Exponential decay properties of Wannier functions and related quantities

Lixin He and David Vanderbilt
Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08855-0849
(January 19, 2001)

The spatial decay properties of Wannier functions and related quantities have been investigated using analytical and numerical methods. We find that the form of the decay is a power law times an exponential, with a particular power-law exponent that is universal for each kind of quantity. In one dimension we find an exponent of $-3/4$ for Wannier functions, $-1/2$ for the density matrix and for energy matrix elements, and $-1/2$ or $-3/2$ for different constructions of non-orthonormal Wannier-like functions.

PACS: 71.15.Ap, 71.20.-b, 71.15.-m

A growing interest in localized real-space descriptions of the electronic structure of solids has been motivated by the development of computationally efficient "linear-scaling" algorithms [1,2] and by the desirability of a local real-space mapping of chemical [3,4] and dielectric [5,6] properties. A primary avenue to such a description is the use of Wannier functions [7–9] (WFs), i.e., a set of localized wavefunctions $w_{\mathbf{R}}(\mathbf{r})$ obtained from the Bloch functions $\psi_{\mathbf{k}}(\mathbf{r})$ by a Fourier-like unitary transformation. A closely related approach is to represent the electronic structure in terms of the density matrix $n(\mathbf{r},\mathbf{r}')$. It is thus not surprising to find considerable recent interest in the localization properties of the WFs [3] and of the density matrix [10,11].

In a classic 1959 paper, Kohn proved, for the case of a centrosymmetric crystal in one dimension (1D), that the WFs have an "exponential decay" $w(x) \approx e^{-hx}$, where h is the distance of a branch point from the real axis in the complex-k plane [8]. More precisely,

$$\lim_{x\to\infty} w(x)\, e^{qx} = \begin{cases} 0, & q < h \\ \infty, & q > h \end{cases} . \qquad (1)$$

The density matrix has a similar decay $n(x,x') \approx e^{-h|x-x'|}$. The exponential decay of the WFs has since been proven for the general 1D [12] and single-band 3D [13] cases, and that of the density matrix (more precisely, of the band projection operator) has been proven in general [12]. The energy matrix elements $E(R) = \langle w_R|H|w_0\rangle$, with $w_R(x) = w(x-R)$ and $R = la$ a lattice vector, are also expected to have a similar decay, $E(R) \sim e^{-hR}$.

The purpose of this Letter is to address two questions. First, Eq. (1) allows considerable freedom; in fact, it is consistent with

$$w(x) \approx x^{-\alpha} e^{-hx} \qquad (2)$$

for *any exponent* α, i.e., a decay which could be faster ($\alpha > 0$) or slower ($\alpha < 0$) than pure exponential. Does such a power-law prefactor exist, and if so, what is the exponent α? Second, it has long been understood that relaxation of the orthogonality constraint $\langle w_0|w_R\rangle = \delta_{0,R}$ can give "more localized" Wannier-like functions [14–16]. In what sense are these more localized – a larger h, or a larger α for the same h, or only a smaller prefactor of the tail? We show that the power-law prefactors of Eq. (2) *do* exist, and that the various quantities have a common inverse decay length h but different exponents α. In 1D we find that $\alpha = 3/4$ for usual (orthonormal) WFs, $\alpha = 1/2$ for $n(x,x')$ and $E(R)$, and $\alpha = 1/2$ or $\alpha = 3/2$ for two different constructions of non-orthonormal Wannier-like functions (NWFs). The NWFs of superior decay ($\sim x^{-3/2}e^{-hx}$) can be constructed by a projection method as duals to a set of trial functions. These results may have important implications for the design and implementation of efficient real-space electronic-structure algorithms.

We first review the central results of the pioneering work of Kohn [8], who considers a centrosymmetric potential of period a in 1D. The WFs are constructed as

$$w_n(x-R) = w_{nR}(x) = \frac{a}{2\pi}\int_{-\pi/a}^{\pi/a} e^{-ikR}\psi_{nk}(x)\, dk \qquad (3)$$

with the phases of the Bloch functions ψ_{nk} chosen as in Sec. 6 of Ref. [8]. The exponential decay of the WFs is then governed by the positions of branch points in the "complex band structure" $E_n(k)$ constructed by regarding complex E_n to be a function of complex k via analytic continuation from the real axis [8,9]. Specifically, there is a Riemann sheet $E_n(k)$ for each band n, and the branch points k_n are the points at which the sheets are connected, $E_n(k_n) = E_{n+1}(k_n)$. These are located at

$$k_n = \begin{cases} \pi/a \pm ih_n, & n \text{ even} \\ \pm ih_n, & n \text{ odd} \end{cases} \qquad (4)$$

and at translational image locations $k_n^{(m)} = k_n + 2\pi m/a$ for integer m. $E_n(k)$ and $\psi_n(k)$ are thus analytic functions in the strip $|\mathrm{Im}(k)| < \bar{h}_n$ where $\bar{h}_n{=}\min(h_{n-1}, h_n)$. Kohn's main result [8] is that the decay of the WF for the n'th band is as $w_n(x) \approx e^{-\bar{h}_n|x|}$ in the sense of Eq. (1). In what follows we restrict our attention to the bottom band (n=0), for which $\bar{h}_0 = h_0$ (henceforth just h). The relevant branch point in the upper half-plane is $k_0 = \pi/a + ih$ and the expected Wannier decay is $w(x) \approx e^{-hx}$.

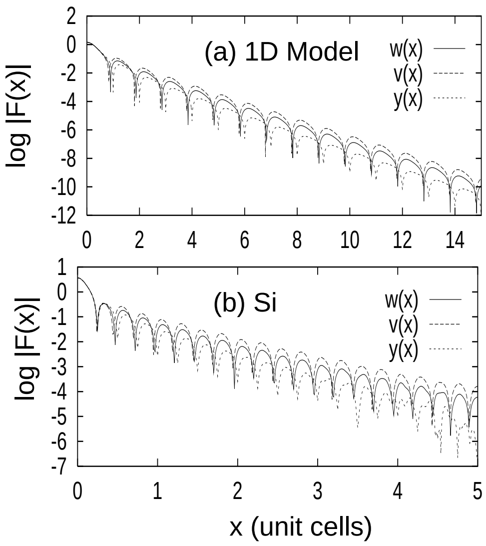

FIG. 1. Decay of normalized WFs $w(x)$ and NWFs $v(x)$ and $y(x)$. (a) 1D model (see text) with b=0.3 and V_0=−10. (b) 3D Si plotted along the [110] direction.

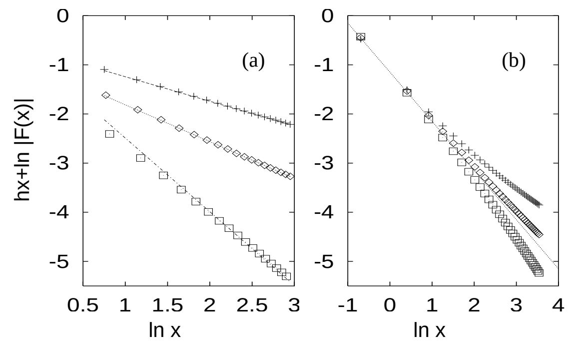

FIG. 2. (a) As in Fig. 1(a) but plotted so that slope reveals exponent $-\alpha$ of Eq. (2). (b) Same for b=0.6 (nearly-free electron case) showing crossover. Pluses, diamonds, and squares represent v, w, and y, respectively.

To confirm this decay numerically, we first choose a simple 1D model Hamiltonian having a periodic potential $U(x) = \sum_m V_{\rm at}(x - ma)$ constructed as a sum of Gaussian "atomic" potentials $V_{\rm at}(x) = (V_0/b\sqrt{\pi})e^{-x^2/b^2}$. Here a is the lattice constant and V_0 and b control the depth and width of $V_{\rm at}$. We choose units such that m=$\hbar$=e=1 and keep a=1 and V_0=−10 fixed while adjusting b to vary the gap. The Bloch functions are computed on a mesh of 200 k points by expanding in 401 plane waves and the WF at $R = 0$ is then constructed according to Eq. (3) using 128-bit arithmetic.

The resulting decay of the WF for b=0.3 is shown as the solid line in Fig. 1(a). In this semilog plot, the approximate linearity of the peaks is consistent with the expected exponential decay, but there is a slight curvature that can be analyzed further. To do so, we first computed the $E_n(k)$ along $\pi/a + i\kappa$ for real κ and defined h to be the value of κ at which $E_0 = E_1$. For b=0.3 we find h=1.28869. In Fig. 2(a) we then plot (diamonds) $hx + \ln|w(x)|$ vs. $\ln(x)$ for each peak of $\ln|w(x)|$. A pure exponential decay $w(x) \approx e^{-hx}$ should yield a horizontal line in such a plot; instead, the data appears linear with a slope of $-3/4$, indicating that

$$w(x) \approx x^{-3/4}\, e^{-hx} \quad . \tag{5}$$

A similar plot (not shown) for

$$E(R) = \langle w_R|H|w_0\rangle = \frac{a}{2\pi}\int dk\, e^{ikR} E(k) \quad . \tag{6}$$

suggests that $E(R)$ shares the same inverse decay length h but has a different power-law exponent,

$$E(R) \approx R^{-3/2}\, e^{-hR} \quad . \tag{7}$$

Naturally h changes if the potential parameter b is varied, but we find that the power-law exponents of $-3/4$ and $-3/2$ do not. It thus appears that these exponents are a universal feature of electron bandstructures in 1D.

In order to gain an analytic understanding of this behavior, we consider first the simpler case of the energy-band Fourier transform $E(R) \leftrightarrow E(k)$. Kohn showed that the expansion of $E(k)$ about $k_0 = \pi/a + ih$ takes the form [8]

$$E(k) = E_0 + \gamma\,(k - k_0)^{1/2} + \cdots \tag{8}$$

with higher terms of order $(k - k_0)^1$, $(k - k_0)^{3/2}$, etc. The form of this expansion arises from the requirement that $E(k)$ come back to itself if k traverses a closed path winding twice around k_0, consistent with the picture of two Riemann sheets touching at k_0.

Now there are well-known mathematical results that relate the behavior of a function near a branch point to the asymptotic decay of its Fourier transform [17]. The following lemma is useful here. Let $f(k)$ be a periodic function $f(k) = f(k + 2\pi/a)$ that has a leading behavior

$$f(k) = f_0 + \gamma\,[i(k - k_0)]^{\beta} \tag{9}$$

when expanded at the branch point $k_0 = \pi/a + ih$. Its Fourier series coefficients are given by

$$F(x) = \int_{C_0} f(k)\, e^{ikx}\, dk \tag{10}$$

at $x = ma$ for integer m. As shown in Fig. 3, the contour C_0 initially lies along the real axis. However, $f(k)\, e^{ikx}$ is invariant under $k \to k + 2\pi/a$, and assuming that no other branch points or poles intervene, the contour can be deformed to become C_1 as shown in Fig. 3. The exponential smallness of e^{ikx} for large x kills the integrals along the horizontal segments, and the dominant contribution to the C_1 integral comes from the vicinity of k_0. Using the contour-integral definition of the Gamma function [18],

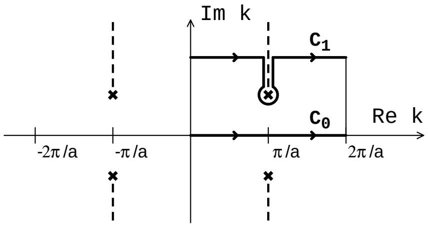

FIG. 3. Branch points (×), cuts (dashed lines), and integration contours (C_0 and C_1) in the complex-k plane.

$$|F(x)| \simeq \gamma\, B(\beta)\, x^{-(1+\beta)}\, e^{-hx} \quad , \tag{11}$$

where $B_\beta = 2\sin(\beta\pi)\,\Gamma(1+\beta)$.

Eq. (11) now allows us to understand the observed behavior of quantities such as $E(R)$ and $w(x)$. For example, since $E(k)$ in Eq. (8) has $\beta = 1/2$, we confirm that $E(R) \approx R^{-\alpha}\, e^{-hR}$ with $\alpha = 1+\beta = 3/2$. Similarly, to understand the decay of $w(x)$, we need to know the behavior of $\psi_k(x)$ regarded as a function of k near the branch point k_0. Once again Kohn [8] provides the needed result $\psi_k \approx (k-k_0)^{-1/4}$. Sure enough, $\beta = -1/4$ gives a decay $w(x+R) \approx R^{-3/4} e^{-hR}$ for small x and $R \gg a$. In other words, $w(x) \approx x^{-3/4} e^{-hx}$ for large x, as obtained numerically from Fig. 2(a).

We can summarize the information about both the k- and x-dependence of $\psi_k(x)$ near the branch point as

$$\psi_k(x) = A_0(x)\, q^{-1/4} + A_1(x)\, q^{1/4} + \cdots , \tag{12}$$

where $q = i(k-k_0)$. (All such terms have powers that are odd-integer multiples of $1/4$, consistent with $\psi_k(x) \to -\psi_k(x)$ when traversing a closed path winding twice around k_0.) $A_0(x)$ and $A_1(x)$ are real functions obeying $A_n(x+a) = e^{ik_0 a}\, A_n(x)$.

The locality of the density matrix (i.e., the band projection operator) is also very important. For example, many linear-scaling algorithms are based on a direct solution for the density matrix [1,2,19]. We can write

$$n(x',x) = \frac{a}{2\pi} \int_{C_0} \psi_{-k}(x')\, \psi_k(x)\, dk \tag{13}$$

where, following Kohn [8], we have substituted $\psi_k^*(x')$ by $\psi_{-k}(x')$ in order that the integrand of Eq. (13) should remain analytic off the real axis. The behavior of $\psi_{-k}(x)$ near the branch point is $\psi_{-k}(x) \approx A_0(-x)\,[i(k-k_0)]^{-1/4}$. The integrand of Eq. (13) then takes the form $\psi_{-k}(x')\psi_k(x) \approx A_0(-x')A_0(x)[i(k-k_0)]^{-1/2}$. Applying Eq. (11) yields $n(0,x) \approx x^{-1/2}\, e^{-hx}$ for large x, and more generally, $n(x',x) \approx (x-x')^{-1/2}\, e^{-h(x-x')}$ for $x \gg x'$. This $\alpha = 1/2$ behavior of the decay has been confirmed from numerical plots (not shown) similar to Fig. 2(a).

We have so far shown that $E(x)$, $w(x)$, and $n(0,x)$ all have a decay of the form $x^{-\alpha}\, e^{-hx}$ with a common h but with different (universal) exponents α_E=3/2, α_w=3/4, and α_n=1/2. The energy matrix elements thus have the fastest decay, and the density matrix the slowest.

One may next ask whether it is possible to find non-orthonormal Wannier-like functions (NWFs) with a faster decay than those of the orthonormal WFs $w(x)$ [14–16]. We explore this question in the context of band-projection methods [20–22]. We find that a naive application of the projection technique actually generates NWFs with a *slower* decay, while a modified "dual construction" approach *does* give improvement as measured by the exponent α.

The basic idea of the projection technique is to start with a trial function $t(x)$ and generate a Wannier-like function $v(x)$ by acting with the band-projection operator $\hat{P} = \sum_k |\psi_k\rangle\langle\psi_k|$, i.e., $|v_R\rangle{=}\hat{P}\,|t_R\rangle$. Here $|t_R\rangle$ corresponds to the translational image $t(x-R)$ of $t(x)$ in cell $R = na$, and similarly for $|v_R\rangle$. The trial functions can be Gaussian functions, atomic or molecular orbitals, etc. The $|v_R\rangle$ are NWFs having overlap $S_{0R} = \langle v_0|v_R\rangle = \langle t_0|\,\hat{P}\,|t_R\rangle$. Numerical investigations on C and Si by Stephan and Drabold indicated that the projected functions $v(x)$ are *not* more localized than the true WFs $w(x)$ [22]. This should not be surprising; introduction of NWFs may give flexibility to generate more localized orbitals, but this flexibility needs to be used to advantage. To do so, we introduce *dual* functions $y(x)$ defined via $|y_0\rangle{=}\sum_R (S^{-1})_{0R}\,|v_R\rangle$, so that $\langle y_0|v_R\rangle = \delta_{0R}$ and also $\langle y_0|t_R\rangle = \delta_{0R}$. This latter equation means that $y(x)$ is orthogonal to the trial function at every site except R=0, suggesting that $y(x)$ may be especially well localized.

Numerical tests of the decay of (normalized versions of) $v(x)$ and $y(x)$ are shown as dashed and dotted curves respectively in Figs. 1(a) and 2(a). The trail function used is a δ-function on the atomic site, but use of other narrow trial functions gives similar results. It clearly appears that $\alpha = 1/2$ and $3/2$ for $v(x)$ and $y(x)$ respectively, to be compared with $\alpha = 3/4$ for $w(x)$. Thus, the simple projected functions $v(x)$ actually have a *slower* decay that the WFs $w(x)$, but the duals $y(x)$ have a much faster decay than either of them.

These results can be explained by the complex analysis of the Bloch-like functions $v_k(x)$ and $y_k(x)$ that are related to $v(x)$ and $y(x)$ in the same way that $\psi_k(x)$ is related to $w(x)$. Defining

$$\eta(k) = \int_{-\infty}^{\infty} \psi_{-k}(x)\, t(x)\, dx \quad , \tag{14}$$

it follows from $|v_k\rangle = |\psi_k\rangle\langle\psi_k|t\rangle$ that $|v_k\rangle = \eta_k\,|\psi_k\rangle$. Also the Fourier transform of $S(0,R)$ can be seen to be $S(k) = \eta^2(k)$, so that $|y_k\rangle = |\psi_k\rangle/\eta(k)$. In the vicinity of k_0 we have

$$\eta(k) = \eta_0\,[i(k-k_0)]^{-1/4} + \eta_1\,[i(k-k_0)]^{1/4} + \cdots \tag{15}$$

where $\eta_n = \int_{-\infty}^{\infty} A_n(-x)\, t(x)\, dx$. Moreover,

$$\begin{aligned} v_k(x) &= \eta_0\, A_0(x)\, q^{-1/2} + \cdots \quad , \\ w_k(x) &= A_0(x)\, q^{-1/4} + \cdots \quad , \\ y_k(x) &= \frac{1}{\eta_0}\, \left\{ A_0(x) + \widetilde{A}_1(x)\, q^{1/2} + \cdots \right\} \quad , \end{aligned} \tag{16}$$

where $q = i(k - k_0)$ and $\widetilde{A}_1(x) = A_1(x) - (\eta_1/\eta_0)\, A_0(x)$. The leading term in $y_k(x)$ gives no singularity, so the real-space decay is determined by the behavior of the next term for which $\alpha = \beta + 1 = 3/2$. To be explicit, we can define $\mathcal{A}_n(x) = A_n(x)\, e^{hx}$ so that $\mathcal{A}$ is anti-periodic, $\mathcal{A}(x + a) = -\mathcal{A}(x)$, and for large x we find

$$
\begin{aligned}
w(x) &\simeq B_{-1/4}\, \mathcal{A}_0(x)\, x^{-3/4}\, e^{-hx} \quad , \\
v(x) &\simeq B_{-1/2}\, \eta_0\, \mathcal{A}_0(x)\, x^{-1/2}\, e^{-hx} \quad , \\
y(x) &\simeq (B_{1/2}/\eta_0)\, \widetilde{\mathcal{A}}_1(x)\, x^{-3/2}\, e^{-hx} \quad .
\end{aligned} \tag{17}
$$

The above conclusions regarding $y(x)$ rely on the absence of zeros of $\eta(k)$ inside the strip $-h < \mathrm{Im}(k) < h$. If such zeros exist, $y_k(x) = \psi_k(x)/\eta_k$ may have new singularities and $y(x)$ will then have poor decay compared to other NWFs. We find that this problem does not arise when using $t(x) = \delta(x)$ or a narrow Gaussian, but can be triggered by use of a too-wide Gaussian for t.

In view of $n(x', x) = \sum_i w_i(x')\, w_i(x)$ it may appear surprising that $n(x', x)$ decays more slowly than $w(x)$ (α_n=1/2 vs. α_w=3/4). The representation of n via w thus has some advantages. Better yet, perhaps, one can represent $n(x', x) = \sum_{ij} S_{ij}\, y_i(x')\, y_j(x)$. Here the slow decay has been transferred to a simple matrix quantity (α_S=1/2) but the NWFs decay very quickly (α_y=3/2).

Is it possible to find a NWF with an even faster decay than $x^{-3/2}\, e^{-hx}$? Yes; define a new NWF $z_k = f(k)\, y_k$ where $f(k)$ is analytic in the strip $|\mathrm{Im}(k)| < h$ and has simple zeros at the branch points $(2n + 1)\pi \pm ih$. The function $f(k) = 1 + \cos(ka)/\cosh(ha)$ is a good candidate [23]. Then the leading singularity of z_k is as $(k - k_0)^{3/2}$, and we expect $z(x) \sim x^{-5/2} e^{-hx}$. We have confirmed numerically that this works. However, since the multiplication by $f(k)$ in k-space corresponds to a convolution in real space, the resulting $z(x)$ is actually *broader* than $y(x)$ or $w(x)$ by almost any other measure (e.g., second moments [3]). Thus, this strategy may be counterproductive in practice.

Before leaving the 1D case, we make two brief comments. First, the extension to the case of non-centrosymmetric potentials in 1D is not difficult, and the results (including values of the α exponents) are unchanged. Second, there is an apparent paradox concerning the nearly-free electron limit. For free electrons the occupied portion of the band gives $w(x) \sim \sin(k_F x)/k_F x$, i.e., $\sim x^{-1}$. One may expect this to go over to $\sim x^{-1}\, e^{-hx}$ in the nearly-free case, but this would be inconsistent with our general result $\alpha = 3/4$. Actually we find there is a crossover behavior, as shown in Fig. 2(b), with α=1 behavior for $x << x_c$ and α=3/4 in the true large-x tail. The crossover distance x_c increases as the gap decreases; as the gap closes, $x_c \to \infty$ and $h \to 0$.

Before concluding, we briefly discuss the 3D case. Here the 6-dimensional space of complex (k_x, k_y, k_z) makes the formal analysis difficult. We have carried out a numerical calculation of WFs and NWFs for Si using an empirical-pseudopotential scheme starting from four bond-centered trial functions. The results are plotted in Fig. 1(b). Plots of $hx + \ln|F(x)|$ vs. $\ln(x)$ (not shown) again show linear behavior, with slopes that appear consistent with the 1D values of $\alpha = 3/4$, $1/2$, and $3/2$ for w, v, and y, respectively. However, in this case we cannot afford to go to very large x values, and we suspect that there may be a crossover to larger α values in the far tails. We leave this as a question for future investigations.

To conclude, we find that in 1D the asymptotic behavior of WFs and related quantities can all be expressed as $x^{-\alpha} e^{-hx}$ with a common h, and with exponents α that take on universal rational values depending on the type of singularity of the relevant function at the branch points in the complex-k plane. It is surprising that this behavior has gone unnoticed since Kohn's seminal 1959 paper. The consequences for linear-scaling calculations, and localized real-space representations of electron structure more generally, remain to be fully explored.

This work was supported by NSF grant DMR-9981193.

[1] G. Galli, Curr. Opin. Sol. State Mater. Sci. **1**, 864 (1996).

[2] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

[3] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[4] P.L. Silvestrelli, N. Marzari, D. Vanderbilt, and M. Parrinello, Solid State Commun. **107**, 7 (1998).

[5] I. Souza, R.M. Martin, N. Marzari, X. Zhao, and D. Vanderbilt, Phys. Rev. B **62**, 15505 (2000).

[6] P.L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999); J. Chem. Phys. **111**, 3572 (1999).

[7] G.H. Wannier, Phys. Rev. **52**, 191 (1937).

[8] W. Kohn, Phys. Rev. **115**, 809 (1959).

[9] E.I. Blount, Solid State Physics **13**, 305 (1962).

[10] S. Goedecker, Phys. Rev. B **58**, 3501 (1998).

[11] S. Ismail-Beigi and T.A. Arias, Phys. Rev. Lett. **82**, 2127 (1999).

[12] J. des Cloizeaux, Phys. Rev. **135**, A685 (1964).

[13] G. Nenciu, Commun. Math. Phys. **91**, 81 (1983).

[14] P.W. Anderson, Phys. Rev. Lett. **21**, 13 (1968).

[15] D.W. Bullett, J. Phys. C **8**, 2695 (1975).

[16] G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992).

[17] See, e.g., F.W.J. Olver, *Asymptotic and Special Functions* (A.K. Peters, Wellesley, MA, 1997), Sec. 3.12.

[18] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series, and Products* (Academic Press, New York, 1980), 8.310.2.

[19] X.-P. Li, R.W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).

[20] J. des Cloizeaux, Phys. Rev. **135**, A698 (1964).

[21] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994); S. Goedecker, J. Comp. Phys. **118**, 261 (1995).

[22] U. Stephan and D.A. Drabold, Phys. Rev. B **57**, 6391 (1998).

[23] Alternatively, the overlap $S(k)$ itself can be used for this purpose via $f(k) = S(k)^{-2}$. This may be useful since one may not know the branch-point locations in advance.