CAMBRIDGE

Cambridge University Press
978-0-521-19118-0 - The Stability of Matter in Quantum Mechanics
Elliott H. Lieb and Robert Seiringer
Excerpt
More information

CHAPTER 1

Prologue

1.1 Introduction

The basic constituents of ordinary matter are electrons and atomic nuclei. These interact with each other with several kinds of forces – electric, magnetic and gravitational – the most important of which is the electric force. This force is attractive between oppositely charged particles and repulsive between like-charged particles. (The electrons have a negative electric charge $-e$ while the nuclei have a positive charge $+Ze$, with $Z = 1, 2, \ldots, 92$ in nature.) Thus, the strength of the attractive electrostatic interaction between electrons and nuclei is proportional to Ze^2, which equals $Z\alpha$ in appropriate units, where α is the dimensionless **fine-structure constant**, defined by

$$\alpha = \frac{e^2}{\hbar c} = 7.297\,352\,538 \times 10^{-3} = \frac{1}{137.035\,999\,68}, \tag{1.1.1}$$

and where c is the speed of light, $\hbar = h/2\pi$ and h is Planck's constant.

The basic question that has to be resolved in order to understand the existence of atoms and the stability of our world is:

Why don't the point-like electrons fall into the (nearly) point-like nuclei?

This problem of classical mechanics was nicely summarized by Jeans in 1915 [97]:

> "There would be a very real difficulty in supposing that the (force) law $1/r^2$ held down to zero values of r. For the force between two charges at zero distance would be infinite; we should have charges of opposite sign continually rushing together and, when once together, no force would be adequate to separate them . . . Thus the matter in the universe would tend to shrink into nothing or to diminish indefinitely in size."

A sensitive reader might object to Jeans' conclusion on the grounds that the non-zero radius of nuclei would ameliorate the collapse. Such reasoning is beside the point, however, because the equilibrium separation of charges observed in nature is not the nuclear diameter (10^{-13} cm) but rather the atomic size (10^{-8} cm) predicted by Schrödinger's equation. Therefore, as concerns the problem of understanding stability, in which equilibrium lengths are of the order of 10^{-8} cm, there is no loss in supposing that all our particles are point particles.

To put it differently, why is the energy of an atom with a point-like nucleus not $-\infty$? The fact that it is not is known as **stability of the first kind**; a more precise definition will be given later. The question was successfully answered by quantum mechanics, whose exciting development in the beginning of the twentieth century we will not try to relate – except to note that the basic theory culminated in Schrödinger's famous equation of 1926 [156]. This equation explained the new, non-classical, fact that as an electron moves close to a nucleus its kinetic energy necessarily increases in such a way that the minimum total energy (kinetic plus potential) occurs at some positive separation rather than at zero separation.

This was one of the most important triumphs of quantum mechanics!

Thomson discovered the electron in 1897 [180, 148], and Rutherford [155] discovered the (essentially) point-like nature of the nucleus in 1911, so it took 15 years from the discovery of the problem to its full solution. But it took almost three times as long, 41 years from 1926 to 1967, before the second part of the stability story was solved by Dyson and Lenard [44].

The second part of the story, known as **stability of the second kind**, is, even now, rarely told in basic quantum mechanics textbooks and university courses, but it is just as important. Given the stability of atoms, is it obvious that bulk matter with a large number N of atoms (say, $N = 10^{23}$) is also stable in the sense that the energy and the volume occupied by $2N$ atoms are twice that of N atoms? Our everyday physical experience tells us that this additivity property, or linear law, holds but is it also necessarily a consequence of quantum mechanics? Without this property, the world of ordinary matter, as we know it, would not exist.

Although physicists largely take this property for granted, there were a few that thought otherwise. Onsager [145] was perhaps the first to consider this kind of question, and did so effectively for classical particles with Coulomb interactions but with the addition of hard cores that prevent particles from getting too close together. The full question (without hard cores) was addressed by Fisher and Ruelle in 1966 [66] and they generalized Onsager's results to smeared out charges. In 1967 Dyson and Lenard [44] finally succeeded in showing that stability of the second kind for truly point-like quantum particles with Coulomb forces holds but, surprisingly, that it need not do so. That is, the *Pauli exclusion principle*, which will be discussed in Chapter 3, and which has no classical counterpart, was essential. Although matter would not collapse without it, the linear law would *not* be satisfied, as Dyson showed in 1967 [43]. Consequently, stability of the second kind does *not* follow from stability of the first kind! If the electrons and nuclei were all bosons (which are particles that do not satisfy the exclusion principle), the energy would not satisfy a linear law but rather decrease like $-N^{7/5}$; we will return to this astonishing discovery later.

The Dyson–Lenard proof of stability of the second kind [44] was one of the most difficult, up to that time, in the mathematical physics literature. A challenge was to find an essential simplification, and this was done by Lieb and Thirring in 1975 [134]. They introduced new mathematical inequalities, now called Lieb–Thirring (LT) inequalities (discussed in Chapter 4), which showed that a suitably modified version of the 1927 approximate theory of Thomas and Fermi [179, 62] yielded, in fact, a lower bound to the exact quantum-mechanical answer. Since it had already been shown, by Lieb and Simon in 1973 [129, 130], that this Thomas–Fermi theory possessed a linear lower bound to the energy, the many-body stability of the second kind immediately followed.

The Dyson–Lenard stability result was one important ingredient in the solution to another, but related problem that had been raised many years earlier. Is it true that the 'thermodynamic limit' of the free energy per particle exists for an infinite system at fixed temperature and density? In other words, given that the energy per particle of some system is bounded above and below, independent of the size of the system, how do we know that it does not oscillate as the system's size increases? The existence of a limit was resolved affirmatively by Lebowitz and Lieb in 1969 [103, 116], and we shall give that proof in Chapter 14.

There were further surprises in store, however! The Dyson–Lenard result was not the end of the story, for it was later realized that there were other sources of instability that physicists had not seriously thought about. Two, in fact. The eventual solution of these two problems leads to the conclusion that, ultimately, stability requires more than the Pauli principle. It also requires an upper bound on both the physical constants α and $Z\alpha$.[1]

One of the two new questions considered was this. What effect does Einstein's relativistic kinematics have? In this theory the Newtonian kinetic energy of an electron with mass m and momentum $\boldsymbol{p}$, $\boldsymbol{p}^2/2m$, is replaced by the much weaker $\sqrt{\boldsymbol{p}^2c^2 + m^2c^4} - mc^2$. So much weaker, in fact, that the simple atom is stable only if the relevant coupling parameter $Z\alpha$ is not too large! This fact was known in one form or another for many years – from the introduction of Dirac's 1928 relativistic quantum mechanics [39], in fact. It was far from obvious, therefore, that many-body stability would continue to hold even if $Z\alpha$ is kept small (but fixed, independent of N). Not only was the linear N-dependence in doubt but also stability of the first kind was unclear. This was resolved by Conlon in 1984 [32], who showed that stability of the second kind holds if $\alpha < 10^{-200}$ and $Z = 1$.

Clearly, Conlon's result needed improvement and this led to the invention of interesting new inequalities to simplify and improve his result. We now know that stability of the second kind holds if and only if *both* α and $Z\alpha$ are not too large. The bound on α itself was the new reality, previously unknown in the physics literature.

Again new inequalities were needed when it was realized that magnetic fields could also cause instabilities, even for just one atom, if $Z\alpha^2$ is too large. The understanding of this strange, and totally unforeseen, fact requires the knowledge that the appropriate Schrödinger equation has '*zero-modes*', as discovered by Loss and Yau in 1986 [139] (that is, square integrable, time-independent solutions with zero kinetic energy). But stability of the second kind was still open until Fefferman showed in 1995 [57, 58] that stability of the second kind holds if $Z = 1$ and α is very small. This result was subsequently improved to robust values of $Z\alpha^2$ and α by Lieb, Loss and Solovej in 1995 [123].

The surprises, in summary, were that stability of the second kind requires bounds on the fine-structure constant and the nuclear charges. In the relativistic case, smallness of α and of $Z\alpha$ is necessary, whereas in the non-relativistic case with magnetic fields, smallness of α and of $Z\alpha^2$ is required.

[1] If $Z \geq 1$, which it always is in nature, a bound on $Z\alpha$ implies a bound on α, of course. The point here is that the necessary bound on α is independent of Z, even if Z is arbitrarily small. In this book we shall not restrict our attention to integer Z.

Given these facts, one can ask if the *simultaneous* introduction of relativistic mechanics, magnetic fields, and the quantization of those fields in the manner proposed by M. Planck in 1900 [149], leads to new surprises about the requirements for stability. The answer, proved by Lieb, Loss, Siedentop and Solovej [127, 119], is that in at least one version of the problem no new conditions are needed, except for expected adjustments of the allowed bounds for $Z\alpha$ and α.

While we will visit all these topics in this book, we will not necessarily follow the historical route. In particular, we will solve the non-relativistic problem by using the improved inequalities invented to handle the relativistic problem, without the introduction of Thomas–Fermi theory. The Thomas–Fermi story is interesting, but no longer essential for our understanding of the stability of matter. Hence we will mention it, and sketch its application in the stability of matter problem, but we will not treat it thoroughly, and will not make further use of it. Some earlier pedagogical reviews are in [108, 115].

1.2 Brief Outline of the Book

An elementary introduction to quantum mechanics is given in **Chapter 2**. It is a thumbnail sketch of the relevant parts of the subject for readers who might want to refresh their memory, and it also serves to fix notation. Readers familiar with the subject can safely skip the chapter.

Chapter 3 discusses the many-body aspects of quantum mechanics and, in particular, introduces the concept of stability of matter in Section 3.2. The chapter also contains several results that will be used repeatedly in the chapters to follow, like the monotonicity of the ground state energy in the nuclear charges, and the fact the bosons have the lowest possible ground state energy among all symmetry classes.

A detailed discussion of Lieb–Thirring inequalities is the subject of **Chapter 4**. These inequalities play a crucial role in our understanding of stability of matter. They concern bounds on the moments of the negative eigenvalues of Schrödinger type operators, which lead to lower bounds on the kinetic energy of many-particle systems in terms of the corresponding semiclassical expressions. This chapter, like Chapters 5 and 6, is purely mathematical and contains analytic inequalities that will be applied in the following chapters.

Electrostatics is an old subject whose mathematical underpinning goes back to Newton's discussion in the *Principia* [144] of the gravitational force, which behaves in a similar way except for a change of sign from repulsive to attractive. Nevertheless, new inequalities are essential for understanding many-body systems, and these are given in **Chapters 5 and 6**. The latter chapter contains a proof of the Lieb–Oxford inequality [125], which gives a bound on the indirect part of the Coulomb electrostatic energy of a quantum system.

Chapter 7 contains a proof of stability of matter of non-relativistic fermionic particles. This is the same model for which stability was first shown by Dyson and Lenard [44] in 1967. The three proofs given here are different and very short given the inequalities derived in Chapters 4–6. As a consequence, matter is not only stable but also extensive, in the sense that the volume occupied is proportional to the number of particles. The instability of the same model for bosons will also be discussed.

The analogous model with relativistic kinematics is discussed in **Chapter 8**, and stability for fermions is proved for a certain range of the parameters α and $Z\alpha$. Unlike in the non-relativistic case, where the range of values of these parameters was unconstrained, bounds on these parameters are essential, as will be shown. The proof of stability in the relativistic case will be an important ingredient concerning stability of the models discussed in Chapters 9, 10 and 11.

The influence of spin and magnetic fields will be studied in **Chapter 9**. If the kinetic energy of the particles is described by the Pauli operator, it becomes necessary to include the magnetic field energy for stability. Again, bounds on various parameters become necessary, this time α and $Z\alpha^2$. It turns out that zero modes of the Pauli operator are a key ingredient in understanding the boundary between stability and instability.

If the kinetic energy of relativistic particles is described by the Dirac operator, the question of stability becomes even more subtle. This is the content of **Chapter 10**. For the Brown–Ravenhall model, where the physically allowed states are the positive energy states of the free Dirac operator, there is always instability in the presence of magnetic fields. Stability can be restored by appropriately modifying the model and choosing as the physically allowed states the ones that have a positive energy for the Dirac operator *with* the magnetic field.

The effects of the quantum nature of the electromagnetic field will be investigated in **Chapter 11**. The models considered are the same as in Chapters 9 and 10, but now the electromagnetic field will be quantized. These models are caricatures of quantum electrodynamics. The chapter includes a self-contained mini-course on the electromagnetic field and its quantization. The stability and instability results are essentially the same as for the non-quantized field, except for different bounds on the parameter regime for stability.

How many electrons can an atom or molecule bind? This question will be addressed in **Chapter 12**. The reason for including it in a book on stability of matter is to show that for a lower bound on the ground state energy only the minimum of the number of nuclei and the number of electrons is relevant. A large excess charge can not lower the energy.

Once a system becomes large enough so that the gravitational interaction can not be ignored, stability fails. This can be seen in nature in terms of the gravitational collapse of stars and the resulting supernovae, or as the upper mass limit of cold stars. Simple models of this gravitational collapse, as appropriate for white dwarfs and neutron stars, will be studied in **Chapter 13**. In particular, it will be shown how the critical number of particles for collapse depends on the gravitational constant G, namely $G^{-3/2}$ for fermions and G^{-1} for bosons, respectively.

The first 13 chapters deal essentially with the problem of showing that the lowest energy of matter is bounded below by a constant times the number of particles. The final **Chapter 14** deals with the question of showing that the energy is really proportional to the number of particles, i.e., that the energy per particle has a limit as the particle number goes to infinity. Such a limit exists not only for the ground state energy, but also for excited states in the sense that at positive temperature the thermodynamic limit of the free energy per particle exists.

CHAPTER 2

Introduction to Elementary Quantum Mechanics and Stability of the First Kind

In this second chapter we will review the basic mathematical and physical facts about quantum mechanics and establish physical units and notation. Those readers already familiar with the subject can safely jump to the next chapter.

An attempt has been made to make the presentation in this chapter as elementary as possible, and yet present the basic facts that will be needed later. There are many beautiful and important topics which will not be touched upon such as self-adjointness of Schrödinger operators, the general mathematical structure of quantum mechanics and the like. These topics are well described in other works, e.g., [150].

Much of the following can be done in a Euclidean space of arbitrary dimension, but in this chapter the dimension of the Euclidean space is taken to be three – which is the physical case – unless otherwise stated. We do this to avoid confusion and, occasionally, complications that arise in the computation of mathematical constants. The interested reader can easily generalize what is done here to the $\mathbb{R}^d$, $d > 3$ case. Likewise, in the next chapters we mostly consider N particles, with spatial coordinates in $\mathbb{R}^3$, so that the total spatial dimension is $3N$.

2.1 A Brief Review of the Connection Between Classical and Quantum Mechanics

Considering the range of validity of quantum mechanics, it is not surprising that its formulation is more complicated and abstract than classical mechanics. Nevertheless, classical mechanics is a basic ingredient for quantum mechanics. One still talks about position, momentum and energy which are notions from Newtonian mechanics.

The connection between these two theories becomes apparent in the semiclassical limit, akin to passing from wave optics to geometrical optics. In its Hamiltonian formulation, classical mechanics can be viewed as a problem of geometrical optics. This led Schrödinger to guess the corresponding wave equation. We refrain from fully explaining the semiclassical limit of quantum mechanics. For one aspect of this problem, however, the reader is referred to Chapter 4, Section 4.1.1.

We turn now to classical dynamics itself, in which a point particle is fully described by giving its **position** $\boldsymbol{x} = (x^1, x^2, x^3)$ in $\mathbb{R}^3$ and its **velocity** $\boldsymbol{v} = \mathrm{d}\boldsymbol{x}/\mathrm{d}t = \dot{\boldsymbol{x}}$ in $\mathbb{R}^3$ at any time t, where the dot denotes the derivative with respect to time.[1] Newton's law of motion says that along any mechanical trajectory its acceleration $\dot{\boldsymbol{v}} = \ddot{\boldsymbol{x}}$ satisfies

$$m\ddot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t), \tag{2.1.1}$$

where $\boldsymbol{F}$ is the **force** acting on the particle and m is the **mass**. With $\boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)$ given, the expression (2.1.1) is a system of second order differential equations which together with the initial conditions $\boldsymbol{x}(t_0)$ and $\boldsymbol{v}(t_0) = \dot{\boldsymbol{x}}(t_0)$ determine $\boldsymbol{x}(t)$ and thus $\boldsymbol{v}(t)$ for all times. If there are N particles interacting with each other, then (2.1.1) takes the form

$$m_i\ddot{\boldsymbol{x}}_i = \boldsymbol{F}_i, \quad i = 1, \ldots, N, \tag{2.1.2}$$

where $\boldsymbol{F}_i$ denotes the sum of all forces acting on the i^{th} particle and $\boldsymbol{x}_i$ denotes the position of the i^{th} particle. As an example, consider the force between two charged particles, whose respective charges are denoted by Q_1 and Q_2, namely the **Coulomb force** given (in appropriate units, see Section 2.1.7) by

$$\boldsymbol{F}_1 = Q_1 Q_2 \frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3} = -\boldsymbol{F}_2. \tag{2.1.3}$$

If $Q_1 Q_2$ is positive the force is repulsive and if $Q_1 Q_2$ is negative the force is attractive. Formula (2.1.3) can be written in terms of the **potential energy function**

$$V(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{Q_1 Q_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}, \tag{2.1.4}$$

noting that

$$\boldsymbol{F}_1 = -\nabla_{\boldsymbol{x}_1} V \quad \text{and} \quad \boldsymbol{F}_2 = -\nabla_{\boldsymbol{x}_2} V. \tag{2.1.5}$$

As usual, we denote the gradient by $\nabla = (\partial/\partial x^1, \partial/\partial x^2, \partial/\partial x^3)$.

[1] We follow the physicists' convention in which vectors are denoted by boldface letters.

2.1.1 Hamiltonian Formulation

Hamilton's formulation of classical mechanics is the entry to quantum physics. **Hamilton's equations** are

$$\dot{\boldsymbol{x}} = \frac{\partial H}{\partial \boldsymbol{p}}, \quad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{x}} \tag{2.1.6}$$

where $H(\boldsymbol{x}, \boldsymbol{p})$ is the **Hamilton function** and $\boldsymbol{p}$ the **canonical momentum** of the particle. Assuming that

$$F(\boldsymbol{x}) = -\nabla V(\boldsymbol{x}) \tag{2.1.7}$$

for some potential V then, in the case that the canonical momentum is given by

$$\boldsymbol{p} = m\boldsymbol{v}, \tag{2.1.8}$$

Eq. (2.1.6) with

$$H = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{x}) \tag{2.1.9}$$

yields (2.1.1). The function

$$T(\boldsymbol{p}) = \frac{\boldsymbol{p}^2}{2m} \tag{2.1.10}$$

is called the **kinetic energy** function. A simple computation using Eq. (2.1.6) shows that along each mechanical trajectory the function $H(\boldsymbol{x}(t), \boldsymbol{p}(t))$ is a constant which we call the **energy**, E.

2.1.2 Magnetic Fields

Not in all cases is the canonical momentum given by (2.1.8). An example is the motion of a charged particle of mass m and charge $-e$ in a magnetic field $\boldsymbol{B}(\boldsymbol{x})$ in addition to a potential, $V(\boldsymbol{x})$. The **Lorentz force** on such a particle located at $\boldsymbol{x}$ and having velocity $\boldsymbol{v}$ is[2]

$$F_{\text{Lorentz}} = -\frac{e}{c}\boldsymbol{v} \wedge \boldsymbol{B}(\boldsymbol{x}). \tag{2.1.11}$$

[2] We use the symbol $\wedge$ for the vector product on $\mathbb{R}^3$, instead of $\times$, since the latter may be confused with $\boldsymbol{x}$.