Vol. 24(1993) *ACTA PHYSICA POLONICA B* No 4

LECTURE ON FEFFERMAN'S PROOF OF THE ATOMIC AND MOLECULAR NATURE OF MATTER*

PH.A. MARTIN

Institut de Physique Théorique
Ecole Polytechnique Fédérale de Lausanne
PHB-Ecublens, CH-1015 Lausanne, Switzerland

(Received December 8, 1992)

An elementary account of Fefferman's analysis of the quantum electron-proton system is presented. Assuming an optimal form of the stability of matter lower bound, it is shown that in an appropriate low density and low temperature limit, the electron-proton system behaves as a free gas of hydrogen atoms.

PACS numbers: 05.30. -d, 02.50. Kd, 03.65. Ge

1. Introduction

In these lectures, I will report on a beautiful piece of work by Fefferman [1, 2] concerning our basic understanding of the atomic and molecular constitution matter in the framework of non relativistic equilibrium quantum statistical mechanics.

But what is to be understood? After all, it is an evident and unquestionable fact of everyday life that matter under usual conditions[1] is made of atoms and molecules. Moreover, we have all been taught that one of the main successes of quantum mechanics is precisely to predict accurately atomic and molecular binding energies. For a gentle start, let me remind you of the hydrogen atom Hamiltonian (one electron of charge e and mass

---

* Presented at the V Symposium on Statistical Physics, Zakopane, Poland, September 21-30, 1992.
A French version of this text can be found in Note di matematica e fisica, CERFIM, Locarno (Switzerland)

[1] By usual conditions, we mean that typical energies are in a range where nuclear reactions and relativistic effects can be neglected.

m_e and one proton of charge $-e$ and mass m_p) written in a coordinate frame where the center of mass is at rest

$$H = \frac{|\underline{p}|^2}{2m} - \frac{e^2}{|\underline{r}|}. \tag{1.1}$$

In (1.1), $|\underline{p}|^2/2m$ is the relative kinetic energy (m is the reduced mass) and $-e^2/|\underline{r}|$ is the Coulomb potential energy of the proton and electron at distance $|\underline{r}|$.

The energy spectrum determined by the stationary Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\Psi(\underline{r}) - \frac{e^2}{|\underline{r}|}\Psi(\underline{r}) = E\Psi(\underline{r}) \tag{1.2}$$

can be divided in two parts. The negative eigenvalues correspond to bound states. In particular to the lowest one, $E_0 = -{}^1\!/\!_2(e^2/a_\mathrm{B})$, corresponds the well known exponentially decaying ground state wave function

$$\Psi_0(\underline{r}) = \frac{1}{\left(8\pi a_\mathrm{B}^3\right)^{1/2}} \exp\left(-\frac{|\underline{r}|}{2a_\mathrm{B}}\right), \tag{1.3}$$

where $a_\mathrm{B} = \hbar^2/me^2$ is the Bohr radius. On the other hand, the non normalizable solutions of (1.2) with positive E describe ionized states where the electron and the proton are dissociated.

Is the very existence of the ground state wave function (1.3) sufficient to justify the formation of atoms in condensed matter? It is far to be the case. In fact, most of the time, we observe atoms and molecules in phases (gas and liquids) which have a non vanishing temperature. The center of mass of the atom has always some kinetic energy and undergoes collisions with other atoms. Under these conditions, taking the principles of equilibrium statistical mechanics for granted, the Gibbs prescription applied to our hydrogen atom tells us that the probability $p(E)$ to find the atom in a state of energy E at inverse temperature $\beta = (k_\mathrm{B}T)^{-1}$ (k_B = Boltzmann constant) is

$$p(E) = \frac{1}{Z}\mathrm{e}^{-\beta E} \qquad Z = \sum_E \mathrm{e}^{-\beta E}. \tag{1.4}$$

The sum in Z runs over all possibles energies of the electron-proton system.

Then we can ask the question: what is the probability for a proton and an electron confined in a region Ω of volume $|\Omega|$ to form an hydrogen atom? The main observation is that the contribution of the ionized states to Z is of the order $|\Omega|$. Indeed, taking Ω a cube of side L with periodic boundary conditions and assuming for simplicity that ionized states are strict plane waves[2] with wave numbers $\underline{k} = 2\pi/L$, the contribution to Z of the positive energy states is of the order

$$\sum_{E \geq 0} \mathrm{e}^{-\beta E} = \sum_{\underline{k}} \exp\left(-\beta \frac{\hbar^2 |\underline{k}|^2}{2m}\right)$$

$$\approx |\Omega| \frac{1}{(2\pi)^3} \int d\underline{k} \exp\left(-\beta \frac{\hbar^2 |\underline{k}|^2}{2m}\right) = \frac{|\Omega|}{\lambda^3}, \tag{1.5}$$

where $\lambda = \left(2\pi\beta\hbar^2/m\right)^{1/2}$ is the thermal wave length. Thus, according to (1.4), the probability to find the hydrogen atom in its ground state is

$$p(E_0) \sim \frac{\lambda^3}{|\Omega|} \mathrm{e}^{-\beta E_0}. \tag{1.6}$$

Since $\lambda \approx 10^{-10}$ cm, this probability is extremely small if Ω is macroscopic. In other words, at fixed non zero temperature, an electron and a proton in a macroscopic volume Ω have no chance to form a bound state. This is an entropic effect: the number of configurations where to two charges are dissociated becomes infinitely large.

What is the reason for this apparent paradox? It is due to the fact that we have considered only a single electron and a single proton in the region Ω. But to reach the thermal equilibrium distribution (1.4), it is necessary that our two charges can exchange energy by collision with other particles, hence to deal with a system of charges with non vanishing density. The number of electrons N and of protons M must be extensive

$$N = \rho_{\mathrm{e}} |\Omega|, \qquad M = \rho_{\mathrm{p}} |\Omega|. \tag{1.7}$$

To ensure global neutrality we take equal electronic and protonic densities $\rho_{\mathrm{e}} = \rho_{\mathrm{p}} = \rho_{\mathrm{a}}$. Then an electron-proton pair has a finite effective volume $1/\rho_{\mathrm{a}}$ and (1.6) does not apply any more: the probability to observe a bound state does not have to be very small in a macroscopic volume. The conclusion is that atoms can form only if there is a non vanishing density of them, and only if they have (at least weak) mutual interactions. One must therefore face the difficult problem that a demonstration of the atomic nature of matter cannot be brought in a first elementary course on quantum mechanics, but necessarily involves the study of the many-body situation. Several aspects of these questions are very well discussed in Chap. 3 of Ref. [3].

---

[2] Coulombic extended states are plane waves with a logarithmic correction. This does not change the conclusion (1.5).

2. Elements in the statistical mechanics of protons and electrons

2.1. The Hamiltonian

As we have just concluded, we are led to consider the Hamiltonian H^{Ω}_{NM} of N electrons and M protons in a region Ω. For brevity, we denote

$$x = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N), \qquad y = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_M), \tag{2.1}$$

the coordinates $\underline{x}_i$, $i = 1, \ldots, N$, $\underline{y}_j$, $j = 1, \ldots, M$, of the electrons and of the protons and

$$\Delta_x = \sum_{i=1}^{N} \Delta \underline{x}_i , \qquad \Delta_y = \sum_{j=1}^{M} \Delta \underline{y}_j \tag{2.2}$$

the corresponding Laplacians. Then

$$H^{\Omega}_{NM} = -\frac{\hbar^2}{2m_e}\Delta_x - \frac{\hbar^2}{2m_p}\Delta_y + \sum_{i<k}^{N} \frac{e^2}{|\underline{x}_i - \underline{x}_k|} + \sum_{i<k}^{M} \frac{e^2}{|\underline{y}_i - \underline{y}_k|} - \sum_{i,k}^{NM} \frac{e^2}{|\underline{x}_i - \underline{y}_k|}, \tag{2.3}$$

where the first two terms are the kinetic energy of the particles, and the three last ones are the repulsive (attractive) Coulomb potentials between charges of the same (opposite) charges. For each value of N and M we have the stationary Schrödinger equation

$$H^{\Omega}_{NM} \Psi^{\Omega}_{NM\alpha}(x, y) = E^{\Omega}_{NM\alpha} \Psi^{\Omega}_{NM\alpha}(x, y) . \tag{2.4}$$

The index α represents all the quantum numbers necessary for the labelling of the energies $E^{\Omega}_{NM\alpha}$ and eigenfunctions $\Psi^{\Omega}_{NM\alpha}(x, y)$. Since the Fermi statistics applies to the electrons and to the protons separately, we only keep solutions of (2.4) that are antisymmetric[3] under the permutations of the x_i and of the y_i. Moreover, since the particles are confined in Ω, we impose Dirichlet conditions on the boundary $\partial\Omega$ of Ω, *i.e.* the wave function vanishes if anyone of its arguments belongs to $\partial\Omega$. The wave functions $\Psi^{\Omega}_{NM\alpha}(x, y)$ are normalized to one in Ω.

---

[3] The antisymmetry applies to the simultaneous exchange of the configurational and spin variables. Aside from the application of the Pauli principle, the spins will play no other role in the sequel, and we shall not indicate them.

2.2. The grand canonical ensemble

The whole statistical mechanical discussion will take place within the grand canonical ensemble. In this ensemble, the probability density to find N electrons at x and M protons at y in Ω is defined by

$$\frac{e^{\beta\mu(N+M)}}{Z(\mu,\beta,\Omega)}\sum_{\alpha}\exp\left[-\beta E^{\Omega}_{NM\alpha}\right]|\Psi^{\Omega}_{NM\alpha}(x,y)|^2, \tag{2.5}$$

where

$$Z(\mu,\beta,\Omega)=1+\sum_{\substack{N,M\\(N,M)\neq(0,0)}}\exp\left[\beta\mu(N+M)\right]\sum_{\alpha}\exp\left[-\beta E^{\Omega}_{NM\alpha}\right] \tag{2.6}$$

is the grand-partition function. In (2.5) and (2.6) μ is the chemical potential associated to the total particle density $|\Omega|^{-1}\langle N+M\rangle$ (see the remark after equation (2.17)). In particular, integrating (2.5) over all positions x, y of the charges in Ω gives the probability $p_{NM}(\mu,\beta,\Omega)$ to find exactly N electrons and M protons anywhere in Ω

$$\begin{aligned}p_{NM}(\mu,\beta,\Omega)&=\frac{e^{\beta\mu(N+M)}}{Z(\mu,\beta,\Omega)}\sum_{\alpha}\exp\left[-\beta E^{\Omega}_{NM\alpha}\right]\\&=\frac{Z(\mu,\beta,N,M,\Omega)}{Z(\mu,\beta,\Omega)},\end{aligned} \tag{2.7}$$

where we have introduced the canonical partition function (up to the factor $\exp[\beta\mu(N+M)]$)

$$Z(\mu,\beta,N,M,\Omega)=e^{\beta\mu(N+M)}\sum_{\alpha}\exp\left[-\beta E^{\Omega}_{NM\alpha}\right]. \tag{2.8}$$

As usual, the density $\rho(\mu,\beta,\Omega)$ and the pressure $P(\mu,\beta,\Omega)$ at fixed chemical potential μ are given by

$$\rho(\mu,\beta,\Omega)=\frac{\langle N+M\rangle}{|\Omega|}=\frac{1}{|\Omega|}\beta^{-1}\frac{\partial}{\partial\mu}\ln Z(\mu,\beta,\Omega), \tag{2.9}$$

and

$$P(\mu,\beta,\Omega)=\beta^{-1}\frac{1}{|\Omega|}\ln Z(\mu,\beta,\Omega). \tag{2.10}$$

Let us recall an important factorization property of the grand-partition function. If one can decompose $\Omega=\Omega_1\cup\Omega_2$ into the disjoint union of two subdomains such that the particles in Ω_1 do not interact with the particles in Ω_2, then

$$Z(\mu,\beta,\Omega) = Z(\mu,\beta,\Omega_1)Z(\mu,\beta,\Omega_2). \tag{2.11}$$

2.3. The stability of matter

The stability of matter property is the statement that the energy per particle in the Coulomb system is bounded from below, *i.e.* there exists a constant $C < \infty$ such that

$$\inf_{N,M,\Psi} \frac{(\Psi, H_{NM}\Psi)}{N+M} \geq -C\,, \tag{2.12}$$

where the infimum is taken over all normalizable antisymmetric wave functions Ψ and all $N, M = 0, 1, 2\ldots$[4] This property was first established in [4] and we refer to [5] for a recent and complete discussion. In the following considerations, the numerical value of the lower bound (2.12) will play an important role. Consider for simplicity the neutral case $N = M$ and the infimum of the energy per electron-proton pair E_*

$$E_* = \inf_{N,\Psi} \frac{(\Psi, H_{NN}\Psi)}{N}\,. \tag{2.13}$$

It is presently proven [5] that, in unit of the ground state energy $|E_0|$ of the hydrogen atom,

$$E_* \geq -5.6\;|E_0|\,. \tag{2.14}$$

The constant 5.6 occurring in (2.14) is presumably not optimal and could be probably lowered to 2 or even less than 2. Indeed, going from a dilute cold plasma (all charges far apart and at rest) to the ground state of a molecular hydrogen crystal, the observed gain in binding energy per electron-proton pair does not exceed $2|E_0|$ (Fig. 1).

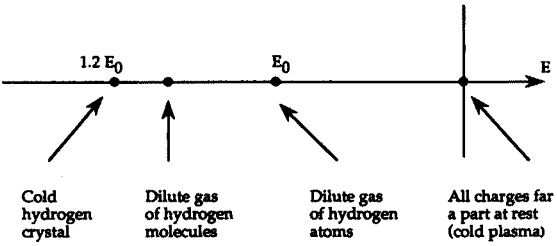

Fig. 1.

---

[4] Here H_{NM} is the Hamiltonian (2.3) in infinite space.

Fefferman's analysis will require a more refined formulation of the stability of matter bound which is stated in the following form for electrons and protons:

There exists a constant K, $0 < K < |E_0|$, such that

$$(\Psi, H_{NM}\Psi) \geq -K(N + M - 1), \quad (N, M) \neq (0,0) \text{ and } (1,1). \tag{2.15}$$

The important point in the assumption (2.15) is the strict inequality $K < |E_0|$ for all cases except of course for the hydrogen atom[5]. This is not yet proven but compatible with all calculated or measured binding energies. For instance we obtain from (2.15) that $E_* \geq -2K > -2|E_0|$, an inequality which is likely true as discussed above. Consider also the ground state energy E_{22} of the hydrogen molecule (taking the spins into account). Then (2.15) implies $E_{22} \geq -3K > -3|E_0|$ or equivalently

$$|E_{22} - 2E_0| < |E_0|, \tag{2.16}$$

i.e. the binding energy gained by the formation of an hydrogen molecule is less than the binding energy of the atom itself, a well known fact. From now on, we shall adopt (2.15) as a plausible assumption.

2.4. The thermodynamic limit

The possibility to describe bulk matter by thermodynamic laws requires the existence of the thermodynamic limit

$$\begin{aligned} \lim_{|\Omega| \to \infty} \rho(\mu, \beta, \Omega) &= \rho(\mu, \beta) \\ \lim_{|\Omega| \to \infty} P(\mu, \beta, \Omega) &= P(\mu, \beta). \end{aligned} \tag{2.17}$$

The proof of the existence of these limits is the content of the Lieb–Lebowitz theorem [5, 6]. We note that we could have introduced two different chemical potentials μ_e and μ_p for the electrons and the protons. Writing

$$\mu_e N + \mu_p M = \left(\frac{\mu_e + \mu_p}{2}\right)(N + M) + \left(\frac{\mu_e - \mu_p}{2}\right)(N - M),$$

it is the conclusion of the theorem that the thermodynamic functions depend only on $\mu = (\mu_e + \mu_p)/2$. This implies neutrality on average (*i.e.* $\rho_e = \rho_p$ in the bulk).

---

[5] For $N = M = 0$, we set $H_{00} = 0$.

3. Fefferman's theorem

Assume that we would be able to evaluate the thermodynamic functions. How do we recognize that we are in presence of an ensemble of hydrogen atoms rather than dissociated electrons and protons? In order to make our goal explicit, let me adopt for a moment the view point of the chemist. When a chemist describes a gas of hydrogen atoms, he considers that the atoms are performed entities. In the most elementary description, the energy of a performed atom consists of the kinetic energy $|\underline{P}|^2/2\overline{m}$ of its center of mass ($\overline{m} = m_e + m_p$) and of the binding energy E_0 of the ground state; moreover, the atoms are independent point particles. In this simplified model the Hamiltonian of the atoms is

$$H_N^{\text{atom}} = \sum_{i=1}^{N} \frac{|\underline{P}_i|^2}{2\overline{m}} + NE_0 . \tag{3.1}$$

Treating the motion of the center of mass as a classical degree of freedom, the corresponding classical grand-partition function is

$$\begin{aligned} Z_{\text{atom}}(\mu, \beta, \Omega) &= \sum_{N=0}^{\infty} \frac{e^{2\beta\mu N}}{N! h^{3N}} \\ &\times \int d\underline{p}_1 \cdots \int d\underline{p}_N \int_{\Omega} d\underline{x}_1 \cdots \int_{\Omega} d\underline{x}_N \exp\left[-\beta H_N^{\text{atom}}\right] \\ &= \sum_{N=0}^{\infty} \frac{e^{2\beta\mu N}}{N!} \left(\frac{\overline{m}}{2\pi\beta\hbar}\right)^{3N/2} |\Omega|^N e^{-\beta N E_0} \\ &= \exp(\rho_a |\Omega|) , \end{aligned} \tag{3.2}$$

with

$$\rho_a = \left(\frac{\overline{m}}{2\pi\beta\hbar}\right)^{3/2} e^{-\beta(E_0 - 2\mu)} . \tag{3.3}$$

Obviously, in this model, the pressure is $P = \beta^{-1}\rho_a$, *i.e.* obeys the law of a perfect gas of atoms with density ρ_a. Note here that we have kept the factor 2 in $\exp(2\beta\mu N)$ by analogy with (2.6) when $N = M$, so that the total number density (electrons and protons) is $\rho = 2\rho_a$ giving as well $P = \beta^{-1}\rho/2$.

To show the existence of an atomic phase, one is therefore led to ask the question: does there exist a density and temperature regime such that the full partition function (2.6) is close to (3.2)? Such a regime can only occur at very low temperatures and densities. If the temperature is not sufficiently low, atoms would have appreciable probabilities to be in excited or even ionized states. If the density is not small, the state of the gas would have no reasons to be perfect. When $\beta \gg 1$, a low density regime is obtained as soon as μ is kept fixed and negative since then all factors $\exp[\beta\mu(N+M)]$ $(N+M \neq 0)$ are small. With these motivations in mind, we can state the thermodynamical version of Fefferman's theorem.

Theorem
Assume that the stability condition (2.15) holds (with $K < |E_0|$). Then, there exists an interval $[\mu_1, \mu_2](E_0 < \mu_1 < \mu_2 < 0)$ such that if
(i) $\mu \in [\mu_1, \mu_2]$
(ii) β and $|\Omega|$ are sufficiently large,
one has

$$Z(\mu, \beta, \Omega) = \exp[\rho_a |\Omega| (1 + O(\beta^{-1}))] . \tag{3.4}$$

The region of the μ–T phase diagram where (3.4) is valid is shown in Fig. 2. Comparing (3.4) with (3.2), one should appreciate that the highly nontrivial part of the theorem lies in the control of the correction $O(\beta^{-1})$ in (3.4). This correction contains all the residual effects due to thermal excitation and interactions between atoms when the temperature and density are not strictly equal to zero.

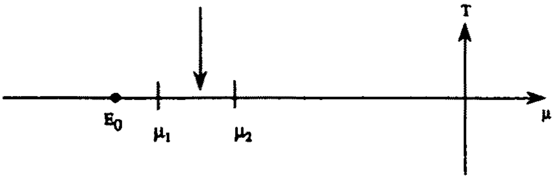

Fig. 2.

A more refined statement on typical configurations of charges can also be found in [1]. It is shown that the distribution (2.5), under the conditions *(i)* and *(ii)* of the theorem, is concentrated on electron-proton pairs having extensions distributed according to the ground state wave function (1.3) of the hydrogen atom.

4. Qualitative ideas for the proof

We describe some of the main ideas of the proof. If these ideas have a simple and intuitive content, needless to say that the full power of the mathematical analysis presented in [1] is necessary to put them at work. The main line of attack is to perform a configurational analysis of the system (instead of trying to solve the complicated eigenvalue equations (2.4)).

Assume that the typical configurations of the charges are indeed those of a dilute gas of hydrogen atoms. Such configurations will be mainly formed of pairs (x_i, y_j), the extension $|x_i - y_j|$ of a pair being of the order of the Bohr radius a_B, whereas two different pairs are at distance $\rho_a^{-1/3}$ $(a_B \ll \rho_a^{-1/3})$. If we set at random in this configuration a ball of radius R with

$$a_B \ll R \ll \rho_a^{-1/3}, \tag{4.1}$$

we will observe that the ball is empty most of the time, but if it contains something, it is exactly one hydrogen atom, except in rare occasions (see Fig. 3). The idea is then to approximately decompose the total value $\Omega = \bigcup_i B_i$ into an union of sub domains (a collection of balls with radii satisfying (4.1)) and to solve the two following problems.

Problem 1

Determine the size of the balls B_i, the temperature T and the chemical potential μ such that the partition function $Z(\mu, \beta, B_i)$ for a single ball is that of a system having at most one hydrogen atom in B_i.

Problem 2

Show that the residual interaction between the different balls B_i is negligible so that by (2.11)

$$Z(\mu, \beta, \Omega) \approx \prod_i Z(\mu, \beta, B_i). \tag{4.2}$$

Thus is we can explicitly evaluate $Z(\mu, \beta, B_i)$ in the simpler situation of problem 1, (4.2) will give access to the full partition function.

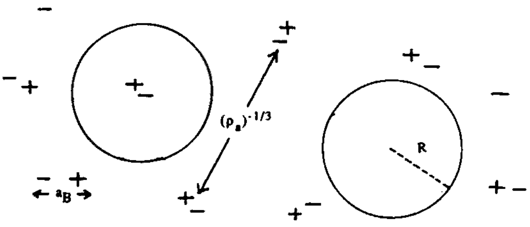

Fig. 3.

Towards a solution of problem 1

According to (2.7), to solve problem 1, one must be able to find μ, β and the radius R of a ball B such that

$$p_{00}(\mu, \beta, B) \gg p_{11}(\mu, \beta, B) \gg p_{NM}(\mu, \beta, B), \tag{4.3}$$

for all (NM) different from $(0,0)$ and $(1,1)$. This means that if B is not void, it is much more probable to find an electron-proton pair in it than any other charge configuration.

Denote $E_{NM} = \inf_{\alpha} E^B_{NM\alpha}$ the ground state energy of the Hamiltonian H^B_{NM} (2.3) for the ball B. Since B will eventually be very large, up to a small correction, we can as well take E_{NM} equal to the infimum of the spectrum of H_{NM} in the infinite space. Then one has $E_{00} = 0$ and $E_{10} = E_{01} = 0$ since the energy of a single charge is purely kinetic, moreover $E_{11} = E_0$ (the ground state energy of the hydrogen atom), and in general $E_{NM} \leq 0$. It is clear that at low temperature and for a large ball, the dominant statistical weight in the sum (2.7) will essentially be $\exp(-\beta E_{NM})$, and therefore,

$$p_{NM}(\mu,\beta,B) \approx \frac{1}{Z(\mu,\beta,B)} \exp[-\beta(E_{NM}-\mu(N+M))]\,, \qquad \beta \gg 1\,. \tag{4.4}$$

Thus to obtain the inequalities (4.3), one must be able to find values of μ such that

$$E_{NM} - \mu(N+M) > E_0 - 2\mu > 0\,, \tag{4.5}$$

for all (N,M) different from $(0,0)$ and $(1,1)$.

Let us examine some implications of (4.5). If $(N,M) = (1,0)$ (a single electron), $E_{10} = 0$ in (4.5) implies

$$E_0 < \mu < 0\,. \tag{4.6}$$

If $(N,M) = (2,2)$ (4.5) gives

$$\mu < \tfrac{1}{2}(E_{22} - E_0)\,, \tag{4.7}$$

where E_{22} is the ground state energy of the hydrogen molecule. One can find μ satisfying (4.6) and (4.7) only if $E_0 < {}^1\!/\!_2(E_{22} - E_0)$ or equivalently if $|E_{22} - 2E_0| < |E_0|$. The latter inequality is known to be true (see the discussion after (2.16)). Thus (4.6) and (4.7) give lower and upper bounds on the possible values of μ, but one must examine all value of (N,M). In fact, one can verify that the possibility to satisfy all the inequalities (4.5) for values of μ in an interval $[\mu_1,\mu_2](E_0 < \mu_1 < \mu_2 < 0)$ is precisely guaranteed by stability of matter bound (2.15). A geometrical interpretation of the inequalities (4.5) is presented in Fig. 4.

Under these conditions, we can now solve problem 1. According to (2.6) and (2.8) we write the grand-partition function as a sum on all particle numbers

$$Z(\mu,\beta,B) = 1 + Z(\mu,\beta,1,1,B) + \sum_{\substack{N,M \\ (N,M)\neq(0,0)\text{ and }(1,1)}} Z(\mu,\beta,N,M,B)\,, \tag{4.8}$$

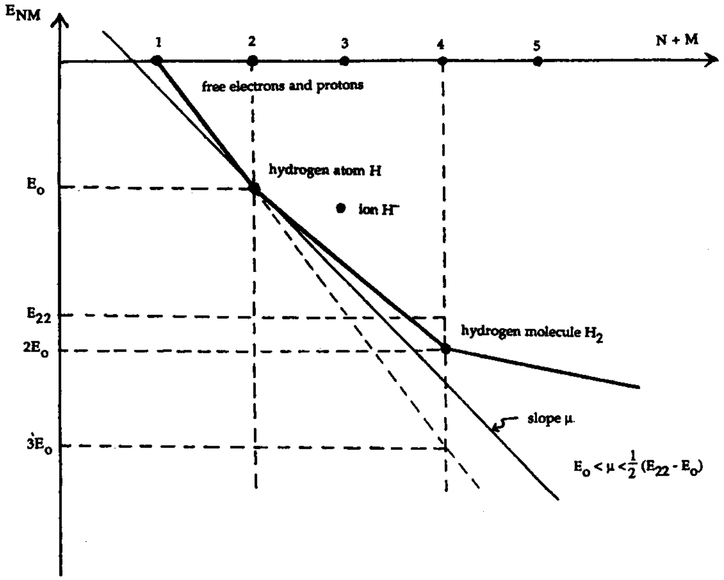

Fig. 4.

where we have singled out the contribution $Z(\mu, \beta, 1, 1, B)$ of a single electron-proton pair. As said before (see (4.1)), in order to see the formation of the atom, the ball B must have an appropriate size: to avoid dissociation, B should not be too big (see the discussion after (1.6)), but nevertheless big enough to leave the atomic wave functions almost unperturbed by finite size and boundary conditions effects. The appropriate choice for R is

$$e^{c_2\beta} \leq R \leq e^{c_1\beta}, \tag{4.9}$$

where $c_1 > c_2 > 0$ are constants (independent of β), and c_1 is sufficiently small. Since the inverse atomic density is expected to grow like $\exp(\beta(E_0 - 2\mu))$ with some μ slightly above E_0 (see (3.3), (4.6) and (4.7)), the condition (4.1) will be satisfied by the choice (4.9).

Then one establishes easily that under the constraint (4.9) and for $\beta \gg 1$ that

$$\begin{aligned} Z(\mu, \beta, 1, 1, B) &= \left(\frac{\overline{m}}{2\pi\beta\hbar^2}\right)^{3/2} e^{-\beta(E_0 - 2\mu)}|B|(1 + O(\beta^{-1})) \\ &= \rho_a|B|(1 + O(\beta^{-1})). \end{aligned} \tag{4.10}$$

Indeed $Z(\mu, \beta, 1, 1, B)$ is the partition function of the two charges problem with Hamiltonian

$$H_{11}^B = \frac{\hbar^2}{2\overline{m}}\Delta_z - \frac{\hbar^2}{2m}\Delta_w - \frac{e^2}{|w|} \tag{4.11}$$

with Dirichlet condition on ∂B. In (4.11), $z, w = x - y$ are the center of mass and relative coordinates, $\overline{m} = m_e + m_p$ and m is the relative mass. The center of mass contribution (kinetic energy) is

$$\mathrm{Tr}_B \exp\left[-\beta\left(-\frac{\hbar^2}{2\overline{m}}\Delta_z\right)\right] = \left(\frac{\overline{m}}{2\pi\beta\hbar^2}\right)^{3/2}|B|(1 + O(R^{-1})). \tag{4.12}$$

The contribution of the hydrogen energy spectrum is for $\beta \gg 1$

$$\mathrm{Tr}_B \exp\left[-\beta\left(-\frac{\hbar^2}{2m}\Delta_w - \frac{e^2}{|w|}\right)\right] = \mathrm{e}^{-\beta E_0}(1 + O(\mathrm{e}^{-c\beta})). \tag{4.13}$$

Since by (4.9) $R^{-1} = O(\mathrm{e}^{-c_2\beta})$, the combination of (4.12) and (4.13) gives (4.10). All the effects due to the excited states and to the finite size of the ball are included in the correction term $O(\beta^{-1})$ in (4.10).

The next step needed to solve problem 1 consists in showing that the last term in (4.8) can be neglected in comparison with the two first ones when μ is chosen in $[\mu_1, \mu_2]$ and $\beta \gg 1$, leading to

$$\begin{aligned} Z(\mu, \beta, B) &= 1 + Z(\mu, \beta, 1, 1, B)\left(1 + O(\beta^{-1})\right) \\ &= 1 + \rho_a|B|\left(1 + O(\beta^{-1})\right) \end{aligned} \tag{4.14}$$

Here the stability bound (2.15) is used in a crucial way. First it insures the existence of an interval $[\mu_1, \mu_2]$ for the chemical potential, as discussed after (4.6) and (4.7), where the hydrogen atom has the largest statistical weight among all (non void) charge configurations. Moreover, it enables to majorize uniformly the terms of the last sum in (4.8) to give (4.14).

Finally, since $\rho_a|B|$ vanishes as $\beta \to \infty$, (choosing c_1 in (4.9) small enough), we can as well write (4.13) in exponential form

$$Z(\mu, \beta, B) = \exp\left[\rho_a|B|\left(1 + O(\beta^{-1})\right)\right] \tag{4.15}$$

This expression is the same as (3.4) with Ω replaced by B. This difference is of importance since B in (4.15) is constrained by the condition (4.9), and is therefore not an independent thermodynamical variable as is the total volume Ω in the grand-canonical ensemble. To obtain the result of the theorem, we have to treat the problem 2.

Towards a solution of problem 2

The problem 2 is a real many body one, of much higher technical difficulty. We shall give a few hints which hardly represent the amount of mathematical work requested to bring them under control.

We decompose the total volume Ω into a disjoint union of balls with radii satisfying the conditions (4.9). For any fixed β, one chooses Ω large enough to contain many balls and one writes $\Omega \sim \bigcup_i B_i$. The symbol $\sim$ means that the covering of Ω is not complete, but that the interstitial region between the balls is sufficiently small to be eventually neglected. Let us assume now that we can switch off the Coulomb interaction between the different balls, introducing an uncorrelated Hamiltonian (relative to the decomposition $\Omega \sim \bigcup_i B_i$) by

$$H_{\text{unc}} = H - U_{\text{cor}} . \tag{4.16}$$

H is the full Hamiltonian (2.3) and U_{cor} is the potential correlation energy between the balls. U_{cor} is constructed in the usual way from the two body potential

$$V_{\text{cor}}(\underline{x}_1, \underline{x}_2) = \sum_{i \neq j} \frac{\chi_{B_i}(\underline{x}_1)\chi_{B_j}(\underline{x}_2)}{|\underline{x}_1 - \underline{x}_2|}, \qquad \chi_B(\underline{x}) = \begin{cases} 1, & \underline{x} \in B \\ 0, & \underline{x} \notin B \end{cases} . \tag{4.17}$$

$V_{\text{cor}}(\underline{x}_1, \underline{x}_2)$ is different from zero only if $\underline{x}_1$ and $\underline{x}_2$ belong to different balls. By construction, the particles located in different balls do not interact in H_{unc}, so the factorization property (2.11) will apply to the grand-canonical partition function $Z_{\text{unc}}(\mu, \beta, \Omega)$ associated with H_{unc}, as well as the result (4.15) of problem 1 for each ball. This leads to (see however the remarks following (4.26))

$$\begin{aligned} Z_{\text{unc}}(\mu, \beta, \Omega) &\sim \prod_i Z(\mu, \beta, B_i) \\ &\sim \exp\Big[\rho_{\text{a}} \sum_i |B_i| \left(1 + O(\beta^{-1})\right)\Big] \\ &\sim \exp\left[\rho_{\text{a}} |\Omega| \left(1 + O(\beta^{-1})\right)\right] . \end{aligned} \tag{4.18}$$

The goal is to prove that the full partition function $Z(\mu, \beta, \Omega)$ has the same behavior (4.18) as β becomes large. In other words, the effective correlation energy U_{cor} (which is temperature dependent through the condition (4.9) on the radii of the balls) can be neglected when $\beta \gg 1$. For this, we note that $V_{\text{cor}}(\underline{x}_1, \underline{x}_2)$ is regular at short distances (in fact it vanishes at $\underline{x}_1 = \underline{x}_2$ since the balls are disjoint), and coulombic at large distances. This behavior is similar to that of a Coulomb potential regularized by extended charges, for instance as

$$V_{\text{reg}}(\underline{x}_1, \underline{x}_2) = \int d\underline{y}_1 \int d\underline{y}_2 \frac{\varphi_R(\underline{y}_1)\varphi_R(\underline{y}_2)}{|(\underline{x}_1 - \underline{y}_1) - (\underline{x}_2 - \underline{y}_2)|}, \tag{4.19}$$

where $\varphi_R(\underline{y})$ represents a smooth charge density supported in B with total charge equal to 1

$$\varphi_R(\underline{y}) = 0\,, \; |y| \geq R\,, \qquad \int\limits_{|\underline{y}|\leq R} d\underline{y}\, \varphi_R(\underline{y}) = 1\,. \tag{4.20}$$

The potential energy of $N+M$ charges $e_i = \pm e$ interacting by the regularized potential (4.19) satisfies

$$\begin{aligned} U_{\mathrm{reg}} &= \sum_{i<j} e_i e_j V_{\mathrm{reg}}(\underline{x}_i - \underline{x}_j) \\ &= \tfrac{1}{2} \sum_{i,j} e_i e_j V_{\mathrm{reg}}(\underline{x}_i - \underline{x}_j) - (N+M)e^2 V_{\mathrm{reg}}(0) \\ &\geq -(N+M)e^2 V_{\mathrm{reg}}(0)\,. \end{aligned} \tag{4.21}$$

The inequality follows from the fact that the potential energy including the self energies is positive, as seen from the Fourier transform representation

$$\sum_{i,j} e_i e_j V_{\mathrm{reg}}(\underline{x}_i - \underline{x}_j) = \frac{1}{2\pi^2} \int d\underline{k} \frac{1}{|\underline{k}|^2} \Big| \sum_j e_j \exp(i\underline{k}\cdot\underline{x}_j)\tilde{\varphi}_R(\underline{k}) \Big|^2 > 0\,. \tag{4.22}$$

Now, because of the normalization (4.20), one has $\varphi_R(\underline{y}) = O(R^{-3})$ implying for the self energy

$$V_{\mathrm{reg}}(0) = \int\limits_{|\underline{y}|\leq R} dy_1 \int\limits_{|\underline{y}|\leq R} dy_2 \frac{\varphi_R(\underline{y}_1)\varphi_R(\underline{y}_2)}{|\underline{y}_1 - \underline{y}_2|} = O\left(\frac{1}{R}\right) \leq O\left(\mathrm{e}^{-c_2\beta}\right)\,, \tag{4.23}$$

where the last inequality follows from (4.9). Taking (4.23) into account in (4.21) gives

$$U_{\mathrm{reg}} \geq -O\left(\mathrm{e}^{-c_2\beta}\right)(N+M)\,. \tag{4.24}$$

If we were allowed to replace U_{cor} by U_{reg} in (4.16), we would obtain

$$\begin{aligned} H - \mu(N+M) &= H_{\mathrm{unc}} + U_{\mathrm{reg}} - \mu(N+M) \\ &\geq H_{\mathrm{unc}} - \bar{\mu}(N+M) \end{aligned} \tag{4.25}$$

with $\bar{\mu} = \mu + O(\mathrm{e}^{-c_2\beta})$. Up to an exponentially small correction to the chemical potential, H_{unc} would provide a lower bound to the full Hamiltonian, and hence, by (2.6) Z_{unc} provides an upper bound to the full partition function

$$\begin{aligned} Z(\mu,\beta,\Omega) &\leq Z_{\mathrm{unc}}(\bar{\mu},\beta,\Omega) \\ &= \exp\left[\rho_{\mathrm{a}}|\Omega|\left(1 + O(\beta^{-1})\right)\right]\,. \end{aligned} \tag{4.26}$$

Here, ρ_a is again given by (3.3), the difference between μ and $\bar{\mu}$ being absorbed in the $O(\beta^{-1})$ correction.

Unfortunately, things are much more complex. First, the difference $U_{\text{error}} = U_{\text{cor}} - U_{\text{reg}}$ has to be taken into account. Moreover, since the covering of Ω by the balls is not complete, one must deal with the interstitial region. Finally the definition (4.16) of H_{unc} is not yet suitable to allow the use of the factorization property (2.11) and the result (4.15) of the problem 1. For this, it is necessary that the Laplacians in H_{unc} have Dirichlet conditions on the boundary of each ball B_i leading also to a perturbation of the kinetic energy. These difficulties are the source of long technical developments that we cannot report here. For instance, in order to control U_{error}, Fefferman averages on a family of coverings of Ω with balls having different radii. To avoid that Dirichlet conditions on the balls cause a to high cost in kinetic energy, one must smooth off the characteristic function in (4.17) and introduce appropriate decompositions of the unity into sums of smooth functions. When the full analysis is eventually performed, the upper bound (4.26) is valid.

The conclusion (3.4) of the theorem follows from a lower bound to $Z(\mu, \beta, \Omega)$ given by the inequality used by Lieb and Lebowitz in their proof of the existence of the thermodynamic limit [6]

$$Z(\mu, \beta, \Omega) \geq \prod_i Z_{\text{neutral}}(\mu, \beta, B_i)\,. \tag{4.27}$$

$Z_{\text{neutral}}(\mu, \beta, B_i)$ is defined as in (2.6), but with the summations restricted to the neutral terms $N = M$. The results of the problem 1 apply as well as to $Z_{\text{neutral}}(\mu, \beta, B_i)$ so that

$$Z(\mu, \beta, \Omega) \geq \exp\left[\rho_a |\Omega| (1 + O(\beta^{-1}))\right] \tag{4.28}$$

The conjunction of the inequalities (4.26) and (4.28) proves the assertion of the theorem.

5. Extensions and perspectives

5.1. *Ionization equilibrium*

There exists an interesting situation, called ionization equilibrium, where a macroscopic fraction of electrons does not bind and remains in thermal equilibrium with the atoms. In the traditional thermodynamical treatment of this situation [7], one considers the unbound electrons (e), the protons (p) and the hydrogen atoms (a) as different chemical species with number densities ρ_e, ρ_p and ρ_a ($\rho_e = \rho_p$ by neutrality). Then the ionization equilibrium phases are characterized by their degree of ionization (the Saha coefficient) defined by

$$\alpha = \frac{\rho_e}{\rho_e + \rho_a} . \tag{5.1}$$

Assuming that each species behaves as a perfect gas, one has the simple equation of state

$$\beta P = \rho_e + \rho_p + \rho_a = \left(\frac{1+\alpha}{2}\right)\rho , \tag{5.2}$$

where $\rho = \rho_e + \rho_p + 2\rho_a$ is the total number density.
Applying the law of mass action to the reaction e + p ↔ a, one obtains the Saha equation for thermal ionization expressing α in terms of the pressure and the temperature

$$\alpha(P,T) = \left[1 + \beta P \left(\frac{2\pi\beta\hbar^2}{m}\right)^{3/2} \mathrm{e}^{-\beta E_0}\right]^{-1/2} . \tag{5.3}$$

We now come back to our fundamental system of protons and electrons, without introducing preformed atoms. We know from Fefferman's theorem that for β large and μ slightly above E_0, the system is a free gas of hydrogen atoms. For μ below E_0 and β large, the system is a fully ionized plasma. The ionization equilibrium phases interpolate between these two situations. More precisely, if the chemical potential approaches $-E_0$ linearly with T in the μ–T phase diagram (see Fig. 4), *i.e.*

$$\mu(\beta) = -E_0 + \sigma\beta^{-1} + O(\beta^{-1}) , \qquad -\infty < \sigma < \infty \tag{5.4}$$

one shows that Eqs (5.2) and (5.3) hold asymptotically [8]. In particular

$$\lim_{\beta\to\infty} \frac{\beta P(\mu(\beta),\beta)}{\rho(\mu(\beta),\beta)} = \frac{1+\alpha}{2} \tag{5.5}$$

and the Saha coefficient is related to the slope σ in (5.4) by

$$\alpha = \left[\frac{(m_e m_p)^{3/4}}{(m_e + m_p)^{3/2}} \mathrm{e}^{\sigma} + 1\right]^{-1} . \tag{5.6}$$

The cases of full binding ($\alpha = 0$) and full dissociation ($\alpha = 1$) correspond to $\sigma = \infty$ and $\sigma = -\infty$ respectively.

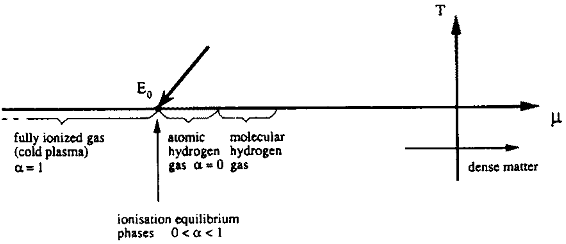

Fig. 5.

5.2. Nuclei and electrons

Conlon *et al.* [9] have generalized the low temperature and low density analysis to the situation where a number of nuclei of arbitrary charges are present. Consider that we have now s types of particles with charges $e_1, \ldots, e_s$ and masses $m_1, \ldots, m_s$ (including the electrons). We denote the corresponding charge, chemical potential and particle number vectors by

$$\begin{aligned} \underline{Q} &= (e_1, \ldots, e_s)\,, \\ \underline{\mu} &= (\mu_1, \ldots, \mu_s)\,, \\ \underline{N} &= (N_1, \ldots, N_s)\,, \end{aligned} \tag{5.7}$$

and let $H_{\underline{N}}$ be the total coulombic Hamiltonian for the $N_1 + \cdots + N_s$ particles. In the thermodynamic limit the density $\rho(\underline{\mu}, \beta)$ and the pressure $\underline{P}(\underline{\mu}, \beta)$ depend only on the component of $\underline{\mu}$ orthogonal to the charge vector (see the comment following (2.17)).

We now fix chemical potentials $\underline{\mu}$ such that $\underline{u} \cdot \underline{Q} = 0$. The dominant weight in the grand-canonical sum $\exp(-\beta E_{\underline{\mu}})$ will correspond to the lowest eigenvalue of

$$\left(H_{\underline{N}}^{\mathrm{rel}} - \underline{\mu} \cdot \underline{N}\right) \phi = E_{\underline{\mu}} \phi\,, \tag{5.8}$$

when we let $\underline{N}$ vary over all particle numbers of the s species. In (5.8) $H_{\underline{N}}^{\mathrm{rel}}$ is the coulomb Hamiltonian H_N with center of mass removed. We may find one, or possibly a finite number of degenerate ground states $\phi_1, \ldots, \phi_g$ satisfying (5.8). These ground states describe g complexes (atoms, ions molecules), each of them constituted of $\underline{N}^{(i)}$ particles, having masses and charges

$$M^{(i)} = \sum_{r=1}^{s} m_r N_r^{(i)}\,, \qquad \underline{Q}^{(i)} = \sum_{r=1}^{s} e_r N_r^{(i)}\,, \tag{5.9}$$

and ground states energies $E^{(i)}$ such that

$$E_{\underline{\mu}} = E^{(i)} - \underline{\mu} \cdot \underline{N}^{(i)}\,, \qquad i = 1, \ldots, g\,. \tag{5.10}$$

The theorem of Conlon *et al.* states that as $\beta \to \infty$, the system behaves like a mixture of perfect gases of g chemical complexes determined by (5.8), with charges and masses (5.9)[6]. Thus, in the region of the $\underline{\mu}$–T phase diagram corresponding to high dilution and low temperature, we find that the thermodynamical laws of mixtures of non interacting chemical species are valid. What kind of chemical species will form for a given $\underline{\mu}$ can of course only be determined by solving (5.8), *i.e.* by a precise numerical knowledge of the possible binding energies.

5.3. Beyond the vanishing temperature and vanishing density limit

The theory presented in this lecture provides a rigorous analysis of the asymptotic behaviour of the quantum Coulomb system as $T \to 0$ and $\rho \to 0$. This so called Saha regime is dominated by the energy-entropy balance. If μ is fixed and very negative, the entropy wins: the state of the system is fully dissociated because of the fast lowering of the density as $\beta \to \infty$. A progressive increase of μ favors higher densities, and the formation of atoms and chemical complexes becomes possible.

The question arises: what is the structure of the first corrections to the asymptotic limit (hidden here in the $O(\beta^{-1})$ estimates of Fefferman). A detailed understanding of these corrections involves conceptual problems (the notion of bound and excited states in a medium with non zero density) as well as practical ones such as the calculation of the next terms of the low density expansion of the equation of state. This leads us into the theory of non ideal plasmas which is of course the subject of an immense physical literature that we cannot quote here, but exact results are still scarce. In this direction, we can mention one attempt to formulate the notion of bound states in the whole range of densities in terms of the spectrum of the reduced density matrices [10], and a study of the low density equation of state [11]. Many basic questions pertaining to the quantum Coulomb system certainly deserve still work and clarification.

REFERENCES

[1] Ch. Fefferman, *Rev. Math. Iberoamericana* **1**, 1 (1985).
[2] Ch. Fefferman, *Comm. on Pure and Applied Mathematics* 67 (1986).
[3] R. Peierls, *Surprises in Theoretical Physics*, Princeton University Press 1979.
[4] F.J. Dyson, A. Lenard, *J. Math. Phys.* **8**, 423 (1967).
[5] E. Lieb, *Rev. Mod. Phys.* **48**, 553 (1976).
[6] E. Lieb, J.L. Lebowitz, *Adv. Math.* **9**, 316 (1972).

---

[6] Also the estimates $O(\beta^{-1})$ are improved to exponential ones $O(e^{-c\beta})$.

[7] L. Landau, E. Lifshitz, *Statistical Physics*, Pergamon Press 1976.
[8] N. Macris, Ph.A. Martin, *J. Stat. Phys.* **60**, 619 (1990).
[9] J. Conlon, E. Lieb, H. Yau, *Comm. Math. Phys.* **125**, 153 (1989).
[10] J. Lebowitz, N. Macris, Ph.A. Martin, *J. Stat. Phys.* **67**, 909 (1992).
[11] A. Alastuey, A. Perez, *Europhys. Letters* **20**, 19 (1992).