Leading correction to the local density approximation of the kinetic energy in one dimension

Kieron Burke, corresponding author, kieron@uci.edu
Departments of Physics and Astronomy and of Chemistry, University of California, Irvine, CA 92697, USA

(Dated: 15 February 2020)

A mathematical framework is constructed for the sum of the lowest $N$ eigenvalues of a potential. Exactness is illustrated on several one-dimensional systems (harmonic oscillator, particle in a box, and Poschl-Teller well). Semiclassical expansion yields the leading corrections for finite systems, identifying the error in common gradient expansions in density functional theory. Some singularities can be avoided when evaluating the correction to the leading term. Correcting the error in the gradient expansion greatly improves accuracy. The relevance to practical density functional calculations is discussed.

In the tens of thousands of density functional calculations published annually, most employ the gradient of the density to estimate the exchange-correlation energy of the Kohn-Sham equations. Such approximations begin from the gradient expansion of a slowly-varying electron gas, which is then 'generalized' to an integral over an energy density with some function of the density gradient. The first such attempt came already in 1968 when Ma and Bruckner showed that severe problems applying this gradient expansion approximation (GEA) for the correlation energy to atoms could be overcome by this procedure. Since then, a variety of procedures and philosophies have been used to construct such generalized gradient approximations (GGAs) and other more sophisticated constructions. Some GGA's are more accurate and popular in chemistry, while others work better for (weakly correlated) materials. This diversity reflects the ambiguity in their derivation. The older, simpler local density approximation, is uniquely determined by the energy of the uniform electron gas.

Long ago, Lieb and Simon proved that, for any electronic system, the relative error in Thomas-Fermi theory vanishes in a well-defined semiclassical limit in which the particle number tends to infinity. Much work since then studies corrections to this limit order-by-order, including extensions of Thomas-Fermi theory. Such work is sometimes limited to atoms where spherical symmetry simplifies the situation. Engel beautifully summarized work with Schwinger on this subject. However, this problem is complicated by the interaction between electrons, the Coulomb attraction to nuclei, and the complexities of semiclassics in three dimensions.

The present work studies the origin of the errors in applying the gradient expansion in the simplest relevant case, namely the kinetic energy of non-interacting electrons in one dimension. This is not of quantitative relevance to realistic electronic structure calculations. The primary purpose is the construction of a mathematical framework in which this question can be directly addressed, and the errors of the gradient expansion explicitly identified and calculated in a systematic expansion in powers of $\hbar$. We show that, for simple model cases, the formalism is exact, and also calculate the order-by-order expansion, finding great quantitative improvements in energies when the corrections are accounted for. We discuss the nature of these corrections and how they might be incorporated in density functional approximations.

Consider a symmetric potential $v(x)$, with zero chosen so that $v(0) = 0$, and which could tend to $D$, the well-depth, at large $x$. Let $\epsilon_j$ be the eigenvalues of the Schrödinger equation, using (Hartree) atomic units (setting $m = \hbar = 1$), and let $M$ be the highest bound state if there is one. The number staircase is

$$N_c(\epsilon) = \sum_{j=1}^{M} \Theta(\epsilon - \epsilon_j),$$

where $\Theta$ is the Heaviside step function, i.e., this is the number of states with $\epsilon_j < \epsilon$. Next, consider a smooth monotonic function $I(\epsilon)$ such that

$$I(\epsilon_j) = J$$

where $J = j - 1/2$. The 1/2 comes from the Maslov index for two turning points. As $\hbar \to 0$, a possible $I(\epsilon)$ is the classical action across the well, divided by $\pi$. We also define $\epsilon(y)$ as the inverse of $I$, so that $\epsilon_j = \epsilon(J)$. Then

$$N_c(\epsilon) = \lceil I(\epsilon) + \frac{1}{2} \rceil,$$

where $\lceil x \rceil$ is the highest positive integer less than $x$. So $N$ is the nearest integer to $I$. Next, define the periodic function

$$\langle \epsilon \rangle = x - \lfloor x + \frac{1}{2} \rfloor,$$

so that

$$N_c(\epsilon) = I(\epsilon) - \langle I(\epsilon) \rangle.$$  

To invert $N_c(\epsilon)$, we turn on a temperature that is much smaller than any energy or difference:

$$N_p(\epsilon) = \sum_{j=1}^{M} f(\beta(\epsilon - \epsilon_j)),$$

where $f^{-1}(x) = 1 + \exp(-x)$ and $\beta$ is inversely proportional to temperature. We then define $\mu_p(M)$ as the inverse of $N_p$. For any finite temperature, $\mu_p(M)$ exists and is well-defined. We take $\beta \to \infty$ at the end of the derivations and stop mentioning the temperature explicitly.
and a well-known semiclassical result is

$$S(N) = \varepsilon N_c(\varepsilon) - G(\varepsilon),$$

i.e., \(N_c(\varepsilon)\) determines the energy staircase. But we really want \(S\) as a function of continuous number, \(\mathcal{N}\), which is

$$S(\mathcal{N}) = S(\mu(\mathcal{N})) = \mathcal{N} \mu(\mathcal{N}) - G(\mu(\mathcal{N})).$$

This expression is well-defined for continuous values of non-negative \(\mathcal{N}\) (for non-zero temperature). As the temperature goes to zero, it becomes piece-wise linear, with changes of slope at integer values of \(\mathcal{N}\), so that knowledge at integer values is sufficient to determine the whole function. Both \(\mu(\mathcal{N})\) and \(G(\mu(\mathcal{N}))\) have step-like features, yielding

$$S(\mathcal{N}) = S_N + (\mathcal{N} - N) \varepsilon_{N+1}, \quad N = [\mathcal{N}].$$

Note that \(\mu(\mathcal{N}) = \varepsilon(N)\) for integers, where \(\varepsilon = \Gamma^{-1}\), i.e., the discontinuous contributions in \(\mu(\mathcal{N})\) vanish identically at integers, so they are not needed to find \(S_N\). Moreover, we change variables in the integration in \(G\). If \(y = I(\varepsilon)\), then

$$S_N = N \varepsilon(N) - G(N), \quad G(N) = \int_0^N \frac{dy}{I'(y)} (y - \langle y \rangle),$$

where \(I' = dI/d\varepsilon\) and only the last term requires an integral over oscillations. Eq. (12) is a central result, providing the machinery to construct the sum of the eigenvalues directly from \(I(\varepsilon)\), in continuous and discontinuous contributions. We define the first continuous term in \(G\) as

$$J(\varepsilon) = \int_{\varepsilon(0)}^\varepsilon d\varepsilon' I(\varepsilon'), \quad G_{\text{disc}}(N) = J(N) - G(N).$$

The value of \(\varepsilon(0), \text{ negative in Fig. 1, is irrelevant to } G(N), \text{ as the step function vanishes for arguments less than } 1/2, \text{ but not to } J(N) \text{ or } G_{\text{disc}}(N). \text{ Because } J' = I, S_N \text{ is fully determined by } J(\varepsilon). \text{ Changing variables to } \varepsilon(y) \text{ and integrating by parts yields the more succinct}\)

$$S^{\Pi}(\mathcal{N}) = \int_0^\mathcal{N} dy \left\{ \varepsilon(y) + \varepsilon'(y) \langle y \rangle \right\},$$

where \(\Pi\) denotes integer-interpolation, i.e., a quantity that matches \(S\) at integers, but may differ inbetween. Fig. 2 plots quantities versus \(\mathcal{N}\) for a PT well, showing that Eq. (14) agrees with \(S(\mathcal{N})\) only at (half)-integers.

A harmonic oscillator is instructive. Here \(I(\varepsilon) = \varepsilon/\omega, \text{ so } \varepsilon(y) = y \omega, I' = 1/\omega, J = \omega N^2/2, \text{ and } G_{\text{disc}} \text{ vanishes because the average of } \langle y \rangle \text{ over one period vanishes if } I' \text{ is constant. For a particle in an infinite well,}

$$I(\varepsilon) = \frac{L \sqrt{2\varepsilon}}{\pi} - \frac{1}{2},$$

Then \(\mu(N) = \pi^2(N + 1/2)^2/(2L^2)\) and \(1/I' = \pi^2(y + 1/2)/L^2\). Then \(J(\mathcal{N})\) is trivial to integrate but, because \(I'\) varies, \(G_{\text{disc}}\) does not vanish:

$$G_{\text{disc}}(N) = \frac{\pi^2}{L^2} \int_0^N dy \left( y + \frac{1}{2} \right) \langle y \rangle$$

The constant term gives no contribution, while the integral over \(y\langle y \rangle\) is \(-N/24\), producing the exact answer \(S_N = 

Fig. 1 illustrates these functions for a Poschl-Teller well with \(D = 5/2\) (see below). The smooth \(I(\varepsilon)\) generates the staircase \(N_c(\varepsilon)\), whose steps are rounded by the temperature, making it invertible. The difference has a sawtooth shape, crossing zero at the eigenvalues, so that \(N_c = I\) when both are (half)-integers.

Fig. 2. Energy staircase as a function of \(\mathcal{N}\) (black), the integer interpolation of Eq. (14) (cyan), continuous contribution, \(\mathcal{N} \varepsilon(\mathcal{N}) - J(\mathcal{N})\) (red), and discontinuous contribution \(G_{\text{disc}}\) (blue) for the same system as Fig. 1 (zero temperature).

We wish to develop an expression for the sum of the eigenvalues, which would be the total energy of \(N\) same-spin fermions in the well. Define the energy staircase:

$$S(\varepsilon) = \sum_{j=1}^M e_j \Theta(\varepsilon - e_j),$$

i.e., the sum of eigenvalues with energy below \(\varepsilon\). It will be especially useful to consider:

$$G(\varepsilon) = \int_{-\infty}^\varepsilon d\varepsilon' N_c(\varepsilon') = \int_{-\infty}^\varepsilon d\varepsilon' \left\lfloor I(\varepsilon') + \frac{1}{2} \right\rfloor,$$

and a well-known semiclassical result is

$$S(\varepsilon) = \varepsilon N_c(\varepsilon) - G(\varepsilon),$$
$$\pi^2 N (N^2 + 3N/2 + 1/2)/(6L^2).$$ A less trivial example is the Poschl-Teller well of depth $D$:

$$v(x) = D - D/\cosh^2(x).$$

Writing $\alpha_e = \sqrt{2D + 1/4}$, then

$$I(\varepsilon) = \alpha_e - \sqrt{2(D - \varepsilon)},$$

yielding the eigenvalues

$$\varepsilon_j = D - (\alpha_e - j^2)/2, \quad j < \alpha_e. \quad (19)$$

The simple result $I' = 1/(\alpha_e - 1)$ makes the calculation easy, using the same integral over $\langle y \rangle$ as before, giving

$$S_N^{\text{PT}} = \alpha_e N^2 - N^3/6 - N/12. \quad (20)$$

So far, this result might be considered a simple tautology. Its real use comes when a semiclassical expansion is performed. We multiply $h$ by a dimensionless number $\eta$, and consider the limit as $\eta \to 0$. Elementary analysis shows $\varepsilon_j(\eta)[v] = \eta^2 \varepsilon_j[v/\eta^2]$ and, as $\eta \to 0$,

$$I(\eta)[v](\varepsilon) = \int \frac{v}{\eta^2} \left( \frac{\varepsilon}{\eta^2} \right) I(0)[v] + \eta \Delta I^{(2)}[v] + ..., \quad (21)$$

where the expansion is known from WKB theory$^{22}$. Here

$$I^{(0)}[v] = \int_{-\infty}^{\infty} dx \frac{p(x)}{\pi} \quad (22)$$

is the classical action divided by $\pi$, $p(x)$ is the real part of the local classical momentum, $\sqrt{2(\varepsilon - v(x))}$, and yields the (zero-order) WKB eigenvalues. As $\varepsilon^{(0)}(0) = 0$,

$$J^{(0)}(\varepsilon) = \int_{-\infty}^{\infty} dx \frac{p^2(x)}{3}. \quad (23)$$

The WKB expansion is an expansion of individual eigenvalues in powers of $\eta$, keeping $\eta^2$ fixed, but the expansion of $S^{(\eta)}(\mathcal{N})$ keeps $\mathcal{N}$ fixed. One can both sum the WKB values to compare with $S_N$ and also compare $S_N - S_{N-1}$ with the $N$-th WKB eigenvalue. In general, these differ order-by-order (but infinite sums are identical). For example, summing WKB eigenvalues for the PT well produces an additional $N/24$ relative to $S^{(0)}(N)$. For $N > 1$, we expect the expansion of $S(N)$ to outperform the sum of WKB eigenvalues to the same order, as the semiclassical approximation is used only at $N$, and not at each individual $j$ up to $N$, where it should be less accurate.

The leading correction is trickier to evaluate, due to a singularity as the turning points are approached. Define

$$B(\varepsilon,a) = \int_{0}^{b(\varepsilon) - a} dx \frac{v''}{p^2(\varepsilon,x)} \quad (24)$$

where $b(\varepsilon)$ is the turning point at energy $\varepsilon$, $0 < a < b(\varepsilon)$. As $a \to 0$, a singularity develops which must be cancelled:

$$\Delta I^{(2)}(\varepsilon,a) = \frac{1}{12\pi} \left( B(\varepsilon,a) + \frac{b''}{b'\sqrt{2ab}} \right). \quad (25)$$

and $\Delta J^{(2)}(\varepsilon)$ is found by taking $a \to 0$. This cumbersome procedure can be elegantly avoided by an integral over a contour surrounding the turning points$^{22}$. Higher-order terms involve even stronger singularities. For a harmonic oscillator, Dunham$^{23}$ showed that all higher-order terms are identically zero, so that WKB yields the exact answers. Likewise for a particle in a box, as all derivatives of $v$ vanish, but $\Delta I^{(2)} = 1/(8\sqrt{D})$ for PT.

But we can instead evaluate the expansion for $J(\varepsilon)$ and perform the energy integration before the spatial integral$^{24}$. Consider

$$A_\delta(\varepsilon) = \int_{\delta}^{\varepsilon} d\varepsilon' \int_{-b(\varepsilon' - \delta)}^{b(\varepsilon' - \delta)} dx \frac{v''(x)}{p^2(\varepsilon',x)}, \quad (26)$$

where $b(\varepsilon)$ is the turning point at energy $\varepsilon$. For positive $\delta$, this has no singularities and the order of integration can be reversed. As $\delta \to 0$, a singular term appears (of order $1/\sqrt{\delta}$) which cancels of that of Eq. (25). Thus

$$\Delta J^{(2)}(\varepsilon) = \lim_{\delta \to 0} \int_{\delta}^{\varepsilon} d\varepsilon' \Delta J^{(2)}(\varepsilon',\varepsilon'(b)/\delta) = -\frac{1}{12\pi} \int_{0}^{b} dx \frac{v''}{p^3} \quad (27)$$

Because of the oscillation, $G_{\text{disc}}$ is already of higher-order than the continuous terms. Thus

$$G^{(2)}_{\text{disc}}(N) = -\frac{1}{24} \frac{d^2(0)}{d\varepsilon_0^2} N \quad (28)$$

Because of the periodicity of $\langle y \rangle$, only the endpoints contribute to the integral as $\eta \to 0$, yielding

$$G^{(2)}_{\text{disc}}(N) = -\frac{1}{24} \frac{d^2(0)}{d\varepsilon_0^2} N \quad (29)$$

Expanding $\varepsilon(N)$ to second order in $J^{(0)}$ yields

$$\Delta S^{(2)}_{\Delta N} = -\Delta J^{(2)}(\varepsilon(0)(N)) + G^{(2)}_{\text{disc}}(N) \quad (30)$$

which is $N^2/(16\sqrt{2D}) - N/12$ for the PT well.

Finally, we are ready to connect with density functional theory (DFT). For 1d same-spin non-interacting fermions in a slowly-varying potential in an extended system, there is a well-known expansion of both the density $n(x)$ and kinetic energy $T$ in gradients of the potential$^{25}$:

$$n^{(0)}(x) = \frac{p_\varepsilon(x)}{\pi}, \quad \Delta n^{(2)}(x) = \frac{v''(x)}{12\pi p_\varepsilon(x)}, \quad (31)$$

and

$$T^{(0)} = \int dx \frac{p^2_\varepsilon(x)}{6\pi}, \quad \Delta T^{(2)} = \int dx \frac{v''(x)}{8\pi p_\varepsilon(x)}, \quad (32)$$

where $p_\varepsilon(x) = p(\varepsilon_\varepsilon,x)$ and $\int dx n(x) = N$ determines $\varepsilon_\varepsilon$. Zero-order is Thomas-Fermi (TF) theory and 2nd order is the gradient expansion. The combination $\mathcal{N} \varepsilon_\varepsilon(\mathcal{N}) - S(\mathcal{N})$ yields $T^{(0)} + 2\Delta T^{(2)}/3$, agreeing with the semiclassical expansion of $J$, showing

$$S^{(2)}[v] = S^{\text{GEA}}[v] + G^{(2)}_{\text{disc}}[v], \quad (33)$$
where GEA denotes (2nd-order) gradient expansion approximation. The semiclassical expansion for $S_N$ reduces to the gradient expansion for extended systems, where the sawtooth contribution vanishes. But $G^{(2)}_{disc}$ produces the exact leading-order correction to the local approximation for finite systems.

TABLE I. Energy sums for a PT well of $D = 9.555$ and errors of DFT approximations in milliHartree on left; errors in eigenvalues on the right. TF is the leading order, GEA is 2nd-order without discontinuous contributions, 2nd includes them, and 4th-order does also. WKB agrees with $S_N$ at 2nd-order.

<table>
<thead>
<tr>
<th>N</th>
<th>$S_N$</th>
<th>Error x 1000</th>
<th>$\delta S_N$</th>
<th>Error x 1000</th>
<th>$\delta S_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TF</td>
<td>GEA</td>
<td>2nd</td>
<td>4th</td>
<td>WKB</td>
</tr>
<tr>
<td>1</td>
<td>1.95</td>
<td>69</td>
<td>-42</td>
<td>0.0</td>
<td>-0.0001</td>
</tr>
<tr>
<td>2</td>
<td>7.30</td>
<td>110</td>
<td>-83</td>
<td>0.2</td>
<td>-0.0001</td>
</tr>
<tr>
<td>3</td>
<td>15.05</td>
<td>122</td>
<td>-125</td>
<td>0.4</td>
<td>-0.0003</td>
</tr>
<tr>
<td>4</td>
<td>24.20</td>
<td>105</td>
<td>-166</td>
<td>0.7</td>
<td>-0.0005</td>
</tr>
</tbody>
</table>

Some results for a generic Poschl-Teller well with $\alpha = 4.4$ are given in Table I. The left side gives errors for the sum of eigenvalues, the right for the individual levels. The second column of errors on the left shows the result of the 2nd-order GEA, which sometimes worsens results relative to TF theory. But inclusion of the correction reduces those errors by two orders of magnitude. Addition of the next order reduces the errors to the microHartree range. Deeper wells are even more favorable. Switching to the right side, for this well even the eigenvalues are better approximated by differences in the sums within TF theory, but comparison of their second-order contributions shows that, in the asymptotic limit, WKB will have smaller errors than $\Delta TF$ for the top 1/6 th of the levels. Table II repeats this calculation for $D = 1$, which binds one particle with energy $-1/2$ relative to the outside, and has a second level right at threshold, far from the semiclassical limit. The trends are the same, but errors are greater, and WKB does better for the eigenvalue at the top of the well.

Lastly, consider density functionals. Simply invert Eq. (31) and insert the result into Eq. (32) to find

\[ T^{GEA}[n] = \frac{\pi^2}{6} \int dx n^3(x) - \frac{1}{24} \int dx \frac{n'(x)^2}{n}. \]  

(34)

It is straightforward to convert Eq. (29) into a functional of the TF density for the present circumstances. For potentials with a parabolic minimum at the origin:

\[ G^{(2)}_{disc}[n^{(0)}] = \frac{1}{24} \left( \frac{\pi^2}{C} - \pi \sqrt{n^{(0)}(x)n^{(0)'(x)}} \right), \]  

(35)

where

\[ C = \int_{-\infty}^{\infty} dx \left( \frac{\pi}{2} \right)^{1/2} n^{(0)(x)}. \]  

(36)

Thus $G^{(2)}_{disc}$ contains both highly local and non-local contributions (integrals over local functionals). Inserting the TF density for the PT well correctly yields $N/24$. Eq. (35) looks like no local correction currently in the literature; it has been derived, not devised.

The local density approximation applies to almost all situations. The potential functional correction to GEA of Eq. (29) applies to many circumstances, such as semi-infinite systems with surfaces, where the Maslov index differs, but must be generalized for e.g., multiple wells. On the other hand, when converted to a density functional, Eq. (35), the form of the functional depends even further on the general class of problem. For example the form differs from Eq. (35) for $v(x) = |x|$.

No general prescription was given for finding $T[\epsilon]$. For the DFT results, one needs only its well-defined asymptotic expansion. For simple model systems, the formulas used here suffice. But adding any other function that vanishes at the eigenvalues generates equally viable candidates. Different $T[\epsilon]$ yield different continuous and discontinuous contributions, but still yield the exact sums.

Many phenomena in DFT have a simple analog within this 1d world, as shown by two examples. The first is the well-known inaccuracy of functional derivatives of reasonably accurate semilocal approximations for the energy. This infamous misbehavior of the LDA XC potential leads to highly inaccurate KS orbital eigenvalues. The analog here is $N$-particle density

\[ \rho(x) = \frac{\delta S_N[v]}{\delta v(x)}. \]  

(37)

The archetype in 1d is the harmonic oscillator in TF theory, which yields the exact eigenenergies (and their sums), but whose density is highly inaccurate. The local approximation is exact for the harmonic potential, but not when small pointwise changes are made, as in Eq. (37). Only smooth changes in the potential should be expected to be correct in a local theory (the first four moments (0-3) of the TF density of the oscillator are exact!). Including the second-order correction yields densities that are singular at the turning points. This simply reflects the incompatibility of the order of limits, by expanding in $\hbar$ before differentiating.

The second is the well-known difficulty of semilocal functionals when bonds are stretched, a specific type of strong correlation. Their failure has been traced to a delocalization error, and related to curvatures of $E$ versus $\mathcal{N}$. The same error shows up more strongly for the 1d kinetic energy. For one particle in two well-separated potentials, half the density ends up in each, leading to a factor of 4 reduction in the kinetic energy relative to the one-well result. However, a model for the double well is

\[ I_{\text{double}}(\epsilon) \approx I_{\text{single}}(\epsilon) + \sum_{j=1}^{2} \Theta(\beta(\epsilon - \epsilon_j)) \]  

(38)

where $\beta$ is now a fixed large number, chosen to mimic the energy splitting between even and odd levels. The local ap-
proximation is much smoother
\[ f_{\text{double}}^{(0)}(\epsilon) = 2f_{\text{single}}^{(0)}(\epsilon) \]  
(39)
and produces a huge overestimate for \( j = 1 \). In fact, the equivalence of TF and WKB approximations breaks down, as there are now four turning points, leading to ambiguities analogous to the symmetry dilemma\(^{29}\) for stretched H\(_2\).

The expansion considered here is the analog of the standard Lieb-Simon (or Thomas-Fermi) expansion as \( Z \to \infty \) of neutral atoms\(^{17,31} \). (Interacting) TF theory yields the dominant term, just as here, but the next two terms (Scott correction and \( Z^{3/2} \) contribution) differ, due to the Coulomb potential, Coulomb interaction, and three dimensions. However, like those terms, one contribution to the leading correction is given by the gradient expansion, and the other involves a difference between a term evaluated at the highest occupied level and at the bottom of the well.

The present work represents a culmination of a series of earlier works\(^{30-32} \) which focused on finding the density as a functional of the potential, but failed to yield systematically improved kinetic energies\(^{33} \). The earlier results should prove useful when understood in the present context. While model results are not directly relevant to realistic calculations, the understanding achieved from previous studies has already had significant practical impact: the derivation of the parameter of the well.

This paper is aimed at the implications of this framework for DFT. Related work focuses on higher-order asymptotics of these sum formulas\(^{31} \). The modern power of asymptotic analysis has recently been demonstrated\(^{37} \) on \( \nu = |x| \), finding contributions that are beyond all finite powers in the semiclassical expansion, with errors of 0.1 mH for just the 2nd level. It is of tremendous interest to apply this machinery in three dimensions and to interacting systems.

ACKNOWLEDGMENTS

I acknowledge funding from NSF (CHE 1856165), and the University of Bristol for a Benjamin Meaker Professorship. I thank Attila Cangi and Raphael Ribeiro, whose earlier work inspired this advance, Michael Berry for instruction in semiclassics, and Chris Hughes for useful discussions.