A Brief Introduction to Thomas-Fermi Model in Partial Differential Equations

Aditya Kumar
Department of Mathematics and Statistics
McGill University, Montreal, QC

December 16, 2012

1 Introduction

Created independently by Llewellyn H. Thomas and Enrico Fermi around 1926, the Thomas-Fermi model is a quantum mechanical theory for the electronic structure of a many-body system. This statistical model was developed separately from wave function theory by being formulated in terms of electron density. The idea of the model is that given a large atom, with many electrons, one can approximately model it by a simple nonlinear problem for a specified charge density. In other words, this statistical model can be used to approximate the distribution of electrons in an atom. In a mathematical framework, one can take the qualitative and quantitative physical assumptions imposed by the Thomas-Fermi model and solve the model through the Direct Method of calculus of variations. This then allows strong mathematical rigor to be applied to the formulation of the model, utilizing functional analysis. The Thomas-Fermi model is defined by the energy functional for the ground state energy of the system for a certain amount of electrons in the atom with a particular charge. In this report, I will prove the uniqueness and existence of a minimizer by filling in some gaps in the proof laid out partially in the textbook *Analysis* by Lieb and Loss. I will then discuss certain issues about the approximation results gained from using this model and how the model has been improved upon. The mathematical rigor applied to the TF model was conducted by research in the 1970s by Elliot Lieb and Barry Simon.

2 Statement of the Thomas-Fermi Model as a Variational Problem

The Thomas-Fermi (TF) theory is defined by an energy functional $\mathcal{E}$ on a certain class of *non-negative* functions ρ on $\mathbb{R}^3$:

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

where Z > 0 is a fixed parameter (physically interpreted as the charge of the atom's nucleus) and

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

is the Coulomb energy of a charge density. This is given by

$$D(f,g) := \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \bar{f}(x)g(y)|x-y|^{2-n}dxdy$$

which we define for $\mathbb{R}^3$, $n \geq 3$, and define for complex-valued functions f & g in $L^1_{loc}(\mathbb{R}^n)$. For the physical interpretation, for n=3, $D(f,f)$ is the true physical energy of a real charge density f. It is the energy needed to assemble f from 'infinitesimal' charges. We now define the class of admissible functions $\mathcal{C}$ for the charge density ρ. The class of admissible functions is

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

Before moving on, I will briefly explain on how to show that $\mathcal{E}(\rho)$ is well defined and finite when ρ is in the class $\mathcal{C}$. This is necessary as this shows that the variational problem is well-posed, to an extent, in terms of the mathematical rigor that we are constructing for the TF model.

In order to show that $\mathcal{E}(\rho)$ is well defined and finite, we have to show each term is well defined and finite. Thus

- $\frac{3}{5}\int_{\mathbb{R}^3} \rho(x)^{\frac{5}{3}} dx$ is, by definition of ρ in the class of admissible functions $\mathcal{C}$, finite. It is well defined because we are minimizing over $L^{\frac{5}{3}}$.
- Split up the integral $\int_{\mathbb{R}^3} \frac{Z}{|x|}\rho(x)dx$ by taking a ball of radius ϵ, $B(0,\epsilon)$, and use Hölder's inequality to bound (to recall Hölder's see Section 3).
- We see that by how $D(\rho,\rho)$ was defined above, it is evident that $D(\rho,\rho)$ is in $L^{\frac{6}{5}}$ and so, we can use Hardy-Littlewood-Sobolev (HLS) inequality to bound (to recall HLS

see Section 3).

For the TF variational problem, we need to minimize $\mathcal{E}(\rho)$ under the condition that $\int \rho = N$, where N is a fixed *positive* number (identified as the "number" of electrons in the atom). We proceed by defining two subsets of $\mathcal{C}$:

$$\mathcal{C}_N := \mathcal{C} \cap \{\rho : \int_{\mathbb{R}^3} \rho = N\} \subset \mathcal{C}_{\leq N} := \mathcal{C} \cap \{\rho : \int_{\mathbb{R}^3} \rho \leq N\}$$

There are two energies corresponding to these two sets: The 'constrained' energy, which is given by

$$E(N) = \inf\{\mathcal{E}(\rho) : \rho \in \mathcal{C}_N\}$$

& the 'unconstrained' energy, which is given by

$$E_{\leq}(N) = \inf\{\mathcal{E}(\rho) : \rho \in \mathcal{C}_{\leq N}\}$$

From this, it is obvious that $E_{\leq}(N) \leq E(N)$. Introducing the unconstrained problem will become evident as we construct the proof for existence portion of the TF minimizer. We note that a minimizer will not exist for the constrained problem when $N > Z$ due to the fact that atoms cannot be negatively charged in TF theory. However, a minimizer will always exist for the unconstrained problem.

3 Useful Inequalities

In this section I will give, without proof, three important inequalities that are very useful when constructing the necessary mathematical rigor for the TF problem. The proofs can be found in any real analysis textbook (see [1]).

Let $L^p(\mathbb{R}^3)$, $1 \leq p \leq \infty$. If $p < \infty$, $L^p(\mathbb{R}^3)$ is a set of measurable functions from $\mathbb{R}^3$ to $\mathbb{C}$ with the property

$$\|f\|_p \equiv \left(\int |f|^p dx \right)^{\frac{1}{p}} < \infty$$

Thus, we have the following useful relations between L^p spaces:

Hölder's Inequality:

$$\|fg\|_r \leq \|f\|_p \|g\|_q$$

if

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} \quad ; \quad 1 \leq p, q, r \leq \infty$$

Young's Inequality:

$$\|f * g\|_r \leq \|f\|_p \|g\|_q$$

if

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q} \quad ; \quad 1 \leq p, q, r \leq \infty$$

Hardy-Littlewood-Sobolev (HLS) Inequality:

Let $p, r > 1$ and $0 < \lambda < n$ with $\frac{1}{p} + \frac{\lambda}{n} + \frac{1}{r} = 2$. Let $f \in L^p(\mathbb{R}^n)$ and $h \in L^r(\mathbb{R}^n)$. Then there exists a sharp constant $C(n, \lambda, p)$, independent of h & f, such that

$$\left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^3} f(x)|x-y|^{-\lambda} h(y) dx dy \right| \leq C(n, \lambda, p) \|f\|_p \|h\|_r$$

These are the three main inequalities that will be used.

4 Uniqueness and Existence of a Minimizer for the Thomas-Fermi Problem

Theorem: Existence of Unconstrained Thomas-Fermi Minimizer
For each $N > 0$ there is a unique minimizing ρ_N for the unconstrained TF problem, i.e. $\mathcal{E}(\rho_N) = E_{\leq}(N)$. The constrained energy $E(N)$ and constrained energy are equal.
Moreover $E(N)$ is convex and non-increasing function of N

We note that the last sentence of the theorem holds because TF problem is defined on all of $\mathbb{R}^3$. If we were to replace $\mathbb{R}^3$ with a bounded subset of $\mathbb{R}^3$, then $E(N)$ would not be a non-increasing function.

PROOF.
First we have to show that $\mathcal{E}(\rho)$ is bounded below on the set $\mathcal{C}_{\leq N}$ so that $E_{\leq}(N) > -\infty$.

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

the $D(\rho, \rho)$ term is a linear functional in $L^{\frac{6}{5}}$, so one can use HLS inequality to see that it is ≥ 0. We take the second term in $\mathcal{E}(\rho)$ and split the integral up and use Hölder's inequality (note: we will deal with the negative sign in front of it after)

$$\int_{\mathbb{R}^3} \frac{Z}{|x|}\rho(x)dx = \int_{\mathbb{R}^3/B(0,\epsilon)} \frac{Z}{|x|}\rho(x)dx + \int_{B(0,\epsilon)} \frac{Z}{|x|}\rho(x)dx$$
$$\leq \frac{1}{\epsilon}\int_{\mathbb{R}^3/B(0,\epsilon)} \rho(x)dx + \|\frac{Z}{|x|}\|_{L^{\frac{5}{2}}(B(0,\epsilon))}\|\rho(x)\|_{L^{\frac{5}{3}}(\mathbb{R}^3)}$$

Then, we note that $\int_{\mathbb{R}^3}\rho \leq N$ from the definition of $\mathcal{C}_{\leq N}$. Hence

$$\mathcal{E}(\rho) \geq \int_{\mathbb{R}^3} \rho(x)^{\frac{5}{3}}dx - \frac{1}{\epsilon}N - \|\frac{Z}{|x|}\|_{L^{\frac{5}{2}}(B(0,\epsilon))}\|\rho(x)\|_{L^{\frac{5}{3}}(\mathbb{R}^3)}$$
$$\geq \frac{1}{2}\|\rho(x)\|_{L^{\frac{5}{3}}(\mathbb{R}^3)} - \frac{1}{\epsilon}N$$

and so,

$$\epsilon\frac{1}{N} + \mathcal{E}(\rho) \geq \frac{1}{2}\|\rho(x)\|_{L^{\frac{5}{3}}\mathbb{R}^3} \implies E_{\leq N} > -\infty$$

Now that we have that $\mathcal{E}(\rho)$ is bounded from below for $\mathcal{C}_{\leq N}$, we are going to show there exists a candidate minimizer, ρ_N, by constructing a weakly convergent minimizing sequence to ρ_N and by showing that all the terms in $\mathcal{E}(\rho)$ are weakly lower semicontinuous (w.l.s.c).

Let ρ_j be a minimizing sequence in $\mathcal{C}_{\leq N}$. Since $\mathcal{E}(\rho)$ is bounded below then

$$\int \rho_j(x)^{\frac{5}{3}} \leq C$$

for some constant C independent of j. Then, there exists a ρ_N, a candidate minimizer, such that

$$\rho_j \rightharpoonup \rho_N$$

meaning that if ρ_j is a minimizing sequence then it converges weakly in $L^{\frac{5}{3}}$ to ρ_N (since bounded sequences have weak limits[1]). By the weak lower semi-continuity of the norm[2]

$$\liminf \int \rho_j(x)^{\frac{5}{3}}dx \ \geq \ \int \rho_N(x)^{\frac{5}{3}}dx$$

From this, we note that $\rho_N \in \mathcal{C}_{\leq N}$, i.e. $\rho_N(x) \geq 0$ and $\int \rho_N(x)dx \leq N$. To show that $\rho_N(x) \geq 0$, we note that for any positive function $f \in L^{\frac{5}{2}}$

$$\int \rho_N(x)f(x)dx = \lim \int \rho_j(x)dx \geq 0 \implies \rho_N(x) \geq 0$$

[1]Refer to Theorem 2.18 in [1]

[2]Refer to Theorem 2.11 (Lower Semi-Continuity of Norms) in [1].

To show that $\int \rho_N(x)dx \leq N$ we shall assume the contrary. Assume that $\int_0 \rho_N(x)dx > N$. Then there exists a set A of finite measure so that

$$\int \rho_N(x)\chi_A(x)dx > N$$

Here, χ_A is the characteristic function of the set A. Because A has finite measure, and $\chi_A \in L^{\frac{5}{3}}$ thus we get

$$\int \rho_N(x)\chi_A(x)dx = \lim \int \rho_j(x)\chi_A(x)dx \leq N; \text{ a contradiction } \implies \int \rho_N(x) \leq N$$

Since we had shown that every term in $\mathcal{E}(\rho)$ is bounded from below, we can then note that $D(\rho_j, \rho_j)$ is also a bounded sequence. This means that we have

$$D(\rho_j, \rho) \to D(\rho_N, \rho)$$

for any $\rho \in L^{\frac{5}{3}} \cap L^1$. To show this weakly converges, we can construct

$$\frac{1}{|x|} * \rho(x) = \int_{|x-y|<1} \frac{1}{|x-y|}\rho(y)dy + \int_{|y|\geq 1} \frac{1}{|y|}\rho(x-y)dy = f_1 + f_2$$

By Young's Inequality

$$\|f_1\|_{L^\infty} \leq \|\frac{1}{|x|}\chi_{|x|<1}\|_{L^{\frac{5}{2}}}\|\rho\|_{L^{\frac{5}{3}}}$$

and

$$\|f_1\|_{L^1} \leq \|\frac{1}{|x|}\chi_{|x|<1}\|_{L^1}\|\rho\|_{L^1}$$

We also have

$$\|f_2\|_q \leq \|\frac{1}{|x|}\chi_{|x|>1}\|_{L^q}\|\rho\|_{L^1}$$

for all $q > 3$. Hence $f_1 + f_2 \in L^q$ for all $q > 3$. The dual of this space is $L^{q'}$ with $q' < \frac{3}{2}$ and since $\rho_j \in L^1 \cap L^{\frac{5}{3}}$ we can assume that $\rho_j \rightharpoonup \rho_N$ weakly in $L^{q'}$ for some $q' < \frac{3}{2}$, thus proving $D(\rho_j, \rho) \to D(\rho_N, \rho)$.

From this, we get that

$$D(\rho_N, \rho_N) = \lim_{j\to\infty} D(\rho_j, \rho_N) \leq \lim\,\inf D(\rho_j, \rho_j)^{\frac{1}{2}} D(\rho_N, \rho_N)^{\frac{1}{2}} \implies \lim\,\inf D(\rho_j, \rho_j) \geq D(\rho_N, \rho_N)$$

Thus this term in $\mathcal{E}(\rho)$ is w.l.s.c.

The potential term is weakly continuous. Write it as

$$\int V(x)\rho_j(x)dx = \int V_<(x)\rho_j(x)dx + \int V_>(x)\rho_j(x)dx$$

where the integral is split into two parts y taking a ball of radius ϵ, $B(0,\epsilon)$, so that the first integral is over $B(0,\epsilon)$ and the second integral is over $\mathbb{R}^3/\ B(0,\epsilon)$. Then this allows us to see that $V_< \in L^{\frac{5}{2}}$ and $V_> \in L^q$ for all $q > 3$. Since ρ_j converges weakly to ρ_N in $L^{\frac{5}{3}}$ we see that the first integral term converges to

$$\int V_<(x)\rho_N(x)dx$$

and, thusly, since ρ_j converges weakly to ρ_N in $L^{q'}$ for some $q' > \frac{3}{2}$ we get that

$$\lim_{j\to\infty}\int V(x)\rho_j(x)dx = \int V(x)\rho_N(x)dx$$

Thus the existence of a minimizer follows from showing that each term in the $\mathcal{E}(\rho_N)$ functional is w.l.s.c.

We will then have that ρ_N is a minimizer because

$$E_{\leq}(N) = \lim_{j\to\infty} \geq \mathcal{E}(\rho_N) \geq E_{\leq}(N)$$

To prove that ρ_N is the unique minimizer, we note that the functional $\mathcal{E}(\rho)$ is strictly convex functional of ρ on the convex sex $\mathcal{C}_{\leq N}$. If there were two different minimizers, say ρ_1 and ρ_2 in $\mathcal{C}_{\leq N}$, then $\rho = \frac{(\rho_1+\rho_2)}{2}$, which is also in $\mathcal{C}_{\leq N}$. We then say that this particular ρ has strictly lower energy (by construction obviously) than $E_{\leq}(N)$, which is a contradiction. With this same reasoning and with the note that was given directly after the theorem, it is obvious that $E_{\leq}(N)$ is a convex function. $E_{\leq}(N)$ is non-increasing due to a simple consequence of its definition noted in Section 2.
To show equality of the constrained and unconstrained energy definitions, it is perfectly laid out in Page 285, Section 11.12 of [1].

Q.E.D

5 Approximation Problems with The Thomas-Fermi Theory and Improvements

The existence of a unique minimizer ρ_N for $\mathcal{E}(\rho)$ energy functional solves the Thomas-Fermi equation written as (written directly from [1]

$$\rho_N(x) = \frac{Z}{|x|} - [\frac{1}{|x|} * \rho_N](x) - \mu \mathrm{if} \rho_N(x) > 0$$

$$0 \geq \frac{Z}{|x|} - [\frac{1}{|x|} * \rho_N](x) - \mu \mathrm{if} \rho_N(x) = 0$$

where $\mu \geq 0$ is a some constant that depends on the parameter N.

This model was an incredible step forward in developing an approximation method to assist in research and development in chemical physics and atomic physics. As noted in [2],[3], it was an extremely important first step in figuring out a method that would avoid explicitly solving the computationally intense and complexly difficult Schrödinger's wave equation for atoms with more than two electrons. The issue with the Thomas-Fermi equation is that the accuracy is limited because the resulting expression for the kinetic energy is only approximate, and because the method does not attempt to represent the exchange energy of an atom as a conclusion of the Pauli principle. But, from a conceptual standpoint, it gave rise to a field of research called density functional theory (DFT)-a way for using quantum mechanical modelling to investigate electronic structure in the many-body systems.

Immediately after the model was constructed, a term for the exchange energy was added by Paul A.M. Dirac in 1928. This upgraded model also had its issues though. The Thomas-Fermi-Dirac (TFD) equation was also quite inaccurate (though quite mathematically rigorous) for most scientific applications. The largest source of error for the TFD theory was in the representation of the kinetic energy, followed by the errors in the exchange energy, and due to the complete neglect of electron correlation. Then in 1935, Carl Friedrich von Weizsäcker added a correction to the kinetic energy term in the TF theory, which can then make a much improved TFDW theory for modeling large atoms and basic atomic interactions. According to Lieb in [2], this theory is quite difficult to show certain properties are mathematically rigorous, i.e. existence of minimizer, and also showing that there exists no unique minimizer is quite difficult. There is also an issue of boundedness of the energy functional for the TFDW model. But, the biggest appreciation is that it helped in generating decent approximations at the time it was constructed.

In current research, there are two big theories in DFT-namely Hatree-Fock method, and the Kohn-Sham theory- and many other particular models that stem from these two theories and from TFDW theory, or even a combination of the two, in terms of application development. Currently, DFT is being studied by mathematical physics as well in terms of justifying certain mathematical structure used in current Density Functional Theories and making sure the formulations are mathematically rigorous. To find more information about these current research in DFT via chemical, physical, or mathematical standpoint, I direct the reader to [2].

6 Conclusion

The Thomas-Fermi theory is a great feat of science. In terms of mathematics, the model is quite mathematical rigorous and justified, and now serves as a key example in introducing mathematical physics researchers to the field of density functional theory. It is also a great example in understanding an the Direct Method of calculus of variations in terms of a mathematical physics perspective. Currently, from an applicative viewpoint, the TF theory is used to draw qualitative trends about the analytical behavior of atoms and molecules, as it is quite fast computationally. As noted above, it sparked an entire field of research from a chemical, physical, and mathematical perspective, by giving rise to DFT. TF theory stands as a great example of the complexity of interelations between chemistry, physics, and mathematical analysis.

References

[1] Lieb, E.H. and Loss, M. Analysis, American Mathematical Society (2001).

[2] Lieb, E.H.: Thomas-Fermi and related theories of atoms and molecules, Rev. Mod. Phys. 53 (1981), 603-641. Errata 54 (1982), 311.

[3] Lieb, E.H. and Simon, B.: The Thomas-Fermi theory of atoms, molecules and solids, Adv. in Math 23 (1977), 22-116.