The Stability of Matter Interacting with Fields

Michael Loss

The Schrödinger equation occupies a central and, at the same time, peculiar place in physics. It is central in the sense that it is acknowledged to be universal; it explains all facts about nonrelativistic matter such as superconductivity. Peculiar it is, since we rarely derive any material properties ab-initio from the Schrödinger equation. Overwhelmed by its mathematical complexity which, of course, reflects the unlimited forms of matter, theoretical physicists dig out the salient bits of mathematics associated with the phenomena and try to deal with those.

There are, however, a number of facts about bulk matter that can be derived rigorously from the basic principles of quantum mechanics. The one phenomenon at the heart of it all is that the physical size of bulk matter and its energy content are proportional to the number of particles involved or, in short, that matter is stable. Matter consists of electrons and nuclei that interact with each other through electrostatic forces, i.e., Coulomb's $1/r^2$ law repelling particles with like charges and attracting particles of opposite charges. Clearly, stability cannot be explained on the basis of classical physics alone, since there is nothing that resists the infinite attraction of Coulomb's law whose lowest energy is reached when particles of opposite charge sit on top of each other. In fact, the development of quantum mechanics was to a certain extent driven by the desire to estimate the size of atoms. While the bulkiness of atoms was understood early on, it was not until the mid-sixties through the work of Fisher, Ruelle, and others in statistical mechanics that a mathematical formulation of the stability of matter problem was even given. A many-particle system is stable if its ground-state energy is proportional to the number of particles involved. This was proved in the late sixties by Dyson and Lenard for a system of nuclei and electrons interacting with each other through electrostatic forces. The foundations on which stability ultimately rests are the **uncertainty principle** (that enters via the Schrödinger equation) and the **Pauli exclusion principle**. The first quantifies the cost of kinetic energy when localizing a quantum mechanical particle and explains successfully the size of hydrogen and helium. The second, which introduces some kind of "repulsion" among electrons, comes into play when dealing with many particles. That it is necessary for stability of bulk matter to hold was shown by Dyson [8].

It is difficult to overstate the importance of these two principles. They seem to be true universal principles that explain facts about nature from the smallest distance scales up to objects of astronomical size (ignoring the somewhat annoying fact that we have not found a true relativistic quantum theory as of yet). For example, they yield the correct magnitude for the Chandrasekhar radius of white dwarfs. For further details, quotes, and more examples about the astonishing range of validity of quantum mechanics the reader should consult the lucid exposition [13], which can also be found in [14] along with many of the articles mentioned in the references.

The goal of this note is to explain a few of the latest developments not covered in [13], and this is stability of matter interacting with fields. First,

Michael Loss is professor of mathematics at the Georgia Institute of Technology. His e-mail address is loss@math.gatech.edu.

we give in a telegraphic style the ingredients necessary to formulate the stability result of Dyson and Lenard.

A piece of matter consists of charged particles, K nuclei, and N electrons interacting with each other according to Coulomb's law. We shall treat the nuclei as classical particles or, what is the same, as particles of infinite mass that are located at the positions $R_1, \dots, R_K$. The state of N electrons is described by a complex–valued wave function Ψ of the space and spin coordinates of each electron; i.e.,

$$\Psi = \Psi(x_1, \sigma_1; \dots; x_N, \sigma_N) .$$

The important point is that nature allows two kinds of Ψ's: the symmetric ones (bosons) or, as in the case of electrons, antisymmetric ones (fermions), the latter being the precise formulation of the Pauli exclusion principle. We denote the Hilbert-space of all N-electron wave functions by

$$\mathcal{H} = \wedge^N L^2(\mathbb{R}^3; \mathbb{C}^2) ,$$

the N-fold antisymmetric tensor product of $L^2(\mathbb{R}^3; \mathbb{C}^2)$. That the wave function of a single electron has values in $\mathbb{C}^2$ is related to the spin of the electron. This will be important later.

The Hamiltonian that describes that system is given, in suitable units, by

$$H = \sum_{j=1}^{N} \tau_j + V_c . \tag{1}$$

The standard choice for the kinetic energy τ is

$$\tau = p^2 = -\Delta , \tag{2}$$

which is the kinetic energy operator of a nonrelativistic electron. The symbol $p = -i\nabla$ stands for the physical momentum, and the subscript j reminds us that τ_j acts on the coordinates of the j-th particle.

V_c is the Coulomb interaction among the particles and is given by

$$V_c = -\sum_{i=1,j=1}^{N,K} \frac{Z_j}{|x_i - R_j|} + \sum_{i<j}^{N} \frac{1}{|x_i - x_j|} + \sum_{i<j}^{K} \frac{Z_i Z_j}{|R_i - R_j|} . \tag{3}$$

The three terms describe, in succession, the attraction of the nuclei and electrons, the repulsion among the electrons, and the repulsion among the nuclei. For simplicity we shall assume that all the nuclei have the same charge Z.

The ground-state energy $E(N, K, Z, R)$ of the system is defined to be the lowest element of the spectrum of the operator H as an operator on $\mathcal{H}$, i.e., on the wave functions that obey the Pauli exclusion principle. As indicated, this number depends on the number of nuclei, their charges and their positions, and the number of electrons.

The main theorem is now:

Theorem 1 (Stability of Matter). The ground-state energy $E(N, K, Z, R)$ of the system defined by (1)– (3) satisfies the estimate

$$E(N, K, Z, R) \geq -C(Z)(N + K) ,$$

where the constant $C(Z)$ is independent of N, K and the position of the nuclei $R_1, \dots, R_K$.

This theorem was first proved by Dyson and Lenard [9] and later simplified and considerably improved by Lieb and Thirring [18].

Various extensions of this fundamental result have been investigated. There is a "relativistic" version due to Conlon [6] where $\tau = \sqrt{p^2 + m^2} - m$, the relativistic kinetic energy. His results were later improved by Fefferman-de La Llave [11] and Lieb–Yau [19]. For a discussion of these results the reader may consult [13] and [15] for a particularly simple approach. The main feature we have to mention is that stability of matter for relativistic systems poses bounds on the fine structure constant α. This constant is a dimensionless combination of the elementary charge e, Planck's constant h, and the speed of light c and is given by $\alpha = 2\pi e^2/hc$ with an approximate value of $1/137.04$. It measures the effective strength of the electromagnetic interaction. This constant will recur in the discussion below.

Now let us turn to the recent developments where matter is coupled to magnetic fields and start with some introductory observations. A static magnetic field is a divergence-free vector field in $\mathbb{R}^3$. We can always find a "vector potential" $A(x)$ such that

$$\mathrm{curl} A(x) = B(x) .$$

Clearly, A is only specified up to a gauge transformation $A \to A + \nabla\chi$ for some function χ.

The electron interacts with the magnetic field through "minimal coupling"; i.e., one replaces the momentum p by $p - A$ and hence obtains for the kinetic energy

$$|p - A|^2 ,$$

where the vector potential A, as in the case of a scalar potential, acts as a multiplication operator. Recall that p is a shorthand for $-i\nabla$. This Hamiltonian describes the interaction of a charged particle with a magnetic field. For a single particle the addition of a magnetic field leads to a raising of the ground-state energy, as can be seen from the pointwise "diamagnetic" inequality

$$|(p - A)\psi(x)| \geq |\nabla|\psi(x)|| .$$

It was shown in [1, 7] that as a consequence of this inequality the estimate of Theorem 1 holds with a constant that is independent of the magnetic field. In general, the field will lower the energy but not by so much as to violate the bound of Theorem 1. The real dependence of the ground-state energy on the magnetic field is an important physical problem that is not very well understood.

A special case that can be solved explicitly is an otherwise free electron in the $x-y$ plane that moves in a constant magnetic field $(0,0,B)$, $B>0$. The lowest energy state of this problem is given by the function

$$\psi(x,y)=(\frac{B}{2\pi})^{1/2}e^{-B(x^2+y^2)/4}\ ,$$

with an energy B. For an electron moving in three space the kinetic energy is raised by adding $-\partial^2/\partial z^2$, and hence the energy of the electron is again bounded below by B. The wave functions with energies close to B tend to be localized perpendicular to the field lines.

The spin of the electron introduces an additional complication. The electron behaves to some extent like a magnetized needle; i.e., it has a magnetic moment μ. The coupling between this moment and the magnetic field is given by $-\mu\cdot B$. Recall that the wave function of a single electron has spin as an additional variable; i.e., it is an element of $L^2(\mathbb{R}^3:\mathbb{C}^2)$. It is convenient to think of it as a "spinor", i.e., a two-component wave function

$$\Psi=\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\ .$$

The "magnetic moment" is now given by the vector $\langle\Psi,\sigma\Psi\rangle$, where σ are the three Pauli matrices

$$\sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\ \sigma_2=\begin{pmatrix}0&-i\\-i&0\end{pmatrix},\ \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

and $\langle,\rangle$ is the standard inner product on $\mathbb{C}^2$. The part of the energy that describes the interaction of the electron with the magnetic field is given by

$$\tau=|p-A|^2-\sigma\cdot B=[\sigma\cdot(p-A)]^2\ , \tag{4}$$

which is a positive matrix-valued operator called the Pauli-Hamiltonian.

If we again consider the case where the magnetic field is constant pointing in the z-direction, we see that the $\sigma\cdot B$ term cancels the ground-state energy B of $|p-A|^2$. The operator is still positive but considerably weakened. An uncertainty principle measures the cost of localizing the wave function; i.e., strong localization is only possible for large kinetic energies. In the present example we see that the cost of localizing the electron is paid in part by the magnetic field; the kinetic energy measures only localization in the z-direction.

Our model for bulk matter is now the many-body Hamiltonian given by (1) and (3), with τ given by (4). We shall refer to this model as the **nonrelativistic model**. Its ground-state energy $E(B,Z,R,N,K)$ will be a complicated function of the magnetic field. A disturbing fact is that by choosing large magnetic fields but keeping all the other parameters fixed, $E(B,Z,R,N,K)$ can be made arbitrarily negative. According to what was mentioned above this may not come as a great surprise, due to the weakening of the kinetic energy. In fact, it is shown in [2] that for hydrogen in a strong *constant* magnetic field B the ground-state energy behaves as $-(\ln B)^2$ to leading order as $B\to\infty$.

Matter is constantly surrounded by magnetic fields; in fact, one knows from quantum electrodynamics that the fluctuations of the electromagnetic field are quite large but are apparently of no concern. Certainly it never happens that matter becomes unstable through a spontaneous fluctuation of the field. The energy cost for creating a magnetic field is measured in terms of the field energy given in suitable units by

$$\frac{1}{8\pi\alpha^2}\int_{\mathbb{R}^3}|B(x)|^2\mathrm{d}x\ .$$

It is therefore physically reasonable that the field energy stabilizes $E(B,Z,R,N,K)$. More precisely, we expect that

$$E(B,Z,R,N,K)+\frac{1}{8\pi\alpha^2}\int_{\mathbb{R}^3}|B(x)|^2\mathrm{d}x \geq -C(Z)(N+K) \tag{5}$$

uniformly in B.

Let us emphasize again that the constant $C(Z)$ should not depend on the number of particles, the position of the nuclei, or the magnetic field.

We usually think of a magnetic field as external; i.e., the magnetic field acting on a piece of matter is produced by some external currents and not by that piece itself. Note, however, that in the above inequalities we did not constrain the magnetic field to be external. Thus one may view the above model as a semiclassical caricature of nonrelativistic quantum electrodynamics, i.e., a "theory" that incorporates the fields as dynamical quantities that fully interact with matter, and (5) amounts to stability of that theory—certainly a first, albeit a small, step toward the full problem with a quantized electromagnetic field.

A surprising fact is that even (5) cannot hold without qualifications; a collapse occurs if $Z\alpha^2$ is large enough. It is instructive to sort this out in the case of a hydrogenic atom [12]. The energy functional for this system is

$$\mathcal{E}(\Psi, A) = \Big(\sigma \cdot (p - A)\Psi, \sigma \cdot (p - A)\Psi\Big) - Z\Big(\Psi, \frac{1}{|x|}\Psi\Big) + \varepsilon \int_{\mathbb{R}^3} |B(x)|^2 \mathrm{d}x. \tag{6}$$

Here $\varepsilon = 1/(8\pi\alpha^2)$ and $\big(\cdot, \cdot\big)$ denotes the inner product in $L^2(\mathbb{R}^3 : \mathbb{C}^2)$. Recall that $B = \mathrm{curl}A$ and that we minimize over A as well as the spinor Ψ. Also recall that we have used (4) to display the interaction of the electron with the magnetic field as a complete square. The catastrophic features of this model come in through the possibility of zero modes of the three-dimensional Dirac equation

$$\sigma \cdot (p - A)\Psi = 0\ . \tag{7}$$

The point about this equation is that the solution Ψ should be square summable and the field B associated with A should have finite energy. A simple exercise in scaling then shows that if we set

$$\Psi_\lambda(x) = \lambda^{\frac{3}{2}}\Psi(\lambda x)$$

and

$$A_\lambda(x) = \lambda A(\lambda x)\ ,$$

then Ψ_λ and A_λ still satisfy the three-dimensional Dirac equation (7). If we choose Ψ_λ and A_λ as trial functions in (6), then

$$\mathcal{E}(\Psi_\lambda, A_\lambda) = \lambda[-Z\Big(\Psi, \frac{1}{|x|}\Psi\Big) + \varepsilon \int_{\mathbb{R}^3} |B(x)|^2 \mathrm{d}x]\ ,$$

which becomes arbitrarily negative provided

$$8\pi Z\alpha^2 > \frac{\int_{\mathbb{R}^3} |B(x)|^2 \mathrm{d}x}{\Big(\Psi, \frac{1}{|x|}\Psi\Big)}\ .$$

The following example, discovered in [21], shows that zero modes exist. Set

$$A(x) = \frac{3}{(1 + x^2)^2}\Big((1 - x^2)w + 2(w \cdot x)x + 2w \times x\Big)$$

and

$$\Psi(x) = \frac{1 + i\sigma \cdot x}{(1 + |x|^2)^{3/2}}\phi_0\ ,$$

where the vector w is given by $w = \langle \phi_0, \sigma\phi_0 \rangle$ and ϕ_0 is a constant spinor of length one. A straightforward computation shows that (7) is satisfied. The magnetic field B that belongs to A is given by

$$B(x) = \frac{12}{(1 + x^2)^3}\Big((1 - x^2)w + 2(w \cdot x)x + 2w \times x\Big)\ .$$

The pattern of its field lines is given by the Hopf-fibration on S^3 viewed on $\mathbb{R}^3$ via stereographic projection. Thus the field lines are Euclidean circles winding around a family of nested tori.[1]

In physical terminology we see here a breakdown of the uncertainty principle; that is, the kinetic energy does not get large if the wave function is squeezed (although the field energy does, but in an insufficient fashion for $Z\alpha^2$ big enough).

Even if $Z\alpha^2$ is small, another collapse can occur. We give here a heuristic reason that can be turned into a rigorous proof quite easily [17]. The constants we get are not correct but serve the purpose of illustration.

Consider a single electron in a field of K nuclei all having the same charge Z. Pick again a field B and a spinor Ψ such that (7) is satisfied—i.e., there is no kinetic energy. For a fixed length scale L the field energy is of the order $(\alpha^2 L)^{-1}$, the interaction energy of the electron with the nuclei is of the order $-ZKL^{-1}$, and the repulsive energy of the nuclei is of the order $(ZK)^2 L^{-1}$. Thus the total energy is of the order

$$L^{-1}(\alpha^{-2} + [ZK - \frac{1}{2}]^2 - \frac{1}{4})\ .$$

If we forget for the moment that Z is an integer and allow arbitrarily small values for it, we can set $ZK = 1/2$ by choosing a large number of nuclei. If α is larger than 2, then the total energy can be made arbitrarily negative as L tends to zero. We thus see that the smallness of $Z\alpha^2$ **and** the fine structure constant α are required for stability.

The full estimate (5) was proved by Fefferman [10] for α sufficiently small, and shortly thereafter the following theorem was proved by Lieb, Loss, and Solovej [16]:

Theorem 2. Estimate (5) holds provided that $Z\alpha^2 \le 0.041$ and $\alpha \le 0.06$. Moreover, if we fix $\alpha = 1/137$, then we have stability for Z up to 1050.

The fact that $Z = 1050$ is far outside the physical range indicates that it is safe to apply perturbation theory for all physically reasonable Z's.

One should emphasize that the above model is of a very rudimentary nature. The electromagnetic field should be quantized; after all, excited atoms can emit radiation spontaneously. The electrons should be treated in a relativistic fashion; after all, in heavy nuclei, electrons acquire relativistic speeds, and the Dirac equation is the true equation that governs a single relativistic electron. Finally, electron-positron pair production can occur that requires the electron "field" to be quantized too. So far a consistent theory of quantum electrodynamics has eluded us. Humbleness requires

[1] *See the cover of the May 1997 issue of the* Notices *for a picture.*

that we limit ourselves to caricatures, with the hope that they contain certain features that will persist in the theory to be; here is a recent one due to Lieb, Siedentop, and Solovej [20].

The Dirac–Hamiltonian of a relativistic electron interacting with a magnetic field can be written as

$$\tau(A) = \alpha \cdot (-i\nabla + eA) + m\beta \,, \tag{8}$$

with four-by-four matrices

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}$$

and

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

This operator, which acts on four spinors, has the obvious problem that it is unbounded below, and it is difficult to make sense of the notion of stability in this context.

The following approach due to Brown and Ravenhall [4], also Bethe and Salpeter [5], yields a model of matter that looks reasonable for energies that are below the electron-positron pair production threshold. In this model one restricts the electron wave function to lie in the positive energy subspace of the operator (8). Call this A-dependent Hilbert space H_A, and define the N-electron space by the antisymmetric tensor product

$$\mathcal{H}_A = \wedge^N H_A \,. \tag{9}$$

If we denote by P_A the projection onto $\mathcal{H}_A$ and by h_A the many-body Hamiltonian given in the usual way by (1) and (3) with the kinetic energy given by (8), then the operator of interest is

$$H(A) = P_A h_A P_A \,. \tag{10}$$

The stability problem is now formulated in complete analogy with the nonrelativistic version (5); i.e., we consider the ground-state energy of this system with a fixed field, add the field energy, and ask whether there is a lower bound to this expression that is proportional to the total particle number with a proportionality constant that is independent of the position of the nuclei, the magnetic field, and the number of particles. In fact, since all the terms have the same dimension, the constant better be zero, because otherwise we could drive the energy easily to $-\infty$ by scaling.

It was proved in [20] that this is indeed the case provided $Z\alpha$ and α are small enough.

Theorem 3. Define $\alpha_c > 0$ to be the solution of

$$1 - (\alpha_c/\beta)^2 = (16\pi L \alpha_c)^{2/3} \,,$$

where

$$\frac{1}{\beta} = (\pi/2)Z + 2.80Z^{2/3} + 1.30 \,,$$

and $L \approx 0.06$. Then $H(A)$ is stable provided $\alpha \le \alpha_c$.

The operator $H(A)$ is a nonlocal operator, and the dependence on A is rather worrisome. Note that the operator (10) is not gauge invariant, but rather gauge covariant; the Hilbert-space has to be transformed too.

An alternative would be to repeat the same procedure but to choose as our one-particle Hilbert space the positive spectral subspace of the **free** Dirac operator, i.e., (9) without the A-field. Note that formally the kinetic energy operator stays the same, but we have restricted the states in a different fashion. This procedure can be rejected on the grounds that it destroys any form of gauge covariance. But things are worse. It was shown in [20] that for any $\alpha > 0$ and for any $Z > 0$ there exist N, K sufficiently large such that $H(A)$ restricted to the positive energy states of the free operator is unbounded below. Thus the stability criterion decides which of the two possible restrictions is the correct choice.

Where to go from here? The main drawbacks of the above models are that neither the electromagnetic field nor the electron field are quantized. There has been some progress in that direction. In [3] the nonrelativistic model is considered with a quantized electromagnetic field. In order that the theory makes sense, however, the field variables have to be cut off for high momenta. As a result, the stability constant depends on this cutoff.

There is usually a sharp division between mathematical rigor and physical intuition. Exponents of both areas are only too willing to attest to that. The division, of course, is artificial, and the research area that I just described makes a good case for that. We now have a clear picture of why matter is stable, and this has been achieved by a combination of physical ideas and nontrivial mathematical concepts. It has to be emphasized that mathematics is vital to the whole enterprise and is not merely a "rigorization" of physical ideas.

Acknowledgment: The author thanks Mary Beth Ruskai and Elliott Lieb for their help in preparing this article.

References

[1] J. AVRON, I. HERBST, and B. SIMON, *Schrödinger operators with magnetic fields: I. General interactions*, Duke. Math. J. **45** (1978), 847–883.

[2] ———, *Schrödinger operators with magnetic fields: III. Atoms in homogeneous magnetic field*, Commun. Math. Phys. **79** (1981), 529–572.

[3] L. BUGLIARO, J. FRÖHLICH, and G.-M. GRAF, *Stability of quantum electrodynamics with nonrelativistic matter*, Phys. Rev. Lett. **77** (1996), 3494–3497.
[4] G. BROWN and D. RAVENHALL, *On the interaction of two electrons*, Proc. Roy. Soc. London A. **208** (A 1095) (1952), 552–559.
[5] H. A. BETHE and E. E. SALPETER, *Quantum mechanics of one- and two-electron atoms*. (S. Flügge, editor), Handbuch der Physik, XXXV, Springer, Berlin, first edition, 1957.
[6] J. G. CONLON, *The ground state energy of a classical gas*, Commun. Math. Phys. **94** (1984), 439–458.
[7] J. M. COMBES, R. SCHRADER, and R. SEILER, *Classical bounds and limits for energy distributions of Hamiltonian operators in electromagnetic fields*, Ann. Phys. III (1978), 1–18.
[8] F. J. DYSON, *Ground-state energy of a finite system of charged particles*, J. Math. Phys. **8** (1967), 1538–1545.
[9] F. J. DYSON and A. LENARD, *Stability of Matter I*, J. Math Phys. **8** (1967), 423–434; *Stability of Matter II*, J. Math Phys. **9** (1967), 698–711.
[10] C. FEFFERMAN, *Stability of Coulomb systems in a magnetic field*, Proc. Nat. Acad. Sci. USA **92** (1995), 5006–5007.
[11] C. FEFFERMAN and R. DE LA LLAVE, *Relativistic stability of matter*, I., Rev. Math. Iberoamericana **2** (1986), 119–215.
[12] J. FRÖHLICH, E. H. LIEB, and M. LOSS, *Stability of Coulomb systems with magnetic fields. I: The one electron atom*, Commun. Math. Phys. **104** (1986), 251–270.
[13] E. H. LIEB, *The stability of matter: From atoms to stars*, Bull. Amer. Math. Soc. (N.S.) **22** (1990), 1–49.
[14] ____, *The stability of matter: From atoms to stars*, Selecta of Elliott H. Lieb (W. Thirring, editor) second edition, Springer Verlag, 1997.
[15] E. H. LIEB, M. LOSS, and H. SIEDENTOP, *Stability of relativistic matter via Thomas-Fermi theory*, Helv. Phys. Acta **69** (1996), 974–984.
[16] E. H. LIEB, M. LOSS, and J.-P. SOLOVEJ, *Stability of matter in magnetic fields*, Phys. Rev. Lett. **75** (1995), 985–989.
[17] E. H. LIEB and M. LOSS, *Stability of Coulomb systems with magnetic fields. II: The many-electron atom and the one-electron molecule*, Commun. Math. Phys. **104** (1986), 271–282.
[18] E. H. LIEB and W. E. THIRRING, *Bound for the kinetic energy of fermions which proves the stability of matter*, Phys. Rev. Lett. **35** (1975), 687–689; Erratum, Phys. Rev. Lett. **36** (1975), 1116.
[19] E. H. LIEB and H.-T. YAU, *The stability and instability of relativistic matter*, Commun. Math. Phys. **118** (1988), 177–213.
[20] E. H. LIEB, H. SIEDENTOP, and J.-P. SOLOVEJ, *Stability and instability of relativistic electrons in classical electromagnetic fields*, to appear in J. Stat. Phys.
[21] M. LOSS and H.-T. YAU, *Stability of Coulomb systems with magnetic fields. III: Zero energy bound states of the Pauli operator*, Commun. Math. Phys. **104** (1986), 283–290.