Chapter 11

Introduction to the Calculus of Variations

11.1 INTRODUCTION

As an illustration of the use of the mathematics developed in this book, we give three additional examples (beyond those of Chapter 4) of solving optimization problems. The first comes from quantum mechanics and is the problem of determining the energy of an atom—primarily the lowest one. The second is a classical type minimization problem—the Thomas–Fermi problem—that arises in chemistry. The third is a problem in electrostatics, namely the capacitor problem. In all cases the difficult part is showing the existence of a minimizer, and hence of a solution to a partial differential equation. Needless to say, the following considerations (known as the direct method in the calculus of variations) for establishing a solution to a differential equation are not limited to these elementary examples, but should be viewed as a general strategy to attack optimization problems.

Historically, and even today in many places, it is customary to dispense with the question of existence as a mere subtlety. By simply assuming that a minimizer or maximizer exists, however, and then trying to derive properties for it, one can be led to severe inconsistencies—as the following amusing example taken from [L. C. Young] and attributed to Perron shows: “Let \( N \) be the largest natural number. Since \( N^2 \geq N \) and \( N \) is the largest natural number, \( N^2 = N \) and hence \( N = 1 \).” What this example tells us is that
even if the ‘variational equation’, here $N^2 = N$, can be solved explicitly, the resulting solution need not have anything to do with the problem we started out to solve.

Let us continue this overview with some general remarks about minimization of functions. A general theorem in analysis says that a bounded continuous real function $f$ defined on a bounded and closed set $K$ in $\mathbb{R}^n$ attains its minimum value. To prove this, pick a sequence of points $x^j$ such that

$$f(x^j) \to \lambda := \inf_{x \in K} f(x) \quad \text{as } j \to \infty.$$ 

Since $K$ is bounded and closed, there exists a subsequence, again denoted by $x^j$, and a point $x \in K$ such that $x^j \to x$ as $j \to \infty$. Hence, since $f$ is continuous,

$$\lambda := \lim_{j \to \infty} f(x^j) = f(x),$$

and the minimum value is attained at $x$.

Instead of $\mathbb{R}^n$, consider now $L^2(\Omega, d\mu)$ and let $\mathcal{F}(\psi)$ be some functional defined on this space. In many examples $\mathcal{F}(\psi)$ is strongly continuous, i.e., $\mathcal{F}(\psi^j) \to \mathcal{F}(\psi)$ as $j \to \infty$ whenever $\|\psi^j - \psi\|_2 \to 0$ as $j \to \infty$. Suppose we wish to show that the infimum of $\mathcal{F}(\psi)$ is attained on $K := \{\psi \in L^2(\Omega, d\mu) : \|\psi\|_2 \leq 1\}$. This set is certainly closed and bounded, but for a bounded sequence $\psi^j \in K$ there need not be a strongly convergent subsequence (see Sect. 2.9).

The idea now is to relax the strength of convergence. Indeed, if we use the notion of weak convergence instead of strong convergence, then, by Theorem 2.18 (bounded sequences have weak limits), every sequence in $K$ has a weakly convergent subsequence. In this way, the set of convergent sequences has been enlarged—but a new problem arises. The functional $\mathcal{F}(\psi)$ need not be weakly continuous—and it rarely is. Thus, to summarize, the more sequences exist that have convergent subsequences the less likely it is that $\mathcal{F}(\psi)$ is continuous on these sequences. The way out of this apparent dilemma is that in many examples the functional turns out to be weakly lower semicontinuous, i.e.,

$$\liminf_{j \to \infty} \mathcal{F}(\psi^j) \geq \mathcal{F}(\psi) \quad \text{if } \psi^j \rightharpoonup \psi \text{ weakly.}$$

Thus, if $\psi^j$ is a minimizing sequence, i.e., if

$$\mathcal{F}(\psi^j) \to \inf\{\mathcal{F}(\psi) : \psi \in C\} = \lambda,$$

then there exists a subsequence $\psi^{j_k}$ such that $\psi^{j_k} \rightharpoonup \psi$ weakly, and hence

$$\lambda = \lim_{j \to \infty} \mathcal{F}(\psi^j) \geq \mathcal{F}(\psi) \geq \lambda.$$

Therefore, $\mathcal{F}(\psi) = \lambda$, and our goal is achieved!
11.2 SCHRODINGER'S EQUATION

The time independent Schrödinger equation [Schrödinger] for a particle in $\mathbb{R}^n$, interacting with a force field $F(x) = -\nabla V(x)$, is

$$-\Delta \psi(x) + V(x)\psi(x) = E\psi(x). \quad (1)$$

The function $V : \mathbb{R}^n \to \mathbb{R}$ is called a potential (not to be confused with the potentials in Chapter 9). The wave function $\psi$ is a complex-valued function in $L^2(\mathbb{R}^n)$ subject to the normalization condition

$$\|\psi\|_2 = 1. \quad (2)$$

The function $\rho_\psi(x) = |\psi(x)|^2$ is interpreted as the probability density for finding the particle at $x$. An $L^2(\mathbb{R}^n)$ solution to (1) may or may not exist for any $E$; often it does not. The special real numbers $E$ for which such solutions exist are called eigenvalues and the solution, $\psi$, is called an eigenfunction.

Associated with (1) is a variational problem. Consider the following functional defined for a suitable class of functions in $L^2(\mathbb{R}^n)$ (to be specified later):

$$\mathcal{E}(\psi) = T_\psi + V_\psi,$$

with

$$T_\psi = \int_{\mathbb{R}^n} |\nabla \psi(x)|^2 \, dx \quad \text{and} \quad V_\psi = \int_{\mathbb{R}^n} V(x)|\psi(x)|^2 \, dx. \quad (4)$$

Physically, $T_\psi$ is called the kinetic energy of $\psi$, $V_\psi$ is its potential energy and $\mathcal{E}(\psi)$ is the total energy of $\psi$.

The variational problem we shall consider is to minimize $\mathcal{E}(\psi)$ subject to the constraint $\|\psi\|_2 = 1$.

As we shall show in Sect. 11.5, a minimizing function $\psi_0$, if one exists, will satisfy equation (1) with $E = E_0$, where

$$E_0 := \inf\{\mathcal{E}(\psi) : \int |\psi|^2 = 1\}.$$ 

Such a function $\psi_0$ will be called a ground state. $E_0$ is called the ground state energy.\(^1\)

Thus the variational problem determines not only $\psi_0$ but also a corresponding eigenvalue $E_0$, which is the smallest eigenvalue of (1).

\(^1\)Physically, the ground state energy is the lowest possible energy the particle can attain. It is a physical fact that the particle will settle eventually into its ground state, by emitting energy, usually in the form of light.
Our route to finding a solution to (1) takes us to the main problem:
Show, under suitable assumptions on $V$, that a minimizer exists, i.e., show that there exists a $\psi_0$ satisfying (2) and such that
\[ E(\psi_0) = \inf\{ E(\psi) : \|\psi\|_2 = 1 \}. \]

There are examples where a minimizer does not exist, e.g., take $V$ to be identically zero.

In Sect. 11.5 we shall prove, under suitable assumptions on $V$, the existence of a minimizer for $E(\psi)$. We shall also solve the corresponding relativistic problem, in which the kinetic energy is given by $(\psi, |p|\psi)$ instead, as defined in Sect. 7.11. In the nonrelativistic case ((1), (4)) it will be shown that the minimizers satisfy (1) in the sense of distributions. Higher eigenvalues will be explained in Sect. 11.6. The content of Sect. 11.7 is an application of the results of Chapter 10 to show that under suitable additional assumptions on $V$, the distributional solutions of (1) are sufficiently regular to yield classical solutions, i.e., solutions that are twice continuously differentiable.

A final question concerns uniqueness of the minimizer. In our Schrödinger example, $E(\psi)$, uniqueness means that the ground state solution to (1) is unique, apart from an 'overall phase', i.e., $\psi_0(x) \rightarrow e^{i\theta} \psi_0(x)$ for some $\theta \in \mathbb{R}$. That uniqueness of the minimizer (proved in Theorem 11.8) implies uniqueness of the solution to (1) with $E = E_0$ is not totally obvious; it is proved in Corollary 11.9. The tool that will enable us to prove uniqueness of the minimizer is the strict convexity of the map $\rho \rightarrow E(\sqrt{\rho})$ for strictly positive functions $\rho : \mathbb{R}^n \rightarrow \mathbb{R}^+$. (See Theorem 7.8 (convexity inequality for gradients).) The hard part is to establish the strict positivity of a minimizer. Theorem 9.10 (lower bounds on Schrödinger wave functions) will be crucial here.

11.3 DOMINATION OF THE POTENTIAL ENERGY BY THE KINETIC ENERGY

Recall that the functional to consider is
\[ E(\psi) = \int_{\mathbb{R}^n} |\nabla \psi(x)|^2 \, dx + \int_{\mathbb{R}^n} V(x)|\psi(x)|^2 \, dx, \]
and the ground state energy $E_0$ is
\[ E_0 = \inf\{ E(\psi) : \|\psi\|_2 = 1 \}. \tag{1} \]

The kinetic energy is defined for any function in $H^1(\mathbb{R}^n)$ and the second term is defined at least for $\psi \in C_c^\infty(\mathbb{R}^n)$ if we assume that $V \in L^1_{\text{loc}}(\mathbb{R}^n)$. The first
necessary condition for a minimizer to exist is that $E(\psi)$ is bounded below by some constant independent of $\psi$ (when $\|\psi\|_2 \leq 1$). The reader can imagine that when, e.g., $V(x) = -|x|^{-3}$, then $E(\psi)$ is no longer bounded below. Indeed, for any $\psi \in C_0^\infty(\mathbb{R}^n)$ with $\|\psi\|_2 = 1$ and $\int V(x)|\psi(x)|^2 \, dx < \infty$, define $\psi_\lambda(x) = \lambda^{n/2}\psi(\lambda x)$ and observe that $\|\psi_\lambda\|_2 = 1$. One easily computes

$$E(\psi_\lambda) = \lambda^2 \int_{\mathbb{R}^n} |\nabla \psi(x)|^2 \, dx - \lambda^3 \int_{\mathbb{R}^n} V(x)|\psi(x)|^2 \, dx.$$ 

Clearly, $E(\psi_\lambda) \to -\infty$ as $\lambda \to \infty$. One sees from this example that the assumptions on $V$ must be such that $V_\psi$ can be bounded below in terms of the kinetic energy $T_\psi$ and the norm $\|\psi\|_2$.

Any inequality in which the kinetic energy $T_\psi$ dominates some kind of integral of $\psi$ (but not involving $\nabla \psi$) is called an uncertainty principle. The historical reason for this strange appellation is that such an inequality implies that one cannot make the potential energy very negative without also making the kinetic energy large, i.e., one cannot localize a particle simultaneously in both $\mathbb{R}^n$ and the Fourier transform copy of $\mathbb{R}^n$. The most famous uncertainty principle, historically, is Heisenberg’s: In $\mathbb{R}^n$

$$(\psi, p^2 \psi) \geq \frac{n^2}{4} (\psi, x^2 \psi)^{-1}$$

for $\psi \in H^1(\mathbb{R}^n)$ and $\|\psi\|_2 = 1$. The proof of this inequality (which uses the fact that $\nabla \cdot x - x \cdot \nabla = n$) can be found in many textbooks and we shall not give it here because (2) is not actually very useful. Knowledge of $(\psi, x^2 \psi)$ tells us little about $T_\psi$. The reason for this is that any $\psi$ can easily be modified in an arbitrarily small way (in the $H^1(\mathbb{R}^n)$-norm) so that $\psi$ concentrates somewhere, i.e., $(\psi, p^2 \psi)$ is not small, but $(\psi, x^2 \psi)$ is huge. To see this, take any fixed function $\psi$ and then replace it by $\psi_\varepsilon(x) = \sqrt{1 - \varepsilon^2} \psi(x) + \varepsilon \psi(x-y)$ with $\varepsilon \ll 1$ and $|y| \gg 1$. To a very good approximation, $\psi_\varepsilon = \psi$ but, as $|y| \to \infty$, $\|\psi_\varepsilon\|_2 \to 1$ and $(\psi_\varepsilon, x^2 \psi_\varepsilon) \to \infty$. Thus, the right side of (2) goes to zero as $|y| \to \infty$ while $T_{\psi_\varepsilon} \approx T_\psi$ does not go to zero.

Sobolev’s inequality (see Sects. 8.3 and 8.5) is much more useful in this respect. Recall that for functions that vanish at infinity on $\mathbb{R}^n$, with $n \geq 3$, there are constants $S_n$ such that

$$T_\psi \geq S_n \left\{ \int_{\mathbb{R}^n} |\psi(x)|^{2n/(n-2)} \, dx \right\}^{(n-2)/n} = S_n \|\rho_\psi\|_{n/(n-2)}$$

for $n = 3$.

For $n = 1$ and $n = 2$, on the other hand, we have

$$T_\psi + \|\psi\|_2^2 \geq S_n, \rho_\psi \|_p \quad \text{for all } 2 \leq p < \infty, \quad n = 2,$$

$$T_\psi + \|\psi\|_2^2 \geq S_1 \|\rho_\psi\|_\infty, \quad n = 1.$$
Moreover, when \( n = 1 \) and \( \psi \in H^1(\mathbb{R}^1) \), \( \psi \) is not only bounded, it is also continuous.

An application of Hölder’s inequality to (3) yields, for any potential \( V \in L^{n/2}(\mathbb{R}^n) \), \( n \geq 3 \),

\[
T_\psi \geq S_n \| \rho_\psi \|_{n/(n-2)} \geq S_n (\psi, V) \| V \|_{n/2}^{-1}.
\]

An immediate application of (6) is that

\[
T_\psi + V_\psi \geq 0
\]

whenever \( \|V\|_{n/2} \leq S_n \).

A simple extension of (6) leads to a lower bound on the ground state energy for \( V \in L^{n/2}(\mathbb{R}^n) + L^\infty(\mathbb{R}^n) \), \( n \geq 3 \), i.e., for \( V \)'s that satisfy

\[
V(x) = v(x) + w(x)
\]

for some \( v \in L^{n/2}(\mathbb{R}^n) \) and \( w \in L^\infty(\mathbb{R}^n) \). There is then some constant \( \lambda \) such that \( h(x) := -(v(x) - \lambda)_- = \min(v(x) - \lambda, 0) \leq 0 \) satisfies \( \|h\|_{n/2} \leq \frac{1}{2} S_n \) (exercise for the reader). In particular, by (6), \( h_\psi \geq -\frac{1}{2} T_\psi \). Then we have

\[
\mathcal{E}(\psi) = T_\psi + V_\psi = T_\psi + (v - \lambda)_\psi + \lambda + w_\psi \\
\geq T_\psi + h_\psi + \lambda + w_\psi \geq \frac{1}{2} T_\psi + \lambda - \|w\|_\infty
\]

and we see that \( \lambda - \|w\|_\infty \) is a lower bound to \( E_0 \). Furthermore (9) implies that the total energy effectively bounds the kinetic energy, i.e., we have that

\[
T_\psi \leq 2(\mathcal{E}(\psi) - \lambda + \|w\|_\infty).
\]

When \( n = 2 \), the preceding argument, together with (4), gives a finite \( E_0 \) whenever \( V \in L^p(\mathbb{R}^2) + L^\infty(\mathbb{R}^2) \) for any \( p > 1 \). Likewise, when \( n = 1 \) we can conclude that \( E_0 \) is finite whenever \( V \in L^1(\mathbb{R}^1) + L^\infty(\mathbb{R}^1) \). In fact, a bit more can be deduced when \( n = 1 \). Since \( \psi \in H^1(\mathbb{R}^1) \) implies that \( \psi \) is continuous, it makes sense to define \( \int \psi(x) \mu(dx) \) when \( \mu = \mu_1 - \mu_2 \) and when \( \mu_1 \) and \( \mu_2 \) are any bounded, positive Borel measures on \( \mathbb{R}^1 \). (‘Bounded’ means that \( \int \mu(dx) < \infty \).) A well-known example in the physics literature is \( \mu(dx) = c \delta(x) dx \) where \( \delta(x) \) is Dirac’s ‘delta function’. More precisely, \( \int \psi(x) \mu(dx) = c \psi(0) \). Then we can define

\[
\mathcal{E}(\psi) = T_\psi + \int_{\mathbb{R}^n} |\psi(x)|^2 \mu(dx)
\]
and then (5) et seq. imply that $E_{0}$, defined as before, is finite. In short, in one dimension a "potential" can be a bounded measure plus an $L^{\infty}(\mathbb{R})$-function.

So far we have considered the nonrelativistic kinetic energy $T_{\psi} = (\psi, p^{2}\psi)$. Similar inequalities hold for the relativistic case $T_{\psi} = (\psi, |p|\psi)$. The relativistic analogues of (3)–(5) are (12) and (13) below (see Sects. 8.4 and 8.5). There are constants $S'_{n}$ for $n \geq 2$ and $S'_{1,p}$ for $2 \leq p < \infty$ such that

$$T_{\psi} \geq S'_{n}\|\rho_{\psi}\|_{n/(n-1)}, \quad n \geq 2,$$

and $S'_{3} = 2^{1/3}\pi^{2/3}$. When $n = 1$,

$$T_{\psi} + \|\psi\|_{2}^{2} \geq S'_{1,p}\|\rho_{\psi}\|_{p} \quad \text{for all} \quad 2 \leq p < \infty, \quad n = 1. \quad (13)$$

The results of this section can be summarized in the following statement.

In all dimensions $n \geq 1$, the hypothesis that $V$ is in the space

$$\begin{align*}
\text{nonrelativistic} & \quad \begin{cases}
L^{n/2}(\mathbb{R}^{n}) + L^{\infty}(\mathbb{R}^{n}), & n \geq 3, \\
L^{1+\varepsilon}(\mathbb{R}^{2}) + L^{\infty}(\mathbb{R}^{2}), & n = 2, \\
L^{1}(\mathbb{R}^{1}) + L^{\infty}(\mathbb{R}^{1}), & n = 1,
\end{cases} \\
\text{relativistic} & \quad \begin{cases}
L^{n}(\mathbb{R}^{n}) + L^{\infty}(\mathbb{R}^{n}), & n \geq 2, \\
L^{1+\varepsilon}(\mathbb{R}^{1}) + L^{\infty}(\mathbb{R}^{1}), & n = 1,
\end{cases}
\end{align*}$$

leads to the following two conclusions:

$$E_{0} \text{ is finite}, \quad (16)$$

$$T_{\psi} \leq CE(\psi) + D\|\psi\|_{2}^{2} \quad (17)$$

when $\psi \in H^{1}(\mathbb{R}^{n})$ (nonrelativistic), or $\psi \in H^{1/2}(\mathbb{R}^{n})$ (relativistic), for suitable constants $C$ and $D$. Furthermore, in the nonrelativistic case in one-dimension, $V$ can be generalized to be a bounded Borel measure.

The existence of minimum energy—or ground state—functions will be proved for the one-body problem under fairly weak assumptions. The principal ingredients are the Sobolev inequality (Theorems 8.3–8.5), and the Rellich–Kondrashov theorem (Theorems 8.7, 8.9). The following definition is convenient:

$$H^{\#}(\mathbb{R}^{n}) \text{ denotes } \begin{cases}
H^{1}(\mathbb{R}^{n}) \text{ in the nonrelativistic case,} \\
H^{1/2}(\mathbb{R}^{n}) \text{ in the relativistic case.}
\end{cases}$$

The main technical result is the following theorem.
11.4 THEOREM (Weak continuity of the potential energy)

Let $V(x)$ be a function on $\mathbb{R}^n$ that satisfies the condition given in 11.3(14) (nonrelativistic case) or 11.3(15) (relativistic case). Assume, in addition, that $V(x)$ vanishes at infinity, i.e.,

$$|\{x : |V(x)| > a\}| < \infty \quad \text{for all } a > 0.$$ 

If $n = 1$ in the nonrelativistic case, $V$ can be the sum of a bounded Borel measure and an $L^\infty(\mathbb{R})$-function $\omega$ that vanishes at infinity. Then $V_\psi$, defined in 11.2(4), is weakly continuous in $H^\#(\mathbb{R}^n)$, i.e., if $\psi^j \rightharpoonup \psi$ as $j \to \infty$, weakly in $H^\#(\mathbb{R}^n)$, then $V_{\psi^j} \rightharpoonup V_\psi$ as $j \to \infty$.

**PROOF.** Note that by Theorem 2.12 (uniform boundedness principle) $\|\psi^j\|_{H^\#}$ is uniformly bounded. First, assume that $V$ is a function.

Define $V^\delta$ (when $V$ is a function) by

$$V^\delta(x) = \begin{cases} 
V(x) & \text{if } |V(x)| \leq 1/\delta, \\
0 & \text{if } |V(x)| > 1/\delta,
\end{cases}$$

and note that $V - V^\delta$ tends to zero as $\delta \to 0$ (by dominated convergence) in the appropriate $L^p(\mathbb{R}^n)$ norm of 11.3(14), resp. 11.3(15). Since $\|\psi^j\|_{H^\#} \leq t$, Theorems 8.3–8.5 (Sobolev’s inequality) imply that

$$\int (V - V^\delta)|\psi^j|^2 < C_\delta,$$

with $C_\delta$ independent of $j$ and, moreover, $C_\delta \to 0$ as $\delta \to 0$. Thus, our goal of showing that $V_{\psi^j} \rightharpoonup V_\psi$ as $j \to \infty$ will be achieved if we can prove that $V_{\psi^j}^\delta \rightharpoonup V_\psi^\delta$ as $j \to \infty$ for each $\delta > 0$. If $n = 1$ and $V$ is a measure, then $V^\delta$ is simply taken to be $V$ itself.

The problem in showing that $V_{\psi^j}^\delta \rightharpoonup V_\psi^\delta$ as $j \to \infty$ comes from the fact that $V^\delta$ is known to vanish at infinity only in the weak sense. Fix $\delta$ and define the set

$$A_\varepsilon = \{x : |V^\delta(x)| > \varepsilon\}$$

for $\varepsilon > 0$. By assumption, $|A_\varepsilon| < \infty$. Then

$$V_{\psi^j}^\delta = \int_{A_\varepsilon} V^\delta|\psi^j|^2 + \int_{A_\varepsilon^C} V^\delta|\psi^j|^2.$$ \hspace{1cm} (1)

The last term is not greater than $\varepsilon \int |\psi^j|^2 = \varepsilon$ (independent of $j$), and hence (since $\varepsilon$ is arbitrary) it suffices to show that the first term in (1) converges, for a subsequence of $\psi^j$’s, to $\int_{A_\varepsilon} V^\delta|\psi|^2$. 

This is accomplished as follows. By Theorem 8.6 (weak convergence implies strong convergence on small sets), on any set of finite measure (that we take to be \( A_x \)) there is a subsequence (which we continue to denote by \( \psi^j \)) such that \( \psi^j \to \psi \) strongly in \( L^r(A_x) \). Here \( 2 \leq r < p \). The reader is invited to check, by using the inequality

\[
||\psi^j|^2 - |\psi|^2| \leq |\psi^j - \psi||\psi^j + \psi|,
\]

that \( |\psi^j|^2 \to |\psi|^2 \) strongly in \( L^{r/2}(A_x) \). Since \( V^\delta \in L^{\infty}(\mathbb{R}^n) \), we have that \( V^\delta \in L^s(A_x) \) for all \( 1 \leq s \leq \infty \). Thus, by taking \( 1/s + 2/r = 1 \), our claim is proved. When \( n = 1 \) we leave it to the reader to check that \( \psi^j(x) \to \psi(x) \) uniformly on bounded intervals in \( \mathbb{R}^1 \), and hence that the same proof goes through in the nonrelativistic case when \( V \) is a bounded measure plus an \( L^\infty(\mathbb{R}^1) \)-function.

11.5 THEOREM (Existence of a minimizer for \( E_0 \))

Let \( V(x) \) be a function on \( \mathbb{R}^n \) that satisfies the condition given in 11.3(14) (nonrelativistic case) or 11.3(15) (relativistic case). Assume that \( V(x) \) vanishes at infinity, i.e.,

\[
|\{x : |V(x)| > a\}| < \infty \quad \text{for all } a > 0.
\]

When \( n = 1 \) in the nonrelativistic case \( V \) can be the sum of a bounded measure and a function \( w \in L^{\infty}(\mathbb{R}) \) that vanishes at infinity. Let \( E(\psi) = T_\psi + V_\psi \) as before and assume that

\[
E_0 = \inf\{E(\psi) : \psi \in H^\#(\mathbb{R}^n), \|\psi\|_2 = 1\} < 0.
\]

By 11.3(16), \( E(\psi) \) is bounded from below when \( \|\psi\|_2 = 1 \).

Our conclusion is that there is a function \( \psi_0 \) in \( H^\#(\mathbb{R}^n) \) such that \( \|\psi_0\|_2 = 1 \) and

\[
E(\psi_0) = E_0.
\]  

(We shall see in Sect. 11.8 that \( \psi_0 \) is unique up to a factor and can be chosen to be positive.) Furthermore, any minimizer \( \psi_0 \) satisfies the Schrödinger equation in the sense of distributions:

\[
H_0 \psi_0 + V \psi_0 = E_0 \psi_0,
\]

where \( H_0 = -\Delta \) (nonrelativistic) and \( H_0 = (-\Delta + m^2)^{1/2} - m \) (relativistic). Note that (2) implies that the function \( V \psi_0 \) is also a distribution; this implies that \( V \psi_0 \in L^1_{\text{loc}}(\mathbb{R}^n) \).
REMARKS. (1) From (2) we see that the distribution \((H_0 + V)\psi_0\) is always a function (namely \(E_0\psi_0\)). This is true in the nonrelativistic case when \(n = 1\), even when \(V\) is a measure!

(2) Theorem 11.5 states that a minimizer satisfies the Schrödinger equation (2). Suppose, on the other hand, that \(\psi\) is some function in \(H^\#(\mathbb{R}^n)\) that satisfies (2) in \(\mathcal{D}'\), but with \(E_0\) replaced by some real number \(E\). Can we conclude that \(E \geq E_0\) and, moreover, that \(E = E_0\) if and only if \(\psi\) is a minimizer? The answer is yes and we invite the reader to prove this by taking a sequence \(\phi^j \in C_c^\infty(\mathbb{R}^n)\) that converges to \(\psi\) as \(j \to \infty\) and testing (2) with this sequence. By taking the limit \(j \to \infty\), one can easily justify the equality \(\mathcal{E}(\psi) = E\|\psi\|_2^2\). The stated conclusion follows immediately.

PROOF. Let \(\psi^j\) be a minimizing sequence, i.e., \(\mathcal{E}(\psi^j) \to E_0\) as \(j \to \infty\) and \(\|\psi^j\|_2 = 1\). First we note that by 11.3(17) \(T_{\psi^j}\) is bounded by a constant independent of \(j\). Since \(\|\psi^j\|_2 = 1\), the sequence \(\psi^j\) is bounded in \(H^\#(\mathbb{R}^n)\). Since bounded sets in \(H^{1/2}(\mathbb{R}^n)\) and \(H^1(\mathbb{R}^n)\) are weakly sequentially compact (see Sect. 7.18), we can therefore find a function \(\psi_0\) in \(H^\#(\mathbb{R}^n)\) and a subsequence (which we continue to denote by \(\psi^j\)) such that \(\psi^j \rightharpoonup \psi_0\) weakly in \(H^\#(\mathbb{R}^n)\). The weak convergence of \(\psi^j\) to \(\psi_0\) implies that \(\|\psi_0\|_2 \leq 1\). This function \(\psi_0\) will be our minimizer as we shall show. Note that, since the kinetic energy \(T_{\psi}\) is weakly lower semicontinuous (see the end of Sect. 8.2), and since, by Theorem 11.4, \(V\) is weakly continuous in \(H^\#(\mathbb{R}^n)\), we have that \(\mathcal{E}(\psi)\) is weakly lower semicontinuous on \(H^\#(\mathbb{R}^n)\). Hence

\[
E_0 = \lim_{j \to \infty} \mathcal{E}(\psi^j) \geq \mathcal{E}(\psi_0)
\]

and \(\psi_0\) is a minimizer provided we know that \(\|\psi_0\|_2 = 1\). By assumption however,

\[
0 > E_0 \geq \mathcal{E}(\psi_0) \geq E_0\|\psi_0\|_2^2.
\]

The last inequality holds by the definition of \(E_0\) and, since \(E_0 < 0\), it follows that \(\|\psi_0\|_2 = 1\). This shows the existence of a minimizer.

To prove that \(\psi_0\) satisfies the Schrödinger equation (2) we take any function \(f \in C_c^\infty(\mathbb{R}^n)\) and set \(\psi^\varepsilon := \psi_0 + \varepsilon f\) for \(\varepsilon \in \mathbb{R}\). The quotient \(\mathcal{R}(\varepsilon) = \mathcal{E}(\psi^\varepsilon)/(\psi^\varepsilon, \psi^\varepsilon)\) is clearly the ratio of two second degree polynomials in \(\varepsilon\) and hence differentiable for small \(\varepsilon\). Since its minimum, \(E_0\), occurs (by assumption) at \(\varepsilon = 0\), \(d\mathcal{R}(\varepsilon)/d\varepsilon = 0\) at \(\varepsilon = 0\). This yields

\[
\left. \frac{d\mathcal{E}(\psi^\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = E_0 \left. \frac{d(\psi^\varepsilon, \psi^\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0},
\]

which implies that

\[
((H_0 + V)f, \psi_0) = E_0(f, \psi_0)
\]
The next theorem is an extension of Theorem 11.5 to higher eigenvalues and eigenfunctions. The ground state energy $E_0$ is the first eigenvalue with $\psi_0$ as the first eigenfunction. Since $\mathcal{E}(\psi)$ is a quadratic form, we can try to minimize it over $\psi$ in $H^1(\mathbb{R}^n)$ (resp. $H^{1/2}(\mathbb{R}^n)$ in the relativistic case) under the two constraints that $\psi$ is normalized and $\psi$ is orthogonal to $\psi_0$, i.e.,

$$
\langle \psi, \psi_0 \rangle = \int_{\mathbb{R}^n} \overline{\psi(x)} \psi_0(x) \, dx = 0.
$$

This infimum we call $E_1$, the second eigenvalue, and, if it is attained, we call the corresponding minimizer, $\psi_1$, the first excited state or second eigenfunction. In a similar fashion we can define the $(k+1)^{th}$ eigenvalue recursively (under the assumption that the first $k$ eigenfunctions $\psi_0, \ldots, \psi_{k-1}$ exist)

$$
E_k := \inf \{ \mathcal{E}(\psi) : \psi \in H^1(\mathbb{R}^n), \| \psi \|_2 = 1 \text{ and } \langle \psi, \psi_i \rangle = 0, \ i = 0, \ldots, k-1 \}.
$$

$H^1(\mathbb{R}^n)$ has to be replaced by $H^{1/2}(\mathbb{R}^n)$ in the relativistic case.

In the physical context these eigenvalues have an important meaning in that their differences determine the possible frequencies of light emitted by a quantum-mechanical system. Indeed, it was the highly accurate experimental verification of this fact for the case of the hydrogen atom (see Sect. 11.10) that overcame most of the opposition to the radical idea of the quantum theory.

### 11.6 THEOREM (Higher eigenvalues and eigenfunctions)

Let $V$ be as in Theorem 11.5 and assume that the $(k+1)^{th}$ eigenvalue $E_k$ given above is negative. (This includes the assumption that the first $k$ eigenfunctions exist.) Then the $(k+1)^{th}$ eigenfunction also exists and satisfies the Schrödinger equation

$$
(H_0 + V)\psi_k = E_k \psi_k
$$

in the sense of distributions. In other words, the recursion mentioned at the end of the previous section does not stop until energy zero is reached. Furthermore each $E_k$ can have only finite multiplicity, i.e., each number $E_k < 0$ occurs only finitely many times in the list of eigenvalues. Conversely, every normalized solution to $(H_0 + V)\psi = E\psi$ with $E \leq 0$ and with $\psi \in H^1(\mathbb{R}^n)$ (respectively, $\psi \in H^{1/2}(\mathbb{R}^n)$) is a linear combination of eigenfunctions with eigenvalue $E$. 
REMARK. There is no general theorem about the existence of a minimizer if \( E_k = 0 \).

PROOF. The proof of existence of a minimizer \( \psi_k \) is basically the same as the one of Theorem 11.5. Take a minimizing sequence \( \psi^j_k, j = 1, 2, \ldots \), each of which is orthogonal to the functions \( \psi_0, \ldots, \psi_{k-1} \). By passing to a subsequence we can find a weak limit in \( H^1(\mathbb{R}^n) \) (resp. \( H^{1/2}(\mathbb{R}^n) \) in the relativistic case) which we call \( \psi_k \). As in Theorem 11.4, \( \mathcal{E}(\psi_k) = E_k \) and \( \|\psi_k\|_2 = 1 \). The only thing we have to check is that \( \psi_k \) is orthogonal to \( \psi_0, \ldots, \psi_{k-1} \). This, however, is a direct consequence of the definition of the weak limit.

The proof of (1) requires a few steps. First, as in the proof of Theorem 11.5, we conclude that the distribution \( D := (H_0 + V - E_k)\psi_k \) is a distribution that satisfies \( D(f) = 0 \) for every \( f \in C_c^\infty(\mathbb{R}^n) \) with the property that \( (f, \psi_i) = 0 \) for all \( i = 0, \ldots, k - 1 \). By Theorem 6.14 (linear dependence of distributions), this implies that

\[
D = \sum_{i=0}^{k-1} c_i \psi_i \tag{2}
\]

for some numbers \( c_0, \ldots, c_{k-1} \). Our goal is to show that \( c_i = 0 \) for all \( i \).

Formally, this is proved by multiplying (2) by some \( \psi_j \) with \( j \leq k - 1 \) and partially integrating to obtain (using the assumed orthogonality)

\[
\int_{\mathbb{R}^n} \nabla \psi_j \cdot \nabla \psi_k + \int_{\mathbb{R}^n} V \overline{\psi}_j \psi_k = c_j \tag{3}
\]

On the other hand, taking the complex conjugate of (1) for \( \psi_j \) and multiplying it by \( \psi_k \) yields

\[
\int_{\mathbb{R}^n} \nabla \psi_j \cdot \nabla \psi_k + \int_{\mathbb{R}^n} V \overline{\psi}_j \psi_k = 0 \tag{4}
\]

The justification of this formal manipulation is left as Exercise 3.

To prove that \( E_k \) has finite multiplicity, assume the contrary. This means that \( E_k = E_{k+1} = E_{k+2} = \cdots \). By the foregoing there is then an orthonormal sequence \( \psi_1, \psi_2, \ldots \) satisfying (1). By 11.3(10) the kinetic energies \( T_{\psi_j} \) remain bounded, i.e., \( T_{\psi_j} < C \) for some \( C > 0 \). Since the \( \psi_j \)'s are orthogonal, they converge weakly to zero in \( L^2(\mathbb{R}^n) \), and hence in \( H^1(\mathbb{R}^n) \) as well, as \( j \to \infty \). But in Theorem 11.4 it was shown that \( V_{\psi_j} \to 0 \) as \( j \to \infty \) and hence \( E_k = \lim_{j \to \infty} T_{\psi_j} + V_{\psi_j} \geq 0 \), which is a contradiction.

The proof that any solution to the Schrödinger equation is a linear combination of eigenfunctions with eigenvalue \( E \) follows the integration by parts argument used for the proof of (1). See Exercise 3.
11.7 THEOREM (Regularity of solutions)

Let \( B_1 \subset \mathbb{R}^n \) be an open ball and let \( u \) and \( V \) be functions in \( L^1(B_1) \) that satisfy

\[-\Delta u + Vu = 0 \quad \text{in } D'(B_1). \tag{1}\]

Then the following hold for any ball \( B \) concentric with \( B_1 \) and with strictly smaller radius:

1. \( n = 1 \): Without any further assumption on \( V \), \( u \) is continuously differentiable.
2. \( n = 2 \): Without any further assumptions on \( V \), \( u \in L^q(B) \) for all \( q < \infty \).
3. \( n \geq 3 \): Without any further assumptions on \( V \), \( u \in L^q(B) \) with \( q < n/(n-2) \).
4. \( n \geq 2 \): If \( V \in L^p(B_1) \) for \( n \geq p > n/2 \), then for all \( \alpha < 2 - n/p \),
   \[ |u(x) - u(y)| \leq C|x - y|^{\alpha} \]
   for some constant \( C \) and all \( x, y \in B \).
5. \( n \geq 1 \): If \( V \in L^p(B_1) \) for \( p > n \), then \( u \) is continuously differentiable and its first derivatives \( \partial_i u \) satisfy
   \[ |\partial_i u(x) - \partial_i u(y)| \leq C|x - y|^{\alpha} \]
   for all \( \alpha < 1 - n/p \), all \( x, y \in B \) and some constant \( C \).
6. Let \( V \in C^{k,\alpha}(B_1) \) for some \( k \geq 0 \) and \( 0 < \alpha < 1 \) (see Remark (2) in Sect. 10.1). Then \( u \in C^{k+2,\alpha}(B) \).

PROOF. The assumption (1) implies that \( Vu \in L^1_{\text{loc}}(B_1) \). As explained in Sect. 10.1 regularity questions are purely local. Thus, applying Theorem 10.2(i), statements (i), (ii) and (iii) are readily obtained. To prove (iv) we use the 'bootstrap' argument. If \( n = 2 \) we know by (ii) that \( u \in L^q(B_2) \) for any \( q < \infty \), and hence \( Vu \in L^r(B_2) \) for some \( r > n/2 \). Here \( B \subset B_2 \subset B_1 \) and \( B_2 \) is concentric with \( B_1 \). Then Theorem 10.2(ii) implies that \( u \) is Hölder continuous, which shows that in fact \( Vu \in L^p(B_3) \). Again \( B \subset B_3 \subset B_2 \) and \( B_3 \) is concentric with \( B_2 \). One more application of Theorem 10.2(ii) yields the result for \( n = 2 \), since the radii of the balls decrease by an arbitrarily small amount.

If \( n \geq 3 \), we proceed as follows. Suppose that \( Vu \in L^{s_1}(B_2) \) for some \( 1 < s_1 < n/2 \) and some ball \( B_2 \) concentric with \( B_1 \) but of smaller radius. By Theorem 10.2(i), \( u \in L^t(B_3) \) for any \( t < ns_1/(n-2s_1) \) and \( B_3 \) concentric with \( B_2 \) with a smaller radius than that of \( B_2 \), but as close as we please. Since \( V \in L^p(B_1) \) for \( n/2 < p \leq n \), we can set \( 1/p = 2/n - \varepsilon \) with \( 0 < \varepsilon \leq 1/n \). By Hölder's inequality \( Vu \in L^{s_2}(B_3) \) for any \( s_2 < s_1/(1-\varepsilon s_1) \) and thus,
in particular, for any $s_2 < s_1/(1 - \varepsilon)$. Iterating this estimate we arrive at the situation where, for some finite $k$, $V u \in L^{s_k}(B_{k+1})$, $s_k > n/2$. Then, by Theorem 10.2(ii), $u$ is Hölder continuous. Now $V u \in L^p(B)$ for some ball concentric with $B_1$ but of smaller radius, and Theorem 10.2(ii) applied once more yields the result.

In the same fashion, by using Theorem 10.3 in addition, the reader can easily prove (v) and (vi).

11.8 THEOREM (Uniqueness of minimizers)

Assume that $\psi_0 \in H^1(\mathbb{R}^n)$ is a minimizer for $\mathcal{E}$, i.e., $\mathcal{E}(\psi_0) = E_0 > -\infty$ and $\|\psi_0\|_2 = 1$. The only assumptions we make are that $V \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $V$ is locally bounded from above (not necessarily from below) and, of course, $V|\psi_0|^2$ is summable. Then $\psi_0$ satisfies the Schrödinger equation 11.2(1) with $E = E_0$. Moreover $\psi_0$ can be chosen to be a strictly positive function and, most importantly, $\psi_0$ is the unique minimizer up to a constant phase.

In the relativistic case the same is true for an $H^{1/2}(\mathbb{R}^n)$ minimizer, but this time we need only assume that $V$ is in $L^1_{\text{loc}}(\mathbb{R}^n)$.

PROOF. Since

$$E_0 = \mathcal{E}(\psi_0) = \int_{\mathbb{R}^n} |\nabla \psi_0|^2 + \int_{\mathbb{R}^n} V(x)|\psi_0(x)|^2$$

and $\psi_0 \in H^1(\mathbb{R}^n)$, we must have that both

$$\int_{\mathbb{R}^n} [V(x)]_+|\psi_0(x)|^2 \, dx \quad \text{and} \quad \int_{\mathbb{R}^n} [V(x)]_-|\psi_0(x)|^2 \, dx$$

are finite. Thus, in particular, $\int_{\mathbb{R}^n} V(x)\psi_0(x)\phi(x) \, dx$ is finite for every $\phi \in C_c^\infty(\mathbb{R}^n)$. Next, we compute for any $\phi \in C_c^\infty(\mathbb{R}^n)$

$$0 \leq \mathcal{E}(\psi_0 + \varepsilon \phi) - E_0\|\psi_0 + \varepsilon \phi\|_2^2$$

$$= \mathcal{E}(\psi_0) - E_0 + 2\varepsilon \text{Re} \int [\nabla \psi_0 \nabla \phi + (V - E_0)\psi_0 \bar{\phi}]$$

$$+ \varepsilon^2 \int [||\nabla \phi||^2 + (V - E_0)|\phi|^2].$$

Every term is finite and, since $\mathcal{E}(\psi_0) = E_0$, the last two terms add up to something nonnegative. Since $\varepsilon$ is arbitrary and can have any sign, this implies that

$$-\Delta \psi_0 + W \psi_0 = 0 \quad \text{in} \ D'(\mathbb{R}^n),$$

(1)

where $W := V - E_0$. 


Sections 11.7–11.9

Next we note that with \( \psi_0 = f + ig \), \( f \) and \( g \) separately are minimizers. Since, by Theorem 6.17 (derivative of the absolute value), \( \mathcal{E}(f) = \mathcal{E}(|f|) \) and \( \mathcal{E}(g) = \mathcal{E}(|g|) \), we also have that \( \phi_0 = |f| + i|g| \) is a minimizer. By Theorem 7.8 (convexity inequality for gradients) \( \mathcal{E}(|\phi_0|) \leq \mathcal{E}(\phi_0) \), and hence there must be equality. The same Theorem 7.8 states that there is equality if and only if \( |f| = c|g| \) for some constant \( c \) provided that either \( |f(x)| \) or \( |g(x)| \) is strictly positive for all \( x \in \mathbb{R}^n \).

Since these functions are minimizers, they satisfy the Schrödinger equation (1) and, since \( V \) is locally bounded, so is \( W \). By Theorem 9.10 (lower bounds on Schrödinger 'wave' functions) \( |f(x)| \) and \( |g(x)| \) are equivalent to strictly positive lower semicontinuous functions \( \tilde{f} \) and \( \tilde{g} \). Thus, up to a fixed sign, \( f = \tilde{f} \) and \( g = \tilde{g} \), and thus \( f = cg \) for some constant \( c \), i.e., \( \psi_0 = (1 + ic)f \).

The proof for the relativistic case is similar except that the convexity inequality, Theorem 7.13, for the relativistic kinetic energy does not require strict positivity of the function involved. \[ \square \]

11.9 COROLLARY (Uniqueness of positive solutions)

Suppose that \( V \) is in \( L^1_{\text{loc}}(\mathbb{R}^n) \), \( V \) is bounded above (uniformly and not just locally) and that \( E_0 > -\infty \). Let \( \psi \neq 0 \) be any nonnegative function with \( \|\psi\|_2 = 1 \) that is in \( H^1(\mathbb{R}^n) \) and satisfies the nonrelativistic Schrödinger equation 11.2(1) in \( \mathcal{D}'(\mathbb{R}^n) \) or is in \( H^{1/2}(\mathbb{R}^n) \) and satisfies the relativistic Schrödinger equation

\[
\left[ \sqrt{\frac{p^2}{m^2}} + m \right] \psi + V\psi = E\psi \quad \text{in} \ \mathcal{D}'(\mathbb{R}^n). \tag{1}
\]

Then \( E = E_0 \) and \( \psi \) is the unique minimizer \( \psi_0 \).

PROOF. The main step is to prove that \( E = E_0 \). The rest will then follow simply from Remark (2) in Sect. 11.5 (existence of a minimizer) and from Theorem 11.8 (uniqueness of minimizers). To prove \( E = E_0 \), we prove that \( E \neq E_0 \) implies the orthogonality relation \( \int \psi\psi_0 = 0 \). (We know that \( E \geq E_0 \) by Remark (2) in 11.5.) Since \( \psi_0 \) is strictly positive and \( \psi \) is nonnegative, this orthogonality is impossible.

To prove orthogonality when \( E \neq E_0 \) in the nonrelativistic case we take the Schrödinger equation for \( \psi_0 \), multiply it by \( \psi \), integrate over \( \mathbb{R}^n \) and obtain (formally)

\[
\int_{\mathbb{R}^n} \nabla \psi \cdot \nabla \psi_0 + \int_{\mathbb{R}^n} (V - E_0)\psi\psi_0 = 0. \tag{2}
\]
To justify this we note, from 11.2(1), that the distribution $\Delta \psi$ is a function and hence is in $L^1_{\text{loc}}(\mathbb{R}^n)$. Moreover, since $\psi$ is nonnegative and $V$ is bounded above, $\Delta \psi = f + g$ for some nonnegative function $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and some $g \in L^2(\mathbb{R}^n)$. Thus (2) follows from Theorem 7.7.

If we interchange $\psi$ and $\psi_0$, we obtain (2) with $E_0$ replaced by $E$. If $E \neq E_0$, this is a contradiction unless $\int \psi \psi_0 = 0$.

The proof in the relativistic case is identical, except for the substitution of 7.15(3) in place of 7.7(2).

11.10 EXAMPLE (The hydrogen atom)

The potential $V$ for the hydrogen atom located at the origin in $\mathbb{R}^3$ is

$$V(x) = -|x|^{-1}. \quad (1)$$

A solution to the Schrödinger equation 11.2(1) is found by inspection to be

$$\psi_0(x) = \exp\left(-\frac{1}{2}|x|\right), \quad E_0 = -\frac{1}{4}. \quad (2)$$

Since $\psi_0$ is positive, it is the ground state, i.e., the unique minimizer of

$$E(\psi) = \int_{\mathbb{R}^3} |\nabla \psi|^2 - \int_{\mathbb{R}^3} \frac{1}{|x|} |\psi(x)|^2 \, dx.$$  

This fact follows from Corollary 11.9 (uniqueness of positive solutions). It is not obvious and is usually not mentioned in the standard texts on quantum mechanics.

We can note several facts about $\psi_0$ that are in accord with our previous theorems.

(i) Since $V$ is infinitely differentiable in the complement of the origin, $x = 0$, the solution $\psi_0$ is also infinitely differentiable in that same region. This result can be seen directly from Theorem 11.7 (regularity of solutions). As a matter of fact, $V$ is real analytic in this region (meaning that it can be expanded in a power series with some nonzero radius of convergence about every point of the region). It is a general fact, borne out by our example, that in this case $\psi_0$ is also real analytic in this region; this result is due to Morrey and can be found in [Morrey].

(ii) Since $V$ is in $L^p_{\text{loc}}(\mathbb{R}^n)$ for $3 > p > 3/2$, we also conclude from Theorem 11.7 that $\psi_0$ must be Hölder continuous at the origin, namely

$$|\psi_0(x) - \psi_0(0)| < c|x|^\alpha$$
for all exponents $1 > \alpha > 0$. In our example, $\psi_0$ is slightly better; it is Lipschitz continuous, i.e., we can take $\alpha = 1$.

- We turn now to our second main example of a variational problem—the Thomas–Fermi (TF) problem. See [Lieb-Simon] and [Lieb, 1981]. It goes back to the idea of L. H. Thomas and E. Fermi in 1926 that a large atom, with many electrons, can be approximately modeled by a simple nonlinear problem for a ‘charge density’ $\rho(x)$. We shall not attempt to derive this approximation from the Schrödinger equation but will content ourselves with stating the mathematical problem.

The potential function $Z/|x|$ that appears in the following can easily be replaced by

$$V(x) := \sum_{j=1}^{K} Z_j |x - R_j|^{-1}$$

with $Z_j > 0$ and $R_j \in \mathbb{R}^3$, but we refrain from doing so in the interest of simplicity.

Unlike our previous tour through the Schrödinger equation, this time we shall leave many steps as an exercise for the reader (who should realize that knowledge does not come without a certain amount of perspiration).

### 11.11 THE THOMAS–FERMI PROBLEM

TF theory is defined by an energy functional $\mathcal{E}$ on a certain class of nonnegative functions $\rho$ on $\mathbb{R}^3$:

$$\mathcal{E}(\rho) := \frac{3}{5} \int_{\mathbb{R}^3} \rho(x)^{5/3} \, dx - \int_{\mathbb{R}^3} \frac{Z}{|x|} \rho(x) \, dx + D(\rho, \rho), \quad (1)$$

where $Z > 0$ is a fixed parameter (the charge of the atom’s nucleus) and

$$D(\rho, \rho) := \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \rho(x) \rho(y) |x - y|^{-1} \, dx \, dy \quad (2)$$

is the Coulomb energy of a charge density, as given by 9.1(2). The class of admissible functions is

$$\mathcal{C} := \left\{ \rho : \rho \geq 0, \int_{\mathbb{R}^3} \rho < \infty, \quad \rho \in L^{5/3}(\mathbb{R}^3) \right\}. \quad (3)$$

We leave it as an exercise to show that each term in (1) is well defined and finite when $\rho$ is in the class $\mathcal{C}$.

Our problem is to minimize $\mathcal{E}(\rho)$ under the condition that $\int \rho = N$, where $N$ is any fixed positive number (identified as the ‘number’ of electrons in the atom). The case $N = Z$ is special and is called the neutral case. We