

Entropic Density Functional Theory

Ahmad Yousefi *  and Ariel Caticha * 

Department of Physics, University at Albany, Albany, NY 12222, USA

* Correspondence: alefyabox@gmail.com (A.Y.); acaticha@albany.edu (A.C.)

Abstract: A formulation of density functional theory (DFT) is constructed as an application of the method of maximum entropy for an inhomogeneous fluid in thermal equilibrium. The use of entropy as a systematic method to generate optimal approximations is extended from the classical to the quantum domain. This process introduces a family of trial density operators that are parameterized by the particle density. The optimal density operator is that which maximizes the quantum entropy relative to the exact canonical density operator. This approach reproduces the variational principle of DFT and allows a simple proof of the Hohenberg–Kohn theorem at finite temperature. Finally, as an illustration, we discuss the Kohn–Sham approximation scheme at finite temperature.

Keywords: density functional theory; Hohenberg–Kohn theorem; entropic inference; method of maximum entropy; inhomogeneous fluids

1. Introduction

Density functional theory (DFT) is one of the most widely used methods for calculations of the structure of inhomogeneous many-body systems including atoms, molecules, liquids, solids, and surfaces [1,2] (for a pedagogic introduction, see [3]). The theory, which finds its earliest roots in the Thomas–Fermi–Dirac model, was first introduced in its modern form by Hohenberg and Kohn, who showed that the ground state of an electron gas in an external potential can be uniquely characterized by the electron density [4], and by Kohn and Sham, who showed how to include the effects of exchange and correlations [5]. The implications of these ideas were soon extended to finite temperatures in the context of the grand canonical framework [6] (see also [7–9] and references therein) and to the statistical mechanics of non-uniform classical fluids, such as the liquid–vapor interface [10–13] (see also [14] for more references).

In previous work, we derived the classical DFT as an application of the method of maximum entropy [15]. A central concept is the use of entropy itself as a tool to generate optimal approximations to probability distributions [16] in terms of those variables that capture the relevant physical information, namely the particle density $n(x)$. We showed that the entropic DFT (eDFT) approach directly leads to Evans’ variational principle of the classical DFT [11].

In this paper, we are concerned with the very foundations of the DFT framework and our main goal is to extend the entropic DFT (eDFT) formalism to the quantum domain. We emphasize that our goal is neither to derive an alternative to DFT nor to develop improvements to the approximations that are inevitably necessary to the successful implementation of DFT in practical applications. Although information theory has been used to quantify chemical concepts [17,18], alternative information-theoretic interpretations of DFT have been suggested [19,20], mostly based on the principle of minimum Fisher information [21]. Our reformulation stands out from three different aspects: First, the maximum entropy foundation on which our formulation is constructed is a completely general method of inference about quantum systems with incomplete information, regardless of the source of the information. In this interpretation of the maximum entropy, information is neither a physical quantity stored in the system, nor is an amount of uncertainty in the probability



Citation: Yousefi, A.; Caticha, A. Entropic Density Functional Theory. *Entropy* **2024**, *26*, 10. <https://doi.org/10.3390/e26010010>

Academic Editor: Carlo Cafaro

Received: 22 November 2023

Revised: 19 December 2023

Accepted: 20 December 2023

Published: 21 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

distribution or the density matrix; it is rather a constraint under which one is to update their degree of rational belief. This is important, especially in the case of DFT. The constraint on the density function arises from a computational assumption that the particle density function of the system is known. However, this information is neither gathered by measurements nor directly obtained from the equilibrium density matrix. Second, we work in the canonical ensemble with a fixed number of particles. This liberates the foundation of DFT from the second quantization. Third and finally, our reformulation is carried out for the more general DFT at finite temperature, which, to our knowledge, had not been tackled by information-theoretic approaches prior to this work.

In Section 2, we review the use of relative entropy as a tool to update density operators in response to new information and we extend the use of entropy as a tool to derive optimal approximations from the classical context [16] to the quantum domain. In Section 3, we construct the entropic DFT formalism and prove a form of the Hohenberg–Kohn theorem at finite temperature within the canonical (fixed number of particles) framework. In Section 4, as an illustration of the eDFT formalism, we discuss the Kohn–Sham model in the local density approximation. Finally, in Section 5, we summarize our conclusions.

2. Preliminaries

The realization that a fundamental theory such as thermodynamics should be interpreted as an application of a general scheme for inference on the basis of information codified into constraints can be traced to Brillouin and Jaynes [22–26]. According to Jaynes—as motivated by Shannon’s axioms [27]—entropy is interpreted as the amount of information that is missing in a probability distribution. The preferred probability distribution is that which agrees with what we know—the information codified into the constraints—but is maximally ignorant about everything else. Thus, one is led to maximize the entropy subject to constraints, a procedure dubbed the MaxEnt method.

A drawback of this approach is that the interpretation of entropy as an amount of missing information is not completely satisfactory. To address this problem, Shore and Johnson [28] proposed that one could avoid invoking questionable measures of information by directly axiomatizing the method for updating probabilities through a variational principle that involved maximizing an entropy functional satisfying certain desirable properties. The question of why one should adopt a variational principle was later clarified by Skilling [29], who proposed a simple ranking strategy; in order to select an optimal distribution (he had in mind the more general case of positive additive distributions which include, e.g., intensities in an image), one proceeds by ranking the distributions according to some preference criteria and then choosing the one which ranks the highest. The ranking scheme is naturally implemented by associating a real number—the entropy—to each distribution with the preference criteria fixed through the axioms of Shore and Johnson. In later work, the nature of the method of maximum entropy was further streamlined as a scheme *designed* to update probabilities when confronted with new information. In this approach, the question “what is information?” receives a very simple answer. *Information is just the constraints we decide to impose on our beliefs*, and there is no need to define “amounts” of information. The motivation behind the design criteria was clarified and their number reduced from five to two [30–33] (reviewed in [34,35]).

2.1. The Quantum MaxEnt Method

The task of extending the method of maximum entropy to the quantum domain as a method to update density operators was carried out by Vanslette [33]. The goal is to update a prior density operator $\hat{\sigma}$ when provided with new information in the form of the expected value of some self-adjoint operators $\langle \hat{A}_i \rangle = A_i$. Vanslette showed that the Umegaki relative entropy [36],

$$S_r[\hat{\rho}|\hat{\sigma}] \stackrel{\text{def}}{=} -\text{Tr}(\hat{\rho} \log \hat{\rho} - \hat{\rho} \log \hat{\sigma}), \quad (1)$$

provides the *unique* criterion to rank density operators $\hat{\rho}$ relative to the prior $\hat{\sigma}$.

The maximization of $S_r[\hat{\rho}|\hat{\sigma}]$ subject to the constraints $\langle \hat{A}_i \rangle = A_i$ and normalization,

$$\delta \left[S_r[\hat{\rho}|\hat{\sigma}] + \alpha_0(1 - \text{Tr}\hat{\rho}) + \sum_i \alpha_i(A_i - \text{Tr}\hat{\rho}\hat{A}_i) \right] = 0, \tag{2}$$

leads to the posterior density operator

$$\hat{\rho}^* = \frac{1}{Z} \exp \left(\log \hat{\sigma} - \sum_i \alpha_i \hat{A}_i \right), \tag{3}$$

where

$$Z(\alpha_i) = e^{\alpha_0} = \text{Tr} \exp \left(\log \hat{\sigma} - \sum_i \alpha_i \hat{A}_i \right). \tag{4}$$

Substituting $\hat{\rho}^*$ back into Equation (1) gives the value of the maximized entropy,

$$S(A_i) \stackrel{\text{def}}{=} S_r[\hat{\rho}^*|\hat{\sigma}] = \sum_i \alpha_i A_i + \log Z. \tag{5}$$

It is widely known that the classical MaxEnt method leads to a mathematical formalism characterized by a contact structure (see, e.g., [37,38]). In a parallel development, the use of Legendre transforms in the context of DFT has also been widely explored [7,39,40]. These results can be extended to the quantum domain, leading to a similar contact structure (see, e.g., [41]). The significance of these results is that the physical content of the formalism is preserved under Legendre transformations quite independently of restrictions to thermal equilibrium and of the physical significance of the so-called “free energies” or Massieu functions.

2.2. Optimal Approximations of Density Operators

The last prerequisite for the construction of the DFT formalism is a systematic method of approximation for density operators. The method we adopt is an extension of the technique developed by Tseng and Caticha in the classical context [16]. The problem is that the exact probability distributions Q obtained using the MaxEnt method are often too intractable to be useful in actual calculations. The solution is to consider a family of more tractable trial distributions P_θ dependent on some parameters θ . The goal is to select the trial distribution P_{θ^*} that best approximates the exact distribution Q . In [16], it was argued that the criterion to select the optimal parameters θ^* is again provided by the method of maximum entropy; the optimal P_{θ^*} is that which is “closest” to the exact Q in the sense that it maximizes the relative entropy $S[P_\theta|Q]$.

Next, we extend this approximation technique to the quantum domain. We consider a family of tractable density operators $\hat{\rho}_\theta$ parameterized by parameters θ . The member of the trial family $\hat{\rho}_\theta$ that best approximates the exact density operator $\hat{\rho}^*$ is the one which maximizes the entropy of $\hat{\rho}_\theta$ relative to $\hat{\rho}^*$:

$$\left. \frac{\partial}{\partial \theta} S_r[\hat{\rho}_\theta|\hat{\rho}^*] \right|_{\theta=\theta^*} = 0. \tag{6}$$

As an example, consider the special case where $\hat{\rho}^*$ and $\hat{\rho}_\theta$ take the exponential form,

$$\hat{\rho}^* = \frac{1}{Z} e^{-\beta \hat{H}} \quad \text{and} \quad \hat{\rho}_\theta = \frac{1}{Z_\theta} e^{-\beta \hat{H}_\theta}; \tag{7}$$

The Gibbs inequality,

$$S_r[\hat{\rho}_\theta|\hat{\rho}^*] \leq 0, \tag{8}$$

reduces to the Bogolyubov inequality,

$$F \leq F_\theta + \langle \hat{H} - \hat{H}_\theta \rangle_\theta, \tag{9}$$

where

$$F = -\frac{1}{\beta} \log Z, \quad F_\theta = -\frac{1}{\beta} \log Z_\theta, \quad \text{and} \quad \langle \cdot \rangle_\theta = \text{Tr}[\hat{\rho}_\theta(\cdot)]. \tag{10}$$

Thus, the argument above shows that the popular approximation method based on the Bogolyubov inequality (see, e.g., [42]) is a special case of the more general approximation method based on entropy maximization.

3. Density Functional Formalism

The goal of the DFT formalism is to find tractable approximations to study the structure of matter. The first crucial step is to recognize that the quantity that captures the desired structural information is the electron density $n(x)$. We wish to design a formalism in which the central role played by the electron density is explicitly displayed.

In the absence of magnetic fields, the time-independent Schrödinger equation for an electron gas of N particles is

$$\hat{H}|\psi\rangle = E|\psi\rangle, \quad (11)$$

where

$$\hat{H}_v = \hat{H}^{(0)} + \hat{V} = \hat{K} + \hat{U} + \hat{V} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} + \frac{e^2}{2} \sum_{j \neq k}^N \frac{1}{|\hat{x}_j - \hat{x}_k|} + \sum_{l=1}^N v(\hat{x}_l), \quad (12)$$

and $|\psi\rangle$ is an antisymmetrized product of N two-spinor orbitals. The potential \hat{U} describes interparticle interactions and the potential \hat{V} describes interactions with nuclei and other external potentials.

3.1. Introducing Density as the Relevant Variable

We are interested in the thermal properties of an inhomogeneous electron fluid, and therefore, we need trial states that describe both thermal equilibrium and inhomogeneity. The former is imposed by a constraint on the expected value of energy and the latter is incorporated by constraints on the expected value $n(x)$ of the electron density $\hat{n}(x)$. The continuous density function $n(x)$ plays a role analogous to the discrete parameters θ in Equations (6)–(10).

Adopting a uniform prior, the relevant trial states are obtained by maximizing the entropy

$$S_r[\hat{\rho}|\hat{1}] = -\text{Tr}\hat{\rho} \log \hat{\rho}, \quad (13)$$

subject to the constraints

$$\text{Tr}\hat{\rho} = 1, \quad (14)$$

$$\text{Tr}\hat{\rho}\hat{H}_v = E, \quad (15)$$

$$\text{and } \text{Tr}\hat{\rho}\hat{n}(x) = n(x), \quad (16)$$

where

$$\hat{n}(x) = \sum_{i=1}^N \delta(\hat{x}_i - x) \quad \text{and} \quad \int d^3x n(x) = N. \quad (17)$$

To be clear, throughout this work, the trace is taken over the Hilbert space of a fixed number N of particles, and in this respect, our formalism *resembles* the canonical ensemble approach. Indeed, all states $|\psi\rangle$ in the Hilbert space are eigenstates of the number operator,

$$\hat{N}|\psi\rangle = \int d^3x \hat{n}(x)|\psi\rangle = N|\psi\rangle \quad \text{so that} \quad \langle\psi|\hat{N}|\psi\rangle = N, \quad (18)$$

but they need not be eigenstates of the density operators $\hat{n}(x)$. Our formalism *differs* from the canonical formalism in that Equation (16) represents an additional infinite number of constraints—one constraint on the expected density function $n(x)$ at each point in space. Due to (18), the expected density function $n(x)$ is not arbitrary; it is constrained to obey (18).

Proceeding to the MaxEnt analog of Equation (3), we find the trial density operator

$$\hat{\rho}_n = \frac{1}{Z_v} \exp\left(-\beta \hat{H}_v - \int d^3x \alpha(x) \hat{n}(x)\right), \tag{19}$$

where

$$Z_v(\beta; \alpha] = \text{Tr} \exp\left(-\beta \hat{H}_v - \int d^3x \alpha(x) \hat{n}(x)\right), \tag{20}$$

and where β and the infinite number of Lagrange multipliers $\alpha(x)$ are implicitly determined by

$$\frac{\partial \log Z_v(\beta; \alpha]}{\partial \beta} = -E \quad \text{and} \quad \frac{\delta \log Z_v(\beta; \alpha]}{\delta \alpha(x)} = -n(x), \tag{21}$$

with the additional constraint (17),

$$\int d^3x n(x) = - \int d^3x \frac{\delta Z_v(\beta; \alpha]}{\delta \alpha(x)} = N. \tag{22}$$

The notation $Z_v(\beta; \alpha]$ indicates that Z is a function of β and a functional of $\alpha(x)$ and depends on $v(x)$ through the Hamiltonian \hat{H}_v . At this point in the argument, there is no implication that the trial states $\hat{\rho}_n$ are in any way more computationally tractable than the exact state $\hat{\rho}^*$ obtained from (19) by setting $\alpha(x)$ to zero.

Next, we calculate the entropy of $\hat{\rho}_n$ relative to the uniform prior to define the trial entropy,

$$S_r[\hat{\rho}_n|\hat{1}] = \beta E + \int d^3x \alpha(x) n(x) + \log Z_v(\beta; \alpha] \stackrel{\text{def}}{=} S_v(E; n]. \tag{23}$$

An important symmetry of the DFT formalism, which is what makes the whole DFT formalism work, arises from the fact that the dependence of $\hat{\rho}_n$ and $Z_v(\beta; \alpha]$ on $v(x)$ and $\alpha(x)$ occurs only through the particular combination

$$\alpha_{\text{int}}(x) \stackrel{\text{def}}{=} \alpha(x) + \beta v(x). \tag{24}$$

The reason for the subscript ‘int’, which denotes ‘intrinsic’, will become clear later in Equation (56). This DFT symmetry implies that a change in the potential $v(x)$ can be compensated by a suitable change in the multiplier $\alpha(x)$ in such a way that $\alpha_{\text{int}}(x)$ and the expected density $n(x)$ remain unaffected. From (12) and (24), we find that (20) can be written as

$$Z_v(\beta; \alpha] = \text{Tr} \exp\left(-\beta \hat{H}^{(0)} - \int d^3x \alpha_{\text{int}}(x) \hat{n}(x)\right) \stackrel{\text{def}}{=} Z(\beta; \alpha_{\text{int}}], \tag{25}$$

so that Equations (19) and (21) become

$$\hat{\rho}_n = \frac{1}{Z(\beta; \alpha_{\text{int}}]} \exp\left(-\beta \hat{H}^{(0)} - \int d^3x \alpha_{\text{int}}(x) \hat{n}(x)\right) \tag{26}$$

and

$$n(x) = - \frac{\delta \log Z(\beta; \alpha_{\text{int}}]}{\delta \alpha_{\text{int}}(x)}. \tag{27}$$

3.2. The Entropic DFT Variational Principle

The exact canonical density operator $\hat{\rho}^*$ is found by maximizing (13) subject to (14) and (15). The result can be read off Equation (19) by setting $\alpha(x) = 0$:

$$\hat{\rho}^* = \frac{1}{Z_v(\beta)} \exp(-\beta \hat{H}_v) \quad \text{and} \quad Z_v(\beta) = \text{Tr} \exp(-\beta \hat{H}_v). \tag{28}$$

(We use a star * to denote exact canonical quantities.) The goal is to approximate $\hat{\rho}^*$ by the best matching member of the family $\{\hat{\rho}_n\}$ with all density operators referring to the same β and N . This involves maximizing the entropy of $\hat{\rho}_n$ relative to $\hat{\rho}^*$:

$$\left. \frac{\delta S_r[\hat{\rho}_n|\hat{\rho}^*]}{\delta n(x)} \right|_{\beta, N} = 0. \tag{29}$$

From (19) and (28), we find

$$S_r[\hat{\rho}_n|\hat{\rho}^*] = \int d^3x \alpha(x)n(x) + \log Z_v(\beta; \alpha] - \log Z_v(\beta). \tag{30}$$

Introducing a Lagrange multiplier α^* to enforce the constraint on N , we have

$$\left. \frac{\delta}{\delta n(x)} \left[S_r[\hat{\rho}_n|\hat{\rho}^*] + \alpha^* \left(N - \int d^3x' n(x') \right) \right] \right|_{\beta} = 0. \tag{31}$$

From the construction above, one might expect that the optimal $\hat{\rho}_n$ coincides with the exact $\hat{\rho}^*$. We can check that this is indeed the case. Substituting Equation (30) into (31), we find

$$\int d^3x' \left[n(x') + \frac{\delta \log Z_v(\beta; \alpha]}{\delta \alpha(x')} \right] \frac{\delta \alpha(x')}{\delta n(x)} = \alpha^* - \alpha(x) \tag{32}$$

The LHS vanishes by Equation (21). Therefore, the optimal $\hat{\rho}_n$ is achieved for $\alpha(x) = \alpha^*$. From (19), (28) and (30), we see that $\alpha^* = 0$, which means that imposing the N constraint was unnecessary; the optimal density reproduces the exact density $n^*(x)$, whether the variations $\delta n(x)$ preserve the total N or not.

We conclude that the entropic DFT variational principle,

$$\left. \frac{\delta S_r[\hat{\rho}_n|\hat{\rho}^*]}{\delta n(x)} \right|_{n^*(x)} = 0, \tag{33}$$

leads to an optimal $\hat{\rho}_n$ which coincides with the exact canonical $\hat{\rho}^*$ in Equation (28):

$$\hat{\rho}_n^{\text{opt}} = \hat{\rho}^*, \quad \text{where} \quad \alpha^{\text{opt}}(x) = \alpha^* = 0. \tag{34}$$

Thus, at this point, our “approximation” scheme is (trivially) exact; by explicit construction, we have demonstrated the existence of a functional of the density $n(x)$, β , and N —the relative entropy $S_r[\hat{\rho}_n|\hat{\rho}^*]$ —that assumes its maximum value at the exact density $n^*(x)$. At this point, however, we have not yet shown that this variational principle is equivalent to the thermal DFT principle derived by Mermin [6]. This we show next.

3.3. The DFT Theorem

Equations (23) and (30) allow us to write

$$S_r[\hat{\rho}_n|\hat{\rho}^*] = -\beta \Omega_v(\beta; n] - \log Z_v(\beta) \tag{35}$$

where we have introduced the “free energy” functional

$$\Omega_v(\beta; n] \stackrel{\text{def}}{=} E - \frac{1}{\beta} S_v(E; n]. \tag{36}$$

The new functional Ω_v ,

$$\Omega_v(\beta; n] = -\frac{1}{\beta} \int d^3x \alpha(x)n(x) - \frac{1}{\beta} \log Z_v(\beta; \alpha], \tag{37}$$

allows us to rewrite the entropic variational principle (31) as

$$\left. \frac{\delta \Omega_v(\beta; n]}{\delta n(x)} \right|_{n^*(x)} = 0. \tag{38}$$

The optimal density $n^*(x)$ is found by minimizing $\Omega_v(\beta; n]$ at fixed β and N . Furthermore, from (37), the multipliers $\alpha(x)$ are obtained from

$$\alpha(x) = -\beta \frac{\delta \Omega_v(\beta; n]}{\delta n(x)}. \tag{39}$$

From Equation (34), $\alpha^{\text{opt}}(x) = \alpha^* = \text{const}$, we obtain

$$\nabla \left. \frac{\delta \Omega_v(\beta; n]}{\delta n(x)} \right|_{n^*(x)} = 0, \tag{40}$$

which has been called the “core integro-differential equation of DFT” [11].

To proceed further, substitute (12), (15), into (36) to find

$$\Omega_v(\beta; n] = \langle \hat{K} + \hat{U} \rangle_{\hat{\rho}_n} + \int d^3x v(x)n(x) - \frac{1}{\beta} S_v(E; n], \tag{41}$$

so that

$$\Omega_v(\beta; n] = F_v(\beta; n] + \int d^3x v(x)n(x), \tag{42}$$

where we have introduced

$$F_v(\beta; n] \stackrel{\text{def}}{=} \langle \hat{K} + \hat{U} \rangle_{\hat{\rho}_n} - \frac{1}{\beta} S_v(E; n]. \tag{43}$$

We are now ready to state the finite temperature DFT theorem.

Theorem 1 (The Density Functional Theorem). *The density functional $F_v[n]$ is independent of the external potential $v(x)$,*

$$\left. \frac{\delta F_v(\beta; n]}{\delta v(x)} \right|_{\beta, n(x)} = 0. \tag{44}$$

This result justifies dropping the index v ,

$$F(\beta; n] \stackrel{\text{def}}{=} F_v(\beta; n], \tag{45}$$

and referring to $F(\beta; n]$ as the intrinsic density functional (The term ‘intrinsic’ indicates that $F(\beta; n]$ is independent of the external potential $v(x)$).

Proof. The crucial observation behind the entropic DFT formalism is that $\hat{\rho}_n$ and $Z_v(\beta; \alpha]$ depend on the external potential $v(x)$ and the Lagrange multiplier function $\alpha(x)$ only through the particular combination $\alpha_{\text{int}}(x)$ defined in (24). Substitute (23)–(25) into (43) to obtain

$$F_v(\beta; n] = -\frac{1}{\beta} \int d^3x \alpha_{\text{int}}(x)n(x) - \frac{1}{\beta} \log Z(\beta; \alpha_{\text{int}}]. \tag{46}$$

Then, the derivative $\delta/\delta v(x')$ at fixed β and $n(x)$ is

$$\left. \frac{\delta F_v(\beta; n]}{\delta v(x')} \right|_{\beta, n(x)} = \int d^3x'' \frac{\delta F_v(\beta; n]}{\delta \alpha_{\text{int}}(x'')} \frac{\delta \alpha_{\text{int}}(x'')}{\delta v(x')} \Big|_{\beta, n(x)}. \tag{47}$$

Equation (26) shows that keeping $n(x)$ fixed is achieved by keeping $\alpha_{\text{int}}(x)$ fixed and vice versa; therefore,

$$\left. \frac{\delta \alpha_{\text{int}}(x'')}{\delta v(x')} \right|_{\beta, n(x)} = \left. \frac{\delta \alpha_{\text{int}}(x'')}{\delta v(x')} \right|_{\beta, \alpha_{\text{int}}(x)} = 0, \tag{48}$$

which implies (44) and concludes the proof. \square

Equations (19) and (40) suggest that (up to an additive constant) the multiplier $\alpha(x)$ plays a role *analogous* to that of a chemical potential. Let us then use Equation (39) to introduce

$$\gamma(x) \stackrel{\text{def}}{=} -\frac{\alpha(x)}{\beta} = \frac{\delta \Omega_v(\beta; n]}{\delta n(x)}, \tag{49}$$

which we shall call the “local chemical potential”. The core Equation (40) has a natural interpretation; the condition for neighboring volume elements to be in equilibrium is that the local chemical potential be uniform:

$$\nabla \gamma(x)|_{n^*} = 0. \tag{50}$$

The optimal value of $\gamma(x)$ is

$$\gamma^* = -\frac{\alpha^*}{\beta} = 0 \quad \text{so that} \quad \nabla \gamma^* = 0. \tag{51}$$

From Equation (42), we have

$$\delta \Omega_v(\beta; n] = \delta F(\beta; n] + \int d^3x [n(x)\delta v(x) + v(x)\delta n(x)], \tag{52}$$

while Equation (49) gives

$$\delta \Omega_v(\beta; n] = \int d^3x \left(\frac{\delta \Omega_v}{\delta v(x)} \delta v(x) + \frac{\delta \Omega_v}{\delta n(x)} \delta n(x) \right) = \int d^3x [n(x)\delta v(x) + \gamma(x)\delta n(x)]. \tag{53}$$

Subtracting these two equations gives

$$\delta F(\beta; n] = \int d^3x [\gamma(x) - v(x)]\delta n(x), \tag{54}$$

which shows that $\delta F/\delta n$ can be interpreted as the *local intrinsic chemical potential*,

$$\frac{\delta F(\beta; n]}{\delta n(x)} \stackrel{\text{def}}{=} \gamma_{\text{int}}(x), \tag{55}$$

with

$$\gamma(x) = \gamma_{\text{int}}(x) + v(x) \quad \text{and} \quad \gamma_{\text{int}}(x) = -\frac{\alpha_{\text{int}}(x)}{\beta}. \tag{56}$$

Evaluating at n^* gives the *equilibrium* intrinsic chemical potential,

$$\left. \frac{\delta F(\beta; n]}{\delta n(x)} \right|_{n^*} = \gamma_{\text{int}}^*(x). \tag{57}$$

(The term ‘intrinsic’ reminds us that both $\gamma_{\text{int}}(x)$ and $\gamma_{\text{int}}^*(x)$ are independent of the external potential $v(x)$).

We mentioned earlier that the multiplier $\alpha(x)$ plays a role *analogous* to that of a chemical potential. We can now be more explicit. Let

$$\mu(x) = \mu(x; n] \quad \text{and} \quad \mu_{\text{int}}(x) = \mu_{\text{int}}(x; n] \tag{58}$$

be the actual chemical potential and the intrinsic chemical potential at location x ,

$$\mu(x; n] = \mu_{\text{int}}(x; n] + v(x). \quad (59)$$

Equilibrium among different volume elements is achieved when

$$\nabla \mu(x; n] |_{n^*} = 0 \implies \mu(x; n^*] = \mu^* = \text{const.} \quad (60)$$

Then, Equations (49)–(51) lead us to identify

$$\gamma(x) = \mu(x) - \mu^* \quad \text{so that} \quad \gamma^* = 0, \quad (61)$$

and

$$\gamma_{\text{int}}(x) = \mu_{\text{int}}(x; n] - \mu(x; n^*]. \quad (62)$$

We can express the eDFT variational principle in terms of F . Using Equations (38) and (42), we find

$$\frac{\delta}{\delta n(x)} \left(F(\beta; n] + \int d^3x' v(x') n(x') \right) \Big|_{n^*(x)} = 0. \quad (63)$$

To summarize, we have reproduced the foundational theorem behind the thermal DFT formalism as an application of maximum entropy methods. This is the main result of this paper. The treatment, so far, has been exact. In the next section, as an illustration of the method, we adapt the well-known Kohn–Sham model to the entropic DFT approach.

4. The Kohn–Sham Approximation Scheme

The exact calculation of $F(\beta; n]$ requires calculating $Z(\beta; \alpha_{\text{int}}]$. Unfortunately, this is just as difficult as calculating the original canonical partition function $Z_v(\beta)$, which was precisely what we wanted to avoid. An analogous problem arises in the standard many-body theory; even for relatively small particle numbers, the calculation of the N -particle wave function becomes impractically difficult because the wave function $\Psi(\vec{r}_1 \dots \vec{r}_N)$ lives in a $3N$ -dimensional configuration space. The DFT framework attempts to evade this problem by focusing attention on the hopefully easier problem of calculating the density $n(x)$, which is a *function* that lives in a mere three dimensions. Unfortunately, the problem is not solved, but merely transferred to the calculation of the *functional* $F(\beta; n]$. Not all is lost, however, because the reformulation in terms of the density $n(x)$ suggests new useful approximations.

The discussion below parallels closely the ground state formulation of Kohn and Sham [5]. It differs from the grand canonical thermal DFT of Mermin [6] in that here we remain within the canonical framework of fixed particle number. In common with the Hartree–Fock approximation, the Kohn–Sham model reduces an interacting many-particle Schrödinger equation to that of a single particle in the presence of an effective potential that includes exchange and correlation effects. An important advantage is that, unlike Hartree–Fock, the Kohn–Sham framework can in principle be exact. In practice, however, the success of the model hinges on whether the approximations for exchange and correlations are sufficiently simple and accurate. Fortunately, the “local density approximation”, which is exact for a uniform electron gas, and should remain valid for slowly varying potentials, has turned out to be quite successful for the prediction of bond lengths and molecular structures even when these involve inhomogeneities at the atomic scale.

Referring to Equation (43), the idea is that $F(\beta; n]$ can be split into three terms,

$$F(\beta; n] = F_0(\beta; n] + U_C[n] + F_{\text{xc}}(\beta; n]. \quad (64)$$

The first term $F_0(\beta; n)$ represents the intrinsic free energy of a gas of *non-interacting* and *uncorrelated* particles at the same temperature and density. The second term $U_C[n]$ is the classical Coulomb interaction,

$$U_C[n] = \frac{e^2}{2} \int d^3x d^3x' \frac{n(x)n(x')}{|x - x'|}, \tag{65}$$

which represents the dominant contribution from the interparticle potential term $\langle \hat{U} \rangle_{\hat{\rho}_n}$ in (43). The third term $F_{xc}(\beta; n)$ is a correction that accounts for all additional exchange and correlations effects. To the extent that we can define $F_{xc}(\beta; n)$ to be the difference

$$F_{xc}(\beta; n) \stackrel{\text{def}}{=} F(\beta; n) - F_0(\beta; n) - U_C[n], \tag{66}$$

Equation (64) is trivially exact.

We are now ready to substitute (64) into the eDFT variational principle (63). The result is

$$\left[\frac{\delta F_0}{\delta n(x)} + v(x) + \int d^3x' \frac{e^2 n(x')}{|x - x'|} + v_{xc}(x; n) \right]_{n^*(x)} = 0, \tag{67}$$

where we introduced

$$v_{xc}(x; n) \stackrel{\text{def}}{=} \frac{\delta F_{xc}}{\delta n(x)}. \tag{68}$$

So far this is exact. However, to make further progress, we note that although exchange correlations are intrinsically non-local, for a thermal system, we can assume that entanglement effects are appreciable only over short distances. Therefore, it might not be unreasonable to approximate F_{xc} by a sum over independent volume elements. Accordingly, we adopt the so-called local density approximation,

$$F_{xc}(\beta; n) \approx F_{xc}^{LDA}(\beta; n) = \int d^3x f_{xc}(n(x))n(x), \tag{69}$$

where the function $f_{xc}(n)$ is assumed to be known; it is the exchange correlation free energy per particle for a uniform electron gas with density n . The corresponding potential

$$v_{xc}(x) = \left. \frac{d}{du} (f_{xc}(u)u) \right|_{u=n(x)} \tag{70}$$

is therefore also known.

To find the optimal density $n^*(x)$ that solves the variational Equation (67), we can use the same trick introduced by Kohn and Sham. They noticed that their variational equation for the ground state—the analogue of our Equation (67)—is exactly of the form one obtains for a gas of non-interacting and uncorrelated particles moving in an effective single-particle potential. This leads us to rewrite (67) as

$$\left[\frac{\delta F_0}{\delta n(x)} + v_{\text{eff}}(x) \right]_{n^*(x)} = 0, \tag{71}$$

where

$$v_{\text{eff}}(x) = v(x) + \int d^3x' \frac{e^2 n(x')}{|x - x'|} + v_{xc}(x). \tag{72}$$

Thus, the problem of N interacting particles has been translated into the problem of a single particle moving in a density-dependent effective potential created by all the other particles. This shows that we can adopt the same iterative procedure followed with the Hartree

self-consistent potential. If $n^{(j)}(x)$ is the density at the j th iteration, use (72) to construct the potential $v_{\text{eff}}^{(j)}(x)$, and solve the single-particle equation,

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}^{(j)}(x)\right]\psi_k^{(j)}(x) = \varepsilon_k^{(j)}\psi_k^{(j)}(x). \quad (73)$$

Then, we construct the density $n^{(j+1)}(x)$ for the next iteration as the thermal average,

$$n^{(j+1)}(x) = \sum_{k=1}^{k_{\text{max}}} \frac{|\psi_k^{(j)}(x)|^2}{1 + \exp[\beta(\varepsilon_k^{(j)} - \mu)]} \quad (74)$$

where the cutoff k_{max} is such that the occupation of orbitals with $k > k_{\text{max}}$ can be neglected and μ is found by imposing $\int d^3x n(x) = N$. The process is repeated until convergence to the optimal n^* is achieved.

Just as in the standard Kohn–Sham model, neither the single-particle potential $v_{\text{eff}}(x)$ nor the wave functions ψ_k and energies ε_k are to be given any real physical interpretation. They are auxiliary quantities whose only purpose is the calculation of the physical density $n^*(x)$.

5. Conclusions

To summarize our conclusions:

We have produced a *reconstruction* of DFT that makes explicit how DFT fits within an ongoing research program that places the concepts of entropy and information at the very foundation for all of physics (see, e.g., [34]). This includes statistical mechanics [22–26], quantum mechanics [35,43], and as we have shown in this work, also the main techniques to study structure—variational principles including mean field methods and DFT.

We extended the use of entropy as a systematic method to generate optimal approximations from the classical to the quantum domain. This allowed an entropic reconstruction of quantum DFT. This process involves a family of trial density operators parameterized by the particle density. The optimal density operator is found by maximizing the quantum entropy relative to the exact canonical density operator. This approach reproduces the variational principle of DFT and allows a proof of the Hohenberg–Kohn theorem at finite temperature that is simpler in that it evades some of the subtleties of the ground state formalism. Our formalism differs from previous approaches in that (i) the central role of entropy is explicit, and (ii) we remain with the canonical ensemble formalism.

Author Contributions: A.Y. and A.C. contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We are thankful to Oleg Lunin, Carlo Cafaro, Herbert Fotsos, and Daniel Robins, for their insightful comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kohn, W. Noble Lecture: Electronic structure of matter—Wave functions and density functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253. [CrossRef]
2. Jones, R.O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897. [CrossRef]
3. Argaman, M.; Makov, G. Density functional theory: An introduction. *Am. J. Phys.* **2000**, *68*, 69. [CrossRef]
4. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864. [CrossRef]
5. Kohn, W.; Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133. [CrossRef]
6. Mermin, D. Thermal properties of inhomogeneous electron gas. *Phys. Rev.* **1965**, *137*, A1441. [CrossRef]
7. Eschrig, H. $T > 0$ ensemble-state density functional theory via Legendre transform. *Phys. Rev. B* **2010**, *82*, 205120. [CrossRef]

8. Pribram-Jones, A.; Pittalis, S.; Gross, E.; Burke, K. Thermal Density Functional Theory in Context. In *Frontiers and Challenges in Warm Dense Matter, Lecture Notes in Computational Science and Engineering*; Graziani, F., Desjarlais, M.P., Redmer, R., Trickey, S.B., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2014; Volume 96, pp. 25–60.
9. Burke, K.; Smith, J.C.; Grabowski, P.E.; Pribram-Jones, A. Exact conditions on the temperature dependence of density functionals. *Phys. Rev. B* **2016**, *93*, 195132. [[CrossRef](#)]
10. Ebner, C.; Saam, W.F.; Stroud, D. Density-functional theory of simple classical fluids. I. Surfaces. *Phys. Rev. A* **1976**, *14*, 2264. [[CrossRef](#)]
11. Evans, R. The nature of the liquid-vapor interface and other topics in the statistical mechanics of non-uniform classical fluids. *Adv. Phys.* **1979**, *28*, 143–200. [[CrossRef](#)]
12. Tarazona, P. Free-energy density functional for hard spheres. *Phys. Rev. A* **1985**, *31*, 2672. [[CrossRef](#)] [[PubMed](#)]
13. Rosenfeld, Y. Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing. *Phys. Rev. Lett.* **1989**, *63*, 980. [[CrossRef](#)] [[PubMed](#)]
14. Evans, R.; Oettel, M.; Roth, R.; Kahl, G. New developments in classical density functional theory. *J. Phys. Condens. Matter* **2016**, *28*, 240401. [[CrossRef](#)] [[PubMed](#)]
15. Yousefi, A.; Caticha, A. An entropic approach to classical density functional theory. *Phys. Sci. Forum* **2021**, *3*, 13.
16. Tseng, C.-Y.; Caticha, A. Using relative entropy to find optimal approximations: An application to simple fluids. *Phys. A* **2008**, *387*, 6759. [[CrossRef](#)]
17. Alipour, M.; Badooei, Z. Information theoretic approach provides a reliable description for kinetic component of correlation energy density functional. *Int. J. Quantum Chem.* **2018**, *118*, 23. [[CrossRef](#)]
18. Rong, C.; Wang, B.; Zhao, D.; Liu, S. Information-theoretic approach in density functional theory and its recent applications to chemical problems. *Wires Comput. Mol. Sci.* **2020**, *10*, 4. [[CrossRef](#)]
19. Nagy, A. Fisher information in density functional theory. *J. Chem. Phys.* **2003**, *119*, 9401. [[CrossRef](#)]
20. Nalewajski, R.F. Information principles in the theory of electronic structure. *Chem. Phys. Lett.* **2003**, *372*, 28–34. [[CrossRef](#)]
21. Freiden, B.R. *Physics from Fisher Information: A Unification*; Cambridge University Press: Cambridge, UK, 1998.
22. Brillouin, L. *Science and Information Theory*; Academic Press: New York, NY, USA, 1952.
23. Brillouin, L. The negentropy principle of information. *J. Appl. Phys.* **1953**, *24*, 1152. [[CrossRef](#)]
24. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620. [[CrossRef](#)]
25. Jaynes, E.T. Information theory and statistical mechanics II. *Phys. Rev.* **1957**, *108*, 171. [[CrossRef](#)]
26. Jaynes, E.T. Where do we stand on maximum entropy? In *The Maximum Entropy Principle*; Levine, R.D., Tribus, M., Eds.; MIT Press: Cambridge, MA, USA, 1979.
27. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 3. [[CrossRef](#)]
28. Shore, J.; Johnson, R. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **1980**, *26*, 26–37. [[CrossRef](#)]
29. Skilling, J. The Axioms of Maximum Entropy. In *Maximum-Entropy and Bayesian Methods in Science and Engineering: Foundations*; Springer: Dordrecht, The Netherlands, 1988; pp. 173–187.
30. Caticha, A. Relative Entropy and Inductive Inference. *arXiv* **2004**, arXiv:physics/0311093.
31. Caticha, A.; Giffin, A. Updating Probabilities. *arXiv* **2006**, arXiv:physics/0608185.
32. Caticha, A. Information and Entropy. *arXiv* **2007**, arXiv:0710.1068.
33. Vanslette, K. Entropic Updating of Probabilities and Density Matrices. *Entropy* **2017**, *19*, 664. [[CrossRef](#)]
34. Caticha, A. Entropic Physics: Probability, Entropy and the Foundations of Physics. Available online: <https://www.arielcaticha.com/> (accessed on 20 November 2023).
35. Caticha, A. Quantum Mechanics as Hamilton-Killing Flows on a Statistical Manifold. *Phys. Sci. Forum* **2021**, *3*, 12.
36. Umegaki, H. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math Sem. Rep.* **1962**, *14*, 59–85. [[CrossRef](#)]
37. Rajeev, S.G. A Hamilton-Jacobi formalism for thermodynamics. *Ann. Phys.* **2008**, *323*, 2265–2285. [[CrossRef](#)]
38. Balian, R.; Valentin, P. Hamiltonian structure of thermodynamics with gauge. *Eur. Phys. J. B* **2001**, *21*, 269–282. [[CrossRef](#)]
39. Lieb, E.H. Density functionals for Coulomb systems. *Int. J. Quantum Chem.* **1983**, *24*, 243. [[CrossRef](#)]
40. Fukuda, R.; Kotani, T.; Suzuki, Y.; Yokojima, S. Density functional theory through Legendre transformation. *Prog. Theor. Phys.* **1994**, *92*, 833–862. [[CrossRef](#)]
41. Yousefi, A. *Entropic Density Functional Theory: Entropic Inference and the Equilibrium State of Inhomogeneous Fluids*; State University of New York at Albany ProQuest Dissertation Publishing: Albany, NY, USA, 2021; p. 28863510.
42. Feynman, R.P. *Statistical Mechanics: A Set of Lectures*; CRC Press: Boca Raton, FL, USA, 1972.
43. Caticha, A. The Entropic Dynamics Approach to Quantum Mechanics. *Entropy* **2019**, *21*, 943. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.