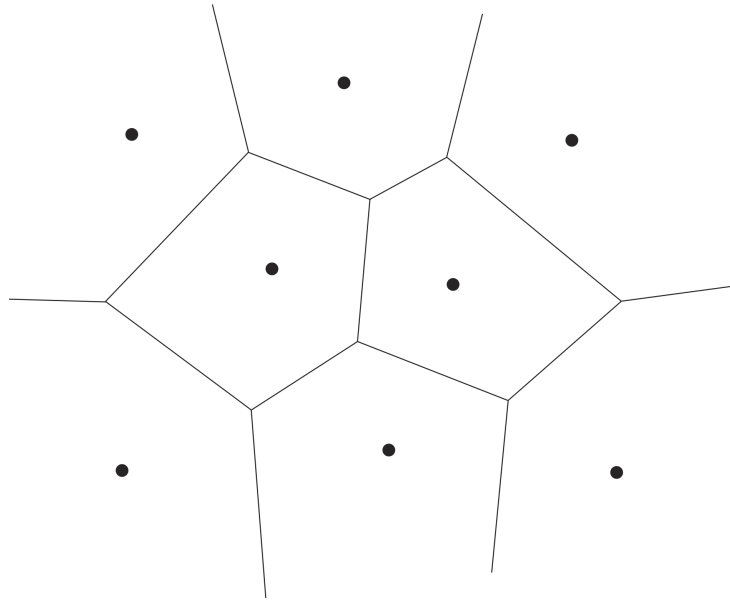


Figure: Voronoi cells in 2D

**Lemma.** *We can write*

$$\Phi(x) = \int_{\mathbb{R}^3} \frac{d\nu(y)}{|x - y|}$$

*with some non-negative measure  $\nu$  on  $\mathbb{R}^3$  supported on the surfaces*

$$\{x \in \mathbb{R}^3 : |x - R_k| = |x - R_\ell| \text{ for some } \ell \neq k\}.$$

*Proof.* We need to prove that

$$-\Delta\Phi = 4\pi\nu.$$

Since  $-\Delta(|x|^{-1}) = 0$  for all  $x \neq 0$ , it is clear that  $-\Delta\Phi(x) = 0$  for all  $x \in \mathbb{R}^3$  except the surfaces. Hence, if  $-\Delta\Phi = 4\pi\nu$  then  $\nu$  must be supported on the surfaces. To compute  $\nu$ , let us take  $f \in C_c^\infty(\mathbb{R}^3)$  and use Voronoi's cells

$$\begin{aligned} \int_{\mathbb{R}^3} \Phi(x)(\Delta f)(x)dx &= \sum_k \int_{\Gamma_k} \Phi(x)(\Delta f)(x)dx \\ &= \sum_k \left( \int_{\Gamma_k} \operatorname{div}(\Phi \nabla f)(x)dx - \int_{\Gamma_k} \nabla \Phi(x) \cdot \nabla f(x)dx \right) \\ &= \sum_k \left( \int_{\partial\Gamma_k} \Phi(x) \mathbf{n}_k \cdot \nabla f(x) dS - \int_{\Gamma_k} \operatorname{div}(f \nabla \Phi)(x)dx + \int_{\Gamma_k} f(x) \Delta \Phi(x)dx \right). \end{aligned}$$

Here  $\mathbf{n}_k$  is the outward normal vector to  $\partial\Gamma_k$  and  $dS$  is the 2D surface measure on  $\partial\Gamma_k$ . Note that on each surface  $|x - R_k| = |x - R_\ell|$ , the outward normal vectors  $\mathbf{n}_k$  and  $\mathbf{n}_\ell$  point in opposite directions. Moreover,  $\Phi$  and  $\nabla f$  are continuous on  $\mathbb{R}^3$ . Thus the contribution from all boundary terms is 0. We also have

$$\int_{\Gamma_k} f(x) \Delta \Phi(x) dx = 0$$

since  $\Delta\Phi = 0$  on  $\Gamma_k$ . Thus

$$\int_{\mathbb{R}^3} \Phi(x)(\Delta f)(x)dx = - \sum_k \int_{\Gamma_k} \operatorname{div}(f \nabla \Phi)(x)dx = \sum_k \int_{\partial\Gamma_k} f(x) \mathbf{n}_k \cdot \nabla \Phi(x)dx.$$

Note that this term is non-zero since  $\nabla\Phi$  is not continuous on  $\mathbb{R}^3$ . A straightforward computation shows that

$$\begin{aligned} \int_{\mathbb{R}^3} \Phi(x)(\Delta f)(x)dx &= 2Z \sum_k \int_{\partial\Gamma_k} f(x) \mathbf{n}_k \cdot \nabla \frac{1}{|x - R_k|} dx \\ &= -2Z \sum_k \int_{\partial\Gamma_k} f(x) \frac{\mathbf{n}_k \cdot (x - R_k)}{|x - R_k|^3} dx. \end{aligned}$$

Here we have a factor 2 because any surface term is counted twice. Thus

$$-\Delta\Phi = 4\pi\nu$$

with

$$4\pi d\nu(x) = 2Z \sum_k \mathbf{1}(x \in \partial\Gamma_k) \frac{\mathbf{n}_k \cdot (x - R_k)}{|x - R_k|^3} \geq 0.$$

Here  $\mathbf{n}_k \cdot (x - R_k) \geq 0$  on  $x \in \partial\Gamma_k$  since  $\Gamma_k$  is convex.  $\square$

**Step 2.** We have the following basic electrostatic inequality.

**Lemma.** *Let  $\mu$  be a measure with  $D(\mu, \mu) < \infty$ . Then*

$$D(\mu, \mu) - \int \Phi(x) \mu(dx) + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} \geq 0.$$

Recall that

$$D(f, g) = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\overline{f(x)}g(y)}{|x - y|} dx dy.$$

This notation extends naturally when  $f, g$  are measures.

*Proof.* Using  $\Phi(x) = |x|^{-1} * \nu$  we have

$$D(\mu, \mu) - \int \Phi(x) \mu(dx) = D(\mu, \mu) - 2D(\mu, \nu) \geq -D(\nu, \nu)$$

since  $D(\mu - \nu, \mu - \nu) \geq 0$ . It remains to calculate  $D(\nu, \nu)$ . Using the equation  $\Phi = |x|^{-1} * \nu$  and the fact

$$\Phi(x) = \sum_{k=1}^M \frac{Z}{|x - R_k|} - \frac{Z}{\mathfrak{D}(x)} \leq \sum_{k=1}^M \frac{Z}{|x - R_k|} = Z \sum_{k=1}^M \int_{\mathbb{R}^3} \delta(y - R_k) \frac{1}{|x - y|} dy$$

we can write

$$\begin{aligned} D(\nu, \nu) &= \frac{1}{2} \int_{\mathbb{R}^3} \Phi(x) d\nu(x) \leq \frac{Z}{2} \sum_{k=1}^M \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \delta(y - R_k) \frac{1}{|x - y|} dy d\nu(x) \\ &= \frac{Z}{2} \sum_{k=1}^M \int_{\mathbb{R}^3} \delta(y - R_k) \left( \int_{\mathbb{R}^3} \frac{1}{|x - y|} d\nu(x) \right) dy = \frac{Z}{2} \sum_{k=1}^M \int_{\mathbb{R}^3} \delta(y - R_k) \Phi(y) dy \\ &= \frac{Z}{2} \sum_{k=1}^M \Phi(R_k) = \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|}. \end{aligned}$$

Here we also used the fact that  $\nu$  is a positive measure. This completes the proof of the lemma.  $\square$

**Step 3.** Now we conclude the proof of Baxter's electrostatic inequality. Recall  $\mathfrak{D}(x) =$

$\min_k |x - R_k|$ . Denote the smeared out spherical charges

$$d\mu_i(x) = \frac{1}{\pi \mathfrak{D}(x_i)^2} \delta(|x - x_i| - \mathfrak{D}(x_i)/2) dx, \quad i = 1, 2, \dots, N.$$

Note that  $\mu_i(\mathbb{R}^3) = 1$  and that  $\mathfrak{D}(x) \geq \mathfrak{D}(x_i)/2$  for every  $x \in \text{supp}\mu_i$ . We apply the above lemma

$$D(\mu, \mu) - \int \Phi(x) \mu(dx) + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} \geq 0$$

with  $\mu = \sum_{i=1}^N \mu_i$ . By Newton's theorem we have

$$\begin{aligned} D(\mu_i, \mu_i) &= \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{d\mu_i(x) d\mu_i(y)}{|x - y|} = \frac{1}{\mathfrak{D}(x_i)}, \\ D(\mu_i, \mu_j) &= \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{d\mu_i(x) d\mu_j(y)}{|x - y|} \leq \frac{1}{2|x_i - x_j|}, \quad \forall i \neq j, \end{aligned}$$

and hence,

$$D(\mu, \mu) = \sum_{i,j} D(\mu_i, \mu_j) \leq \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{i=1}^N \frac{1}{\mathfrak{D}(x_i)}.$$

Using  $\mathfrak{D}(x) = \min_k |x - R_k| \geq \mathfrak{D}(x_i)/2$  for every  $x \in \text{supp}\mu_i$ , by Newton's theorem we have

$$\begin{aligned} \int \Phi(x) d\mu(x) &= \sum_{i=1}^N \int \Phi(x) d\mu_i(x) = \sum_{k=1}^M \sum_{i=1}^N \int \frac{Z}{|x - R_k|} d\mu_i(x) - \sum_{i=1}^N \int \frac{Z}{\mathfrak{D}(x)} d\mu_i(x) \\ &= \sum_{k=1}^M \sum_{i=1}^N \frac{Z}{|x_i - R_k|} - \sum_{i=1}^N \int \frac{Z}{\mathfrak{D}(x)} d\mu_i(x) \\ &\geq \sum_{k=1}^M \sum_{i=1}^N \frac{Z}{|x_i - R_k|} - \sum_{i=1}^N \frac{2Z}{\mathfrak{D}(x_i)}. \end{aligned}$$

Thus in summary

$$\begin{aligned} 0 &\leq D(\mu, \mu) - \int \Phi(x) \mu(dx) + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} \\ &\leq \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{k=1}^M \sum_{i=1}^N \frac{Z}{|x_i - R_k|} + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} + \sum_{i=1}^N \frac{2Z + 1}{\mathfrak{D}(x_i)} \end{aligned}$$

which is equivalent to the desired inequality. □

## 11.4 Second proof of the stability of matter

Come back again to the Hamiltonian on  $L_a^2(\mathbb{R}^{3N})$ :

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i}) - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{|x_i - R_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}.$$

We prove that the ground state energy

$$E(M, N) = \inf_{\{R_k\}} \inf_{\|\Psi\|_{L_a^2(\mathbb{R}^{3N})} = 1} \langle \Psi, H_{M,N} \Psi \rangle$$

satisfies the lower bound  $E(M, N) \geq -C(M + N)$ .

*Second proof of the stability of matter. Step 1.* We consider the simple case where  $Z_k = Z$  for all  $k = 1, 2, \dots, M$ . We use Baxter's electrostatic inequality

$$\sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=1}^N \sum_{k=1}^M \frac{Z}{|x_i - R_k|} + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} \geq - \sum_{i=1}^N \frac{2Z + 1}{\mathfrak{D}(x_i)}$$

where  $\mathfrak{D}(x) = \min_{1 \leq k \leq M} |x - R_k|$  is the distance to the nearest nucleus. Thus

$$H_{M,N} \geq \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{2Z + 1}{\mathfrak{D}(x_i)} \right).$$

Hence, by Pauli's exclusion principle and the Lieb–Thirring inequality, for every  $\mu > 0$  we have

$$\begin{aligned} H_{M,N} &\geq \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{2Z + 1}{\mathfrak{D}(x_i)} + \mu \right) - \mu N \\ &\geq \text{Tr} \left[ -\Delta - \frac{2Z + 1}{\mathfrak{D}(x)} + \mu \right]_- - \mu N \\ &\geq -L_{1,3} \int_{\mathbb{R}^3} \left[ \frac{2Z + 1}{\mathfrak{D}(x)} - \mu \right]_+^{5/2} dx - \mu N. \end{aligned}$$

Since  $\mathfrak{D}(x) = \min_{1 \leq k \leq M} |x - R_k|$ , we have

$$\left[ \frac{2Z + 1}{\mathfrak{D}(x)} - \mu \right]_+^{5/2} \leq \sum_{k=1}^M \left[ \frac{2Z + 1}{|x - R_k|} - \mu \right]_+^{5/2}.$$

and hence

$$\begin{aligned} \int_{\mathbb{R}^3} \left[ \frac{2Z+1}{\mathfrak{D}(x)} - \mu \right]_+^{5/2} dx &\leq \sum_{k=1}^M \int_{\mathbb{R}^3} \left[ \frac{2Z+1}{|x-R_k|} - \mu \right]_+^{5/2} dx \\ &= M \int_{\mathbb{R}^3} \left[ \frac{2Z+1}{|x|} - \mu \right]_+^{5/2} dx = M \frac{5\pi^2(2Z+1)^3}{4\sqrt{\mu}}. \end{aligned}$$

In summary,

$$H_{M,N} \geq -M \frac{5\pi^2(2Z+1)^3}{4\sqrt{\mu}} - \mu N.$$

This gives the desired lower bound  $E(M, N) \geq -C(M+N)$ . Actually by optimizing over  $\mu > 0$  we find that

$$E(M, N) \geq -C(2Z+1)^2 M^{2/3} N^{1/3}.$$

**Step 2.** Now we come to the general case when  $Z_k \leq Z$  for all  $k = 1, 2, \dots, M$ . The proof in this case follows from Step 1 and the following **monotonicity in nuclear charges**.

**Lemma.** Denote  $E(M, N, \{Z_k\})$  be the ground state energy of  $H_{M,N}$  with given nuclear charges  $\{Z_k\}$ . If  $Z_k \leq \tilde{Z}_k$  for all  $k = 1, 2, \dots, M$ , then

$$E(M, N, \{Z_k\}) \geq E(M, N, \{\tilde{Z}_k\}).$$

This observation is due to **Daubechies** and **Lieb** (1983).

*Proof of the lemma.* Note that for every  $\ell \in \{1, 2, \dots, M\}$ , the mapping  $Z_\ell \rightarrow H_{M,N}$  is linear. Therefore, the mapping

$$Z_\ell \mapsto E(M, N, \{Z_k\})$$

is concave (the concavity holds separately for each  $Z_\ell$ , not jointly for  $\{Z_k\}$ ). Under the condition  $0 \leq Z_\ell \leq \tilde{Z}_\ell$  we can write

$$Z_\ell = t \cdot 0 + (1-t) \cdot \tilde{Z}_\ell, \quad \text{for some } t \in [0, 1].$$

Hence, the concavity implies that

$$E(M, N, \{Z_k\}) \geq tE(M, N, \{Z_k\})_{|Z_\ell=0} + (1-t)E(M, N, \{Z_k\})_{|Z_\ell=\tilde{Z}_\ell}.$$

On the other hand, setting  $Z_\ell = 0$  is equivalent to putting  $R_\ell$  at infinity, and hence

$$E(M, N, \{Z_k\})|_{Z_\ell=0} \geq E(M, N, \{Z_k\})|_{Z_\ell=\tilde{Z}_\ell}.$$

Thus

$$E(M, N, \{Z_k\}) \geq E(M, N, \{Z_k\})|_{Z_\ell=\tilde{Z}_\ell}.$$

By induction, we find that

$$E(M, N, \{Z_k\}) \geq E(M, N, \{\tilde{Z}_k\}).$$

This completes the proof of the lemma.  $\square$

From the condition  $Z_k \leq Z$  for all  $k = 1, 2, \dots, M$  and the lemma, we find that

$$E(M, N, \{Z_k\}) \geq E(M, N, \{Z, Z, \dots, Z\}).$$

By Step 1, we have

$$E(M, N, \{Z, Z, \dots, Z\}) \geq -C_Z(M + N).$$

This implies the same lower bound for  $E(M, N, \{Z_k\})$ . Thus  $E(M, N) \geq -C(M + N)$ .  $\square$

## 11.5 Existence of thermodynamic limit

Consider the Hamiltonian on  $L_a^2(\mathbb{R}^{3N})$

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i}) - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{|x_i - R_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}.$$

and the ground state energy

$$E(M, N) = \inf_{\{R_k\}} \inf_{\|\Psi\|_{L_a^2(\mathbb{R}^{3N})} = 1} \langle \Psi, H_{M,N} \Psi \rangle.$$

**Theorem** (Existence of thermodynamic limit). *Assume  $Z_k = Z$  for all  $k = 1, 2, \dots, M$ .*



For every  $\eta \in (0, 1)$ , we have

$$\lim_{\substack{M, N \rightarrow \infty \\ N/(M+N) \rightarrow \eta}} \frac{E(M, N)}{M + N} = e(\eta).$$

The function  $e : (0, 1) \rightarrow \mathbb{R}$  is bounded and convex (consequently it is continuous).

*Proof. Step 1.* We have the following **sub-additivity**

$$E(M_1 + M_2, N_1 + N_2) \leq E(M_1, N_1) + E(M_2, N_2).$$

This is an easy consequence of the variational principle. More precisely, given two wave functions  $\Psi_{N_1} \in L_a^2(\mathbb{R}^{3N_1})$  and  $\Psi_{N_2} \in L_a^2(\mathbb{R}^{3N_2})$  we can construct a trial wave function  $\Psi_{N_1, N_2}^{(y)}$  in  $L_a^2(\mathbb{R}^{3(N_1+N_2)})$  by antisymmetrizing the product

$$\Psi_{N_1}(x_1, \dots, x_{N_1}) \Psi_{N_2}(x_{N_1+1} + y, \dots, x_{N_1+N_2} + y).$$

Then

$$\begin{aligned} E(M_1 + M_2, N_1 + N_2) &\leq \lim_{|y| \rightarrow \infty} \langle \Psi_{N_1, N_2}^{(y)}, H_{M_1+M_2, N_1+N_2} \Psi_{N_1, N_2}^{(y)} \rangle \\ &= \langle \Psi_{N_1}, H_{M_1, N_1} \Psi_{N_1} \rangle + \langle \Psi_{N_2}, H_{M_2, N_2} \Psi_{N_2} \rangle. \end{aligned}$$

Technically, if we have the nuclei  $\{R_k\}_{k=1}^{M_1}$  for  $H_{M_1, N_1}$  and  $\{\tilde{R}_k\}_{k=1}^{M_2}$  for  $H_{M_2, N_2}$ , then we take the nuclei

$$\{R_1, R_2, \dots, R_{M_1}, \tilde{R}_1 + y, \tilde{R}_2 + y, \dots, \tilde{R}_{M_2} + y\}$$

for  $H_{M_1+M_2, N_1+N_2}$ . Optimizing over  $\Psi_{N_1}, \Psi_{N_2}$  and the locations of the nuclei, we obtain the desired inequality.

**Step 2.** From the sub-additivity and the negativity of atomic energy we find that  $E(M, N) \leq 0$  and  $(M, N) \rightarrow E(M, N)$  is decreasing. From the stability of matter, we obtain

$$0 \geq \frac{E(M + N)}{M + N} \geq -C.$$

Hence, for every  $\eta \in (0, 1)$ , we can find a sequence  $(M_j, N_j)$  such that

$$M_j, N_j \rightarrow \infty, \quad \frac{N_j}{M_j + N_j} \rightarrow \eta, \quad \lim_{j \rightarrow \infty} \frac{E(M_j, N_j)}{M_j + N_j} \text{ exists.}$$

It remains to show that for any  $(M'_j, N'_j)$  with  $M'_j, N'_j \rightarrow \infty$ ,  $\frac{N'_j}{M'_j + N'_j} \rightarrow \eta$  we also have

$$\lim_{j \rightarrow \infty} \frac{E(M'_j, N'_j)}{M'_j + N'_j} = \lim_{j \rightarrow \infty} \frac{E(M_j, N_j)}{M_j + N_j}$$

Indeed, by passing to a subsequence of  $(M'_j, N'_j)$  if necessary, we can assume that  $M'_j/M_j \rightarrow \infty$  and  $N'_j/N_j \rightarrow \infty$ . Define

$$L_j := \min \left\{ \left[ \frac{M'_j}{M_j} \right], \left[ \frac{N'_j}{N_j} \right] \right\} \in \mathbb{N}$$

where  $[t]$  is the integer part of  $t$  (i.e.  $[t] \leq t < [t] + 1$ ). Then

$$M'_j \geq L_j M_j, \quad N'_j \geq L_j N_j, \quad \frac{L_j(M_j + N_j)}{M'_j + N'_j} \rightarrow 1.$$

The last convergence comes from the fact that  $\frac{N'_j}{M'_j + N'_j}$  and  $\frac{N_j}{M_j + N_j}$  have the same limit. Thus by the monotonicity and the sub-additivity of  $E(M, N)$  we have

$$\frac{E(M'_j, N'_j)}{M'_j + N'_j} \leq \frac{E(L_j M_j, L_j N_j)}{M'_j + N'_j} \leq \frac{L_j E(M_j, N_j)}{M'_j + N'_j} = \frac{L_j(M_j + N_j)}{M'_j + N'_j} \cdot \frac{E(M_j, N_j)}{M_j + N_j}.$$

Thus

$$\limsup_{j \rightarrow \infty} \frac{E(M'_j, N'_j)}{M'_j + N'_j} \leq \lim_{j \rightarrow \infty} \frac{E(M_j, N_j)}{M_j + N_j}$$

Similarly, by passing to a subsequence of  $(M_j, N_j)$  if necessary, we can assume that  $M_j/M'_j \rightarrow \infty$  and  $N_j/N'_j \rightarrow \infty$ . The same argument as above

$$\liminf_{j \rightarrow \infty} \frac{E(M'_j, N'_j)}{M'_j + N'_j} \geq \lim_{j \rightarrow \infty} \frac{E(M_j, N_j)}{M_j + N_j}.$$

In summary,

$$\lim_{j \rightarrow \infty} \frac{E(M'_j, N'_j)}{M'_j + N'_j} = \lim_{j \rightarrow \infty} \frac{E(M_j, N_j)}{M_j + N_j}.$$

Hence, the limit

$$e(\eta) := \lim_{\substack{M, N \rightarrow \infty \\ N/(M+N) \rightarrow \eta}} \frac{E(M, N)}{M + N}$$

exists.

**Step 3.** Since  $0 \geq E(M + N) \geq -C(M + N)$ , we have

$$0 \geq e(\eta) \geq -C$$

for every  $\eta \in (0, 1)$ . Thus  $e(\eta)$  is bounded. The convexity of  $\eta \rightarrow e(\eta)$  follows from the sub-additivity of  $E(M, N)$ . Indeed, let  $\eta, \eta' \in (0, 1)$  and let  $N, M = M(N), M' = M'(N) \rightarrow \infty$  such that

$$\frac{N}{M+N} \rightarrow \eta, \quad \frac{N}{M'+N} \rightarrow \eta'.$$

Then

$$\frac{N(M+N) + N(M'+N)}{2(M+N)(M'+N)} = \frac{1}{2} \left( \frac{N}{M'+N} + \frac{N}{M+N} \right) \rightarrow \frac{\eta + \eta'}{2}.$$

Note that  $2(M+N)(M'+N) - N(M+N) - N(M'+N) = NM + NM' + 2MM'$ . Hence,

$$\frac{E(NM + NM' + 2MM', N(N+M) + N(N+M'))}{2(M+N)(M'+N)} \rightarrow e\left(\frac{\eta + \eta'}{2}\right).$$

On the other hand, by the sub-additivity of the ground state energy, we have

$$\begin{aligned} & E(NM + NM' + 2MM', N(N+M) + N(N+M')) \\ & \leq E(NM, N^2) + E(NM', N^2) + E(MM', NM) + E(MM', NM') \\ & \leq NE(M, N) + NE(M', N) + ME(M', N) + M'E(M, N) \\ & = (M' + N)E(M, N) + (M + N)E(M', N). \end{aligned}$$

Dividing both sides by  $2(M+N)(M'+N)$  we obtain

$$\begin{aligned} & \frac{E(NM + NM' + 2MM', N(N+M) + N(N+M'))}{2(M+N)(M'+N)} \\ & \leq \frac{1}{2} \left( \frac{E(M, N)}{M+N} + \frac{E(M', N)}{M'+N} \right). \end{aligned}$$

Taking the limit we conclude that

$$e\left(\frac{\eta + \eta'}{2}\right) \leq \frac{1}{2} (e(\eta) + e(\eta')).$$

Since  $e(\eta)$  is uniformly bounded in  $(0, 1)$ , the latter bound implies the convexity, namely

$$e((1-t)\eta + t\eta') \leq (1-t)e(\eta) + te(\eta'), \quad \forall t \in (0, 1).$$

□

**Exercise.** Let  $f : (0, 1) \rightarrow \mathbb{R}$  be a bounded function such that

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2}, \quad \forall a, b \in (0, 1).$$

Prove that  $f$  is convex in  $(0, 1)$ .

## 11.6 Grand-canonical stability

There is also the stability in the **grand-canonical setting**, where the ground state energy is computed without the particle number constraint, but with a volume constraint. For simplicity, let us consider a system of  $N$  particles of charge  $-1$  and  $M$  particles of charge  $+1$  in an open bounded set  $\Omega \subset \mathbb{R}^3$ . The system is described by the Hamiltonian

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i}) + \sum_{k=1}^M (-\Delta_{y_k}) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq k < \ell \leq M} \frac{1}{|y_k - y_\ell|} - \sum_{i=1}^N \sum_{k=1}^M \frac{1}{|x_i - y_k|}$$

on  $L_a^2(\Omega^N) \otimes L_a^2(\Omega^M)$ . The grand canonical ground state energy is

$$E(\Omega) = \inf_{M,N} \inf_{\substack{\Psi \in C_c^\infty(\Omega^{M+N}) \\ \|\Psi\|_{L^2} = 1}} \langle \Psi, H_{M,N} \Psi \rangle.$$

**Theorem** (Grand-canonical stability). *We have*

$$E(\Omega) \geq -C|\Omega|$$

with a finite constant  $C > 0$  independent of  $\Omega$ .

This result holds under a more general assumption, where the masses and the charges of the particles can be different.

*Proof.* By the canonical stability we have

$$\sum_{i=1}^N \left(-\frac{1}{2}\Delta_{x_i}\right) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq k < \ell \leq M} \frac{1}{|y_k - y_\ell|} - \sum_{i=1}^N \sum_{k=1}^M \frac{1}{|x_i - y_k|} \geq -C(M+N).$$

Note that we used only half of the kinetic energy of the first  $N$  particles and did not use the kinetic energy of the last  $M$  particles. By the Lieb–Thirring inequality (or the Berezin–Li–Yau inequality as we have the Dirichlet boundary condition) we have

$$\sum_{i=1}^N \left( -\frac{1}{2} \Delta_{x_i} \right) \geq \frac{K}{|\Omega|^{2/3}} N^{5/3}, \quad \sum_{k=1}^M \left( -\frac{1}{2} \Delta_{y_j} \right) \geq \frac{K}{|\Omega|^{2/3}} M^{5/3}$$

with a constant  $K > 0$ . Thus in summary,

$$H_{M,N} \geq \frac{K}{|\Omega|^{2/3}} (M^{5/3} + N^{5/3}) - C(M + N).$$

Thus

$$E(\Omega) \geq \inf_{M,N} \left( \frac{K}{|\Omega|^{2/3}} (M^{5/3} + N^{5/3}) - C(M + N) \right) \geq -C|\Omega|.$$

□

Note that the energy  $E(\Omega)$  satisfies the following properties:

- **Translation-invariant**  $E(\Omega + z) = E(\Omega)$  for all  $z \in \mathbb{R}^3$ .
- **Sub-additivity**  $E(\Omega_1 \cup \Omega_2) \leq E(\Omega_1) + E(\Omega_2)$  if  $\Omega_1 \cap \Omega_2 = \emptyset$ .
- **Stability**  $E(\Omega) \geq -C|\Omega|$ .

All that implies

**Theorem** (Existence of thermodynamic limit). *The limit*

$$\lim_{\substack{\Omega=[-L,L]^3 \\ L \rightarrow \infty}} \frac{E(\Omega)}{|\Omega|}$$

*exists and it is finite.*

The proof of this theorem is left as an exercise. Actually the existence of the thermodynamic limit holds for a much bigger class of domains  $\Omega$ . For example, **tilling domains** are allowed.

## 11.7 Instability for bosons

In the proofs of the stability of matter, the fermionic property is crucial. Indeed, the stability fails if Pauli's exclusion principle is turned off. Let us consider the Hamiltonian

$$H_{M,N} = \sum_{i=1}^N (-\Delta_{x_i}) - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{|x_i - R_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq M} \frac{Z_\ell Z_k}{|R_\ell - R_k|}.$$

on  $L^2(\mathbb{R}^{3N})$  (without the anti-symmetry assumption). The corresponding ground state energy is

$$E_B(M, N) = \inf_{\{R_k\}} \inf_{\|\Psi\|_{L^2(\mathbb{R}^{3N})}=1} \langle \Psi, H_{M,N} \Psi \rangle.$$

Remark: It is well-known that by a convexity argument (c.f. the diamagnetic inequality), the ground state energy on the full Hilbert space  $L^2(\mathbb{R}^{3N})$  is the same with the restriction to the bosonic/symmetric space  $L_s^2(\mathbb{R}^{3N})$ .

**Theorem** ( $N^{5/3}$  instability). *Let  $M = N$  and  $Z_k = 1$  for all  $k$ . Then*

$$-CN^{5/3} \leq E_B(M, N) \leq -C^{-1}N^{5/3}$$

for a constant  $C > 0$  independent of  $N$ .

The lower bound was proved by **Dyson** and **Lenard** (1967). The upper bound was proved by **Lieb** (1979).

*Proof. Lower bound.* By Baxter's electrostatic inequality

$$\sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=1}^N \sum_{k=1}^M \frac{Z}{|x_i - R_k|} + \sum_{1 \leq \ell < k \leq M} \frac{Z^2}{|R_\ell - R_k|} \geq - \sum_{i=1}^N \frac{3}{\mathfrak{D}(x_i)}$$

where  $\mathfrak{D}(x) = \min_{1 \leq k \leq M} |x - R_k|$  we have

$$H_{M,N} \geq \sum_{i=1}^N \left( -\Delta_{x_i} - \frac{3}{\mathfrak{D}(x_i)} \right).$$

Thus the desired inequality follows from the one-body operator inequality

$$-\Delta - \frac{3}{\mathfrak{D}(x)} \geq -CN^{2/3} \quad \text{on } L^2(\mathbb{R}^3).$$

By Sobolev's inequality (c.f. CLR bound), there exists  $\varepsilon_0 > 0$  such that for every  $\mu > 0$  satisfying

$$\int_{\mathbb{R}^3} \left[ \frac{3}{\mathfrak{D}(x)} - \mu \right]_+^{3/2} dx \leq \varepsilon_0.$$

we have

$$-\Delta - \frac{3}{\mathfrak{D}(x)} + \mu \geq 0.$$

By the definition of  $\mathfrak{D}(x)$ , we can bound

$$\begin{aligned} \int_{\mathbb{R}^3} \left[ \frac{3}{\mathfrak{D}(x)} - \mu \right]_+^{3/2} dx &\leq \sum_{k=1}^N \int_{\mathbb{R}^3} \left[ \frac{3}{|x - R_k|} - \mu \right]_+^{3/2} dx \\ &= N \int_{\mathbb{R}^3} \left[ \frac{3}{|x|} - \mu \right]_+^{3/2} dx = CN\mu^{-3/2}. \end{aligned}$$

Thus the condition  $CN\mu^{-3/2} \leq \varepsilon_0$  is satisfied when  $\mu \sim N^{2/3}$ , as desired.

**Upper bound.** We take the trial function

$$\Psi(x_1, \dots, x_N) = u^{\otimes N}(x_1, \dots, x_N) = u(x_1)u(x_2)\dots u(x_N).$$

with a normalized function  $u \in L^2(\mathbb{R}^3)$ . Then

$$\begin{aligned} E_B(M, N) &\leq \langle u^{\otimes N}, H_{M,N} u^{\otimes N} \rangle = N \int_{\mathbb{R}^3} |\nabla u|^2 + \frac{N(N-1)}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \\ &\quad - N \sum_{i=1}^N \int_{\mathbb{R}^3} \frac{|u(x)|^2}{|x - R_k|} dx + \sum_{1 \leq \ell < k \leq N} \frac{1}{|R_\ell - R_k|}. \end{aligned}$$

Note that this upper bound holds for any choice of the nuclear positions  $\{R_k\}_{k=1}^N$ . Hence, we can average over  $\{R_k\}_{k=1}^N$ .

**First try.** Integrating the above variational inequality against

$$|u(R_1)|^2 \dots |u(R_N)|^2 dR_1 \dots dR_N, \quad R_k \in \mathbb{R}^3$$

we obtain

$$\begin{aligned} E_B(M, N) &\leq N \int_{\mathbb{R}^3} |\nabla u|^2 + \frac{N(N-1)}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \\ &\quad - N^2 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(R)|^2}{|x-R|} dx dR + \frac{N(N-1)}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(R_1)|^2 |u(R_2)|^2}{|R_1 - R_2|} dR_1 dR_2 \end{aligned}$$

$$= N \left( \int_{\mathbb{R}^3} |\nabla u|^2 - \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \right).$$

This leads to the upper bound  $E_b(M, N) \leq -CN$ .

**Second try.** We divide the support of  $u$  into  $N$  disjoint sets  $\{\Omega_k\}_{k=1}^N$  such that

$$\int_{\Omega_k} |u|^2 = \frac{1}{N}, \quad \forall k = 1, 2, \dots, N.$$

Then integrating the above variational inequality

$$\begin{aligned} E_b(M, N) &\leq N \int_{\mathbb{R}^3} |\nabla u|^2 + \frac{N(N-1)}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \\ &\quad - N \sum_{i=1}^N \int_{\mathbb{R}^3} \frac{|u(x)|^2}{|x-R_k|} dx + \sum_{1 \leq \ell < k \leq N} \frac{1}{|R_\ell - R_k|} \end{aligned}$$

against

$$\left( N|u(R_1)|^2 \right) \dots \left( N|u(R_N)|^2 \right) dR_1 \dots dR_N, \quad R_k \in \Omega_k$$

we obtain

$$\begin{aligned} E_B(M, N) &\leq N \int_{\mathbb{R}^3} |\nabla u|^2 + \frac{N(N-1)}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \\ &\quad - N^2 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(x)|^2 |u(R)|^2}{|x-R|} dx dR \\ &\quad + \frac{N^2}{2} \left( \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|u(R_1)|^2 |u(R_2)|^2}{|R_1 - R_2|} dR_1 dR_2 - \sum_{k=1}^N \int_{\Omega_k} \int_{\Omega_k} \frac{|u(R_1)|^2 |u(R_2)|^2}{|R_1 - R_2|} dR_1 dR_2 \right) \\ &\leq N \int_{\mathbb{R}^3} |\nabla u|^2 - N^2 \sum_{k=1}^N \int_{\Omega_k} \int_{\Omega_k} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy. \end{aligned}$$

We can choose the support of  $u$  like a cube  $[0, L]^3$ , and any set  $\Omega_k$  has diameter  $\sim L/N^{1/3}$ .

This gives

$$N^2 \int_{\Omega_k} \int_{\Omega_k} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \geq N^2 \int_{\Omega_k} \int_{\Omega_k} \frac{|u(x)|^2 |u(y)|^2}{(CL/N^{1/3})} dx dy = \frac{N^{1/3}}{CL}.$$

Moreover, if  $u$  behaves as a constant inside its support, then  $\|\nabla u\|_{L^2}^2 \sim L^{-2}$  (the spectral



gap of the Laplacian). Thus

$$\begin{aligned} E_B(M, N) &\leq N \int_{\mathbb{R}^3} |\nabla u|^2 - N^2 \sum_{k=1}^N \int_{\Omega_k} \int_{\Omega_k} \frac{|u(x)|^2 |u(y)|^2}{|x-y|} dx dy \\ &\leq C^{-1} \frac{N}{L^2} - C \frac{N^{4/3}}{L}. \end{aligned}$$

Optimizing over  $L > 0$  we obtain the desired upper bound

$$E_B(M, N) \leq -C^{-1} N^{5/3}.$$

□

In the above we have ignore the kinetic energy of the nuclei. The situation changes a bit when we consider nuclei with finite masses, however the instability remains. In the following let us consider the Hamiltonian

$$H_N = \sum_{i=1}^N (-\Delta_{x_i}) + \sum_{k=1}^N (-\Delta_{y_k}) - \sum_{i=1}^N \sum_{k=1}^N \frac{1}{|x_i - y_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq \ell < k \leq N} \frac{1}{|y_\ell - y_k|}.$$

on  $L^2(\mathbb{R}^{6N})$  (without the anti-symmetry assumption) and the ground state energy

$$E_B(N) = \inf_{\|\Psi\|_{L^2(\mathbb{R}^{6N})}=1} \langle \Psi, H_N \Psi \rangle.$$

**Theorem** ( $N^{7/5}$  instability). *We have*

$$\lim_{N \rightarrow \infty} \frac{E_B(N)}{N^{7/5}} = \inf_{\|u\|_{L^2(\mathbb{R}^3)}=1} \int_{\mathbb{R}^3} (2|\nabla u(x)|^2 - I_0 |u(x)|^{5/2}) dx$$

where

$$I_0 = \left(\frac{2}{\pi}\right)^{3/4} \int_0^\infty \frac{1}{1+t^4+t^2\sqrt{t^4+2}} dt = \frac{4^{5/4}\Gamma(3/4)}{5\pi^{1/4}\Gamma(4/5)}.$$

The upper bound  $-C^{-1}N^{7/5}$  was first proved by **Dyson** (1966) and the lower bound  $-CN^{7/5}$  was proved by **Conlon, Lieb, and Yau** (1988). The exact limit of  $E_B(N)/N^{7/5}$  was proved much later by **Lieb–Solovej** (2004, lower bound) and **Solovej** (2006, upper bound), based on a justification of **Bogoliubov’s theory**.

*Sketch of Solovej's proof of the sharp upper bound. Step 1.* First we work on Fock space where the number of particles is not fixed. We consider a trial state of the form

$$\Psi = W\mathbb{U}|0\rangle$$

with **Weyl unitary transformation**

$$W = e^{\sqrt{2N}(a^*(u)-a(u))}, \quad 0 \leq u \in L^2(\mathbb{R}^3), \quad \|u\|_{L^2} = 1$$

and a (bosonic) **Bogoliubov unitary transformation**  $\mathbb{U}$ . Note that applying Weyl's transformation to the vacuum gives us a **coherent state**

$$W|0\rangle = \sum_{n=0}^{\infty} e^{-N} \frac{(2N)^{n/2}}{n!} (a^*(u))^n |0\rangle$$

whose particle number expectation is exactly  $2N$ . On the other hand, applying Bogoliubov transformation to the vacuum gives us a **quasi-free state** which can be fully characterized by the one-body density matrix

$$\langle f, \gamma g \rangle = \langle 0 | \mathbb{U}^* a^*(g) a(f) \mathbb{U} | 0 \rangle.$$

The expectation of the particle number of  $\mathbb{U}|0\rangle$  is  $\text{Tr } \gamma$ . Thus the expectation of the total particle number of  $\Psi$  is  $2N + \text{Tr } \gamma$ . In our choice,  $\text{Tr } \gamma \leq CN^{3/5} \ll N$ .

Note that  $u$  stands for the Bose–Einstein condensation and  $\gamma$  stands for excited particles. In this **two–component charged Bose gas**, there is a huge cancellation in Hartree theory (due to the screening effect of Coulomb potential), and the contribution from the excited particles are crucial.

A lengthy but straightforward computation gives us the energy expectation of  $\Psi$ :

$$E = \langle \Psi, \bigoplus_{n=0}^{\infty} H_n \Psi \rangle = 2N \int_{\mathbb{R}^3} |\nabla u|^2 + \text{Tr}(-\Delta \gamma) + N \text{Tr} \left[ K(\gamma - \sqrt{\gamma(1+\gamma)}) \right]$$

where  $K$  is a non-negative Hilbert–Schmidt operator on  $L^2(\mathbb{R}^3)$  with kernel

$$K(x, y) = \frac{u(x)u(y)}{|x - y|}.$$

Finally, let us construct an explicit formula for  $\gamma$  using the **coherent state method**. Let  $G$  be non-negative even function with  $\|G\|_{L^2(\mathbb{R}^3)} = 1$  and denote

$$f_{k,y}(x) = e^{2\pi i k \cdot x} G(x - y).$$

Recall the coherent identity

$$\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |f_{k,y}\rangle \langle f_{k,y}| dk dy = \mathbf{1}_{L^2(\mathbb{R}^3)}.$$

Then we define

$$\gamma = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) |f_{k,y}\rangle \langle f_{k,y}| dk dy$$

with a suitable function  $M(k, y)$ . The **semiclassical approximation** to the energy

$$\text{Tr}(-\Delta\gamma) + N \text{Tr} \left[ K(\gamma - \sqrt{\gamma(1 + \gamma)}) \right]$$

is

$$\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \left( |2\pi k|^2 M(k, y) + \frac{N}{\pi|k|^2} |u(y)|^2 (M(k, y) - \sqrt{M(k, y)(1 + M(k, y))}) \right) dk dy$$

For example, we have

$$\begin{aligned} \text{Tr}[K\gamma] &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \langle f_{k,y}, K f_{k,y} \rangle dk dy \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \left( \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} e^{-2\pi i k \cdot x} G(x - y) \frac{u(x)u(z)}{|x - z|} e^{2\pi i k \cdot z} G(y - z) dx dz \right) dk dy \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \frac{e^{-2\pi i k \cdot (x-z)}}{|x - z|} G(x - y) u(x) u(z) G(y - z) dx dz dk dy \\ &\approx \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \frac{e^{-2\pi i k \cdot (x-z)}}{|x - z|} |G(x - y)|^2 |u(x)|^2 dx dz dk dy \quad (x \approx z) \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \frac{e^{-2\pi i k \cdot z}}{|z|} (G^2 * u^2)(y) dz dk dy \\ &\approx \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} M(k, y) \frac{1}{\pi|k|^2} |u(y)|^2 dk dy \quad (G^2 \approx \delta). \end{aligned}$$

Optimizing over  $M(k, y)$  the latter expression we get

$$E \approx 2N \int_{\mathbb{R}^3} |\nabla u|^2 - N^{5/4} I_0 \int_{\mathbb{R}^3} |u|^{5/2}.$$

We can take

$$u(x) = \ell^{3/2} u_0(\ell x)$$

with a normalized function  $u_0 \in L^2(\mathbb{R}^3)$ , and obtain

$$E \approx 2N\ell^2 \int_{\mathbb{R}^3} |\nabla u_0|^2 - N^{5/4} \ell^{3/4} I_0 \int_{\mathbb{R}^3} |u_0|^{5/2}$$

(the power  $3/4$  comes from  $\frac{3}{2} \cdot \frac{5}{2} - 3$ ). We choose  $\ell$  such that

$$N\ell^2 = N^{5/4} \ell^{3/4} \iff \ell = N^{1/5}.$$

Thus

$$E \approx N^{7/5} \left( 2 \int_{\mathbb{R}^3} |\nabla u_0|^2 - I_0 \int_{\mathbb{R}^3} |u_0|^{5/2} \right).$$

Optimizing over  $u_0$  gives us the desired upper bound.

**Step 2.** It remains to find a trial state in the  $2N$ -particle sector. We can repeat the choice of the trial state

$$\Psi = WU|0\rangle$$

but now we use the **Weyl unitary transformation**

$$W = e^{\sqrt{M}(a^*(u) - a(u))}, \quad 0 \leq u \in L^2(\mathbb{R}^3), \quad \|u\|_{L^2} = 1$$

with  $M = 2N - CN^{3/5}$ , so that

$$\langle \Psi, \mathcal{N}\Psi \rangle = M + \text{Tr } \gamma \leq 2N - N^{3/5}.$$

Recall that with our choice of  $\gamma$ , the number of excited particles is  $\text{Tr } \gamma = O(N^{3/5})$ . Moreover, we have

$$\langle \Psi, \mathcal{N}^2\Psi \rangle - \langle \Psi, \mathcal{N}\Psi \rangle^2 = \left\langle \Psi, \left( \mathcal{N} - \langle \Psi, \mathcal{N}\Psi \rangle \right)^2 \Psi \right\rangle \leq CN.$$

Similarly to Step 1, we have

$$E = \left\langle \Psi, \bigoplus_{n=0}^{\infty} H_n \right\rangle \leq N^{7/5} \inf_{\|u\|_{L^2(\mathbb{R}^3)}=1} \int_{\mathbb{R}^3} (2|\nabla u(x)|^2 - I_0|u(x)|^{5/2}) dx + o(N^{7/5}).$$

At the moment,  $\Psi$  is still a state on Fock space. To go to fixed particle sectors, we denote

$$\Psi = (\Psi_n)_{n=0}^\infty, \quad \Psi_n \in L^2(\mathbb{R}^{3n}), \quad \sum_{n=0}^\infty \|\Psi_n\|^2 = 1.$$

Then

$$E = \left\langle \Psi, \bigoplus_{n=0}^\infty H_n \right\rangle = \sum_{n=0}^\infty \langle \Psi_n, H_n \Psi_n \rangle \geq \sum_{n=0}^\infty \|\Psi_n\|^2 E_B(n).$$

To conclude, we use the fact that  $n \mapsto E_B(n)$  is decreasing and that  $0 \geq E_B(n) \geq -Cn^{7/5}$  (we do not need a sharp lower bound here). Then

$$\begin{aligned} E_B(N) &\leq \sum_{n \leq 2N} \|\Psi_n\|^2 E_B(N) \leq \sum_{n \leq 2N} \|\Psi_n\|^2 E_B(n) \\ &= \sum_{n=0}^\infty \|\Psi_n\|^2 E_B(n) - \sum_{n > 2N} \|\Psi_n\|^2 E_B(n) \leq E + C \sum_{n > 2N} \|\Psi_n\|^2 n^{7/5}. \end{aligned}$$

The error term with  $n > 2N \geq \langle \Psi, \mathcal{N}\Psi \rangle + N^{4/5}$  can be estimated by the variance bound and the Cauchy–Schwarz inequality

$$\begin{aligned} \sum_{n > 2N} \|\Psi_n\|^2 n^{7/5} &\leq \sum_{n > 2N} \|\Psi_n\|^2 n^{7/5} \left( \frac{|n - \langle \Psi, \mathcal{N}\Psi \rangle|}{N^{3/5}} \right)^{3/5} \\ &\leq \left( \sum_{n=0}^\infty \|\Psi_n\|^2 n^2 \right)^{7/10} \left( \sum_{n=0}^\infty \|\Psi_n\|^2 \left( \frac{|n - \langle \Psi, \mathcal{N}\Psi \rangle|}{N^{3/5}} \right)^2 \right)^{3/10} \\ &= (\langle \Psi, \mathcal{N}^2 \Psi \rangle)^{7/10} \left( \frac{\left\langle \Psi, (\mathcal{N} - \langle \Psi, \mathcal{N}\Psi \rangle)^2 \Psi \right\rangle}{N^{6/5}} \right)^{3/10} \\ &\leq C(N^2)^{7/10} \left( \frac{N}{N^{6/5}} \right)^{3/10} = CN^{\frac{7}{5} - \frac{3}{50}} \ll N^{7/5}. \end{aligned}$$

Thus we conclude that

$$E_B(N) \leq E + o(N^{7/5}) \leq N^{7/5} \inf_{\|u\|_{L^2(\mathbb{R}^3)}=1} \int_{\mathbb{R}^3} (2|\nabla u(x)|^2 - I_0|u(x)|^{5/2}) dx + o(N^{7/5}).$$

This is the desired upper bound in the  $2N$ -particle sector.  $\square$