Thomas-Fermi theory of screening of a positive point charge in an insulator

J. Oliva
Department of Physics and Astronomy, California State University at Long Beach, Long Beach, California 90840
(Received 2 September 1986)

The Thomas-Fermi treatment of screening of a point positive charge Ze in a model insulator is developed. In contrast to the corresponding familiar problem for a metal, the density of states, which enters into the Thomas-Fermi analysis, is here appropriate to a model band structure with two bands and a gap. The induced electron density in the conduction band near the point charge, expressed in terms of the electrostatic potential, is used in forming the Thomas-Fermi equation. The partial filling of the conduction band near the point charge gives rise to a Z-dependent screening in addition to that of the usual form due to the background dipole density of the insulator (the latter is treated phenomenologically). This additional screening in turn leads to the asymptotic incomplete screening form for the potential: $\phi(r) \approx Ze/\epsilon_0 r$, where ϵ_0 is Z dependent and thus *not* equal to the bulk dielectric constant associated with the background.

I. INTRODUCTION

The present paper develops a new Thomas-Fermi theoretic approach for screening of a positive point charge within an insulator. We work at temperature T equal to zero. The approach is strictly akin to the familiar Thomas-Fermi[1,2] (TF) approach for a metal:[3] the induced electron density is first expressed in terms of the electrostatic potential $\phi(r)$ by using the principle of minimization of the total energy. This relation is then used in Poisson's equation in order to obtain a differential equation for $\phi(r)$. Now the relation between the induced electron density and $\phi(r)$ involves the density of electron states. For a metal, the usual approach is to use a single parabolic band. For the insulator, in contrast, we use here a simple model band structure with a gap and naturally having a qualitatively different density of states. The TF procedure is then applied using this different density of states.

The approach used here is to be contrasted with a recent popular TF approach for an insulator based on Resta's work.[4-8] In the Resta model[4] there is only a single band. Further, in solving the Poisson's equation, an *ad hoc* boundary condition is used to introduce the insulator character: this is associated with continuity requirements on $\phi(r)$ and its derivative at some radius R (to be determined *through* the continuity requirements), beyond which the electron density is taken to be constant. The Resta model entails a number of qualitatively reasonable features but does not, unlike the present model, build in the important features of conduction band, valence band, and gap.

II. MODEL

In our simple insulator model we use the band structure of Fig. 1. Both valence band and conduction band are parabolic though the detailed form of the valence band turns out to be unimportant here. The valence band has a finite width W; the conduction band extends to infinity. The effective mass of electrons in the conduction band is denoted by m_c. The conduction band is taken to have a degeneracy of d_c. For the present case of a positive point charge impurity the valence-band effective mass and degeneracy do not enter. Finally the energy of the direct gap of the model is denoted by E_g.

Now in the TF modeling we are concerned with electron occupation of the valence and conduction bands. We also have implicitly an ion-core system. At a phenomenological level we assume that the valence band-electron—ion-core system may be polarized in the usual sense of there then being associated with it an induced *dipole* density in the presence of an electric field. Thus for a weak field, our model insulator only responds through a change in the dipole density according, say, to the usual bulk linear dielectric constant. On the other hand, in strong enough fields (as encountered near the point charge) electrons can be present in the conduction band; this, then, gives an additional contribution to the screen-

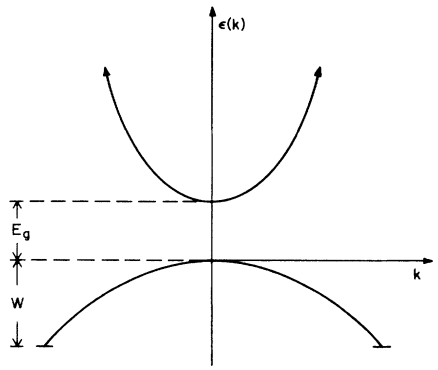

FIG. 1. Qualitative form of band structure in our insulator model: both valence and conduction bands are parabolic; the valence band has finite width W and the conduction band has infinite width.

ing response of the system. Whereas, the dipolar screening associated with the valence band electrons and ion cores is treated phenomenologically, the conduction-electron screening is analyzed within the TF approach.

Let us finally note that the above band structure is to be regarded as a local band structure within the framework of the TF theory: the band structure is shifted locally up or down in energy depending upon the local value of the electrostatic potential. Also note that for purposes of the TF aspect of our modeling we assume a uniform positive neutralizing background.

III. THOMAS-FERMI ANALYSIS

Working at $T=0$, we first obtain the relation between the conduction band occupation and the self-consistent local electrostatic potential $\phi(r)$ associated with the presence of the positive point charge. The dipolar density background component of our model is not treated within TF but will be added into the eventual TF equation for the conduction electrons at a phenomenological level.

We nominally make the standard slow variation assumption within TF. Later for semiquantitative purposes, we will apply the TF results to the rapid variation region near the point impurity; this is typically done in the TF treatment of a point impurity in a metal.

Figure 2 qualitatively indicates the shifting of the local band structure in the self-consistent positive potential field resulting from the impurity (taken at $r=0$). For large r the band structure tends to the unshifted uniform limit. As r decreases there is first a range of r where the local conduction-band edge is between the asymptotic conduction and valence-band edges. In this range the conduction band remains unoccupied. Finally, as r is further decreased, a critical value r_0 is reached where the local conduction band edge is at the level of the asymptotic valence band edge.

Now the asymptotic local valence band is completely filled. We assume that the solid is very large but finite (we can ignore surface effects), and therefore we have an effectively infinite sink of electrons far from the impurity. This assumption is normally made in the corresponding case of a metal. Thus when the local conduction-band edge dips below the asymptotic valence-band edge (i.e., for $r<r_0$) the conduction band becomes partly occupied; further, the conduction band is filled only up to the level of the (essentially unchanged) asymptotic valence-band edge. The valence band remains filled for all r. Note that the above considerations on band filling are based on the principle of minimization of the total energy of the many electron system.

Let us point out that in this model study we assume that the local-band-structure picture described above

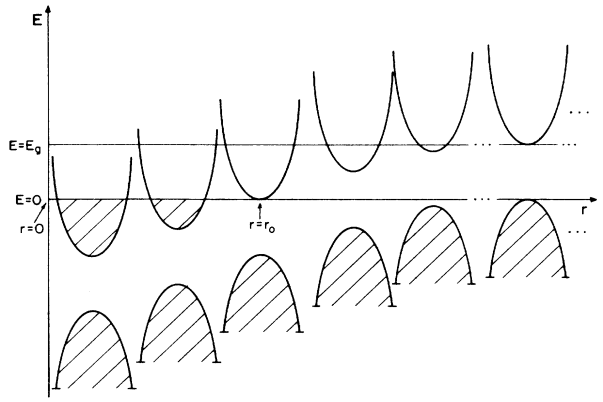

FIG. 2. Qualitative behavior of local band structure as a function of r for case of a positive point charge at $r=0$. At infinity, upper valence-band edge tends to the value $E=0$. At $T=0$ all available states up to $E=0$ are filled for all r. As r decreases $\phi(r)$ increases causing a downward shift of the band structure of magnitude $e\phi(r)$. In the strong field region $r<r_0$ the conduction band is partly occupied. (Shading denotes electron filling of states.)

holds for very large fields. Other breakdown effects are also ignored.[9] Both of these assumptions are normally made in the usual TF treatment of a point charge in a metal.

Now the conduction-band occupation is zero for $r>r_0$ and therefore the total electron density is constant in this region while $\phi(r)\neq 0$. From Fig. 2, the range of values of ϕ for which the conduction band can be partly filled (i.e., for $r<r_0$) is seen to satisfy $e\phi(r)>E_g$.

For $r\leq r_0$, we then have for the conduction-band occupation δn_c:

$$\delta n_c(r)=\int_{E_g}^{e\phi(r)}\rho_c(\epsilon)d\epsilon,\quad e\phi(r)\geq E_g\ , \tag{1}$$

where $\rho_c(\epsilon)$, the conduction-band density of states per unit volume, is given by[10]

$$\rho_c(\epsilon)=\frac{d_c}{2\pi^2}\left[\frac{2m_c}{\hbar^2}\right]^{3/2}(\epsilon-E_g)^{1/2}\quad(\epsilon\geq E_g)\ . \tag{2}$$

(Here the energy ϵ is reckoned from the local value of the valence-band maximum.) Evaluating (1) using (2) we get

$$\delta n_c(r)=\frac{d_c}{3\pi^2}\left[\frac{2m_c}{\hbar^2}\right]^{3/2}[e\phi(r)-E_g]^{3/2},\quad e\phi(r)\geq E_g\ . \tag{3}$$

We may then use Eq. (3) in Poisson's equation to obtain the TF equation for our overall model:

$$\epsilon_x\nabla^2\phi(r)=\begin{cases}\dfrac{4e}{3\pi}d_c\left[\dfrac{2m_c}{\hbar^2}\right]^{3/2}[e\phi(r)-E_g]^{3/2}-4\pi Ze\delta^3(\mathbf{r}), & e\phi(r)\geq E_g \quad (4a)\\ -4\pi Ze\delta^3(\mathbf{r}), & 0<e\phi(r)<E_g\ . \quad (4b)\end{cases}$$

Note that since the valence band is filled throughout, the valence electron density cancels the contribution from the uniform positive background on the right-hand side of (4). The phenomenological dipole background polarization effect comes into Eq. (4) via the quantity ϵ_x, nominally taken to be the bulk dielectric constant. Now Eq. (4a) entails the boundary condition: as $r\to 0$, $r\phi(r)\to Ze/\epsilon_x$. Thus, in our simple model we are building in strict linear dielectric screening by the background for all r, i.e., we ignore breakdown effects in the background, etc.

Note that the TF Eq. (4) for an insulator is naturally rather different in form from that appropriate to a point charge in a metal: For a metal we have[3]

$$\nabla^2\phi(r)=a\{[\epsilon_F+e\phi(r)]^{3/2}-\epsilon_F^{3/2}\}-4\pi Ze\delta^3(\mathbf{r}) , \quad (5)$$

where a is a constant and ϵ_F is the Fermi energy. Whereas, Eq. (5) can be linearized in ϕ, Eq. (4) is intrinsically nonlinearizable in ϕ. Naturally the presence of the gap and the two-band character implied in Eq. (4) also contrast with Eq. (5).

IV. SOLUTION OF THE TF EQUATION

Equation (4) can be written in dimensionless form as

$$\frac{d^2}{dx^2}v(x)=\begin{cases}\dfrac{1}{x^{1/2}}v^{3/2}(x), & v(x)>0 \quad (6a)\\ 0, & v(x)\le 0 . \quad (6b)\end{cases}$$

Here we have removed from the equation the explicit appearance of the δ-function associated with the point positive charge. Rather the effect of the δ-function will now come in through the boundary condition at the impurity. In going from Eq. (4) to Eq. (6) we have the transformation relations

$$r=\beta x \quad (7a)$$

with

$$\beta=\left[\frac{3\pi}{4}\right]^{2/3}\frac{\hbar^2\epsilon_x^{2/3}}{2Z^{1/3}e^2m_cd_c^{2/3}} \quad (7b)$$

and

$$v(x)=\frac{r(x)\phi[r(x)]}{Ze}-\gamma x \quad (7c)$$

with

$$\gamma=\left[\frac{3\pi}{4}\right]^{2/3}\frac{\hbar^2E_g\epsilon_x^{2/3}}{2m_ce^4Z^{4/3}d_c^{2/3}} . \quad (7d)$$

In regard to Eqs. (4) and (6) the $r\approx 0$ and $r>r_0$ behaviors are:

$$\phi(r\approx 0)\approx\frac{Ze}{\epsilon_x r}\Longrightarrow v(0)=\frac{1}{\epsilon_x} , \quad (8a)$$

$$r>r_0:\ \phi(r)=\frac{Ze}{\epsilon_0 r}$$
$$\Longrightarrow x>x_0:\ v(x)=-\gamma(x-x_0) , \quad (8b)$$

where

$$\epsilon_0=\frac{1}{\gamma x_0} . \quad (8c)$$

Equations (8b) follow from Eqs. (4b) and (6b). The condition for continuity of $\phi'(r)$ at $r=r_0$ reads as

$$\phi'(r_0^-)=-\frac{Ze}{\epsilon_0 r_0^2}\Longrightarrow v'(x_0^-)=-\gamma , \quad (8d)$$

where $x_0=r_0/\beta$; x_0 is the point such that $v(x_0)=0$. The constant ϵ_0 in the outer solution relates to the total (incomplete) screening charge around the impurity, due to conduction electrons as well as background; in general, $\epsilon_0\neq\epsilon_x$. Note that the $\phi(r)\sim 1/r$ behavior for $r>r_0$, i.e., incomplete nometallic screening, automatically come out of the TF equation here; for a metal we get complete screening and, in the linear approximation, $\phi(r)\sim e^{-\alpha r}/r$, α a constant.

Equation (6) is the same as the standard TF equation for an atom or ion but the boundary conditions and parameters are of course different here.[11]

The numerical solution procedure for Eqs. (6) with Eqs. (8) is as follows. We first obtain the series form of the solution for small x (see Appendix) and use that to start the numerical integration out from a point x_s near the origin. [The second derivative of $v(x)$ is infinite at the origin.] The initial slope $b_2(<0)$ of $v(x)$ is at our disposal at this point. We integrate out from x_s to the point x_0, where $v(x_0)=0$. Note that if $|b_2|$ is too small, x_0 may not exist: the solution may start to diverge away from the x axis as the integration proceeds. At the point x_0 we ask whether the continuity condition $v'(x_0)=-\gamma$ [Eq. (8d)] is satisfied. We repeat the overall procedure varying b_2 each time until we find a b_2 for which the last equality is satisfied. This then gives the desired solution $v(x)$. [Note that for the range of parameters relevant here and with keeping terms up to $\sim x^6$ in the expansion for $v(x)$ (see Appendix), adequate accuracy was achieved with $x_s\sim 10^{-3}$–10^{-2}].

We consider a brief qualitative application of the above discussion. First note that $v(x)$ is concave upward for $v(x)>0$; also $v(x)$ and $|v'(x)|$ are monotonically decreasing for $v(x)>0$. A typical solution $v(x)$ is shown in Fig. 3. Let us then increase the gap E_g. This implies

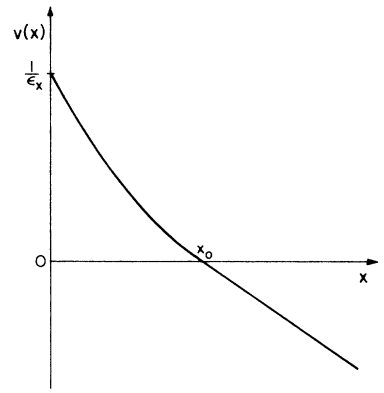

FIG. 3. Qualitative form of solution for $v(x)$. For $v(x)>0$ the curve satisfies Eq. (6a). For $x>x_0$, $v(x)$ is linear; $v'(x)$ is continuous at $x=x_0$ and $v'(x_0)=-\gamma$.

from Eq. (7d) that γ increases. This then implies that the slope of the outer (i.e., $x > x_0$) solution gets more negative, i.e., steeper. Referring to Fig. 3, we see that continuity of $v'(x)$ at x_0 would then imply that x_0 would have to move inward. From Eq. (7a) this implies that r_0, the radius of the free (i.e., conduction) electron cloud decreases. This is of course what we would physically expect for increasing the gap (thereby making it "harder" for electrons to move into the conduction band).

We emphasize that the asymptotic behavior of $\phi(r)$ is $\phi(r) \approx Ze/\epsilon_0 r$ where ϵ_0 is an effective *Z-dependent* dielectric constant which does not equal the ordinary bulk value and which reflects the presence of the conduction electron screening near the point charge (this is intrinsically nonlinear). The deviation of ϵ_0 from ϵ_x is a further point of contrast with respect to the Resta model wherein ϵ_0 is in effect forced to equal ϵ_x.

V. RESULTS AND DISCUSSION

Some results for our model appear in Figs. 4—10. Results for r_0 versus Z for various E_g and for the values $\epsilon_x = 10$, $m_c = 1$ a.u., and $d_c = 1$ appear in Fig. 4. Naturally, for all E_g values $r_0 = 0$ for $Z = 0$. As Z increases the number of induced conduction electrons attracted to the region near the point charge increases implying an increase in r_0. Further, as evident in the plot the smaller the gap, the "easier" it is for the conduction electron cloud to build up as Z is increased.

Results for ϵ_0 versus Z for various E_g and for the values $\epsilon_x = 10$, $m_c = 1$ a.u., and $d_c = 1$ appear in Fig. 5. For $Z = 0$, $\epsilon_0 = \epsilon_x$ for all E_g: For $Z \to 0$, the conduction electron density decreases faster than $\sim Z$, and thus all screening is due to the dipolar background. As Z increases the nonlinear conduction-electron screening builds up: this further screening beyond that of the dipolar background leads to an increase in ϵ_0 above ϵ_x. In addition, it is clear from the graph that decreasing E_g increases ϵ_0; this follows since a smaller gap must give rise to a larger conduction electron occupation and therefore more screening.

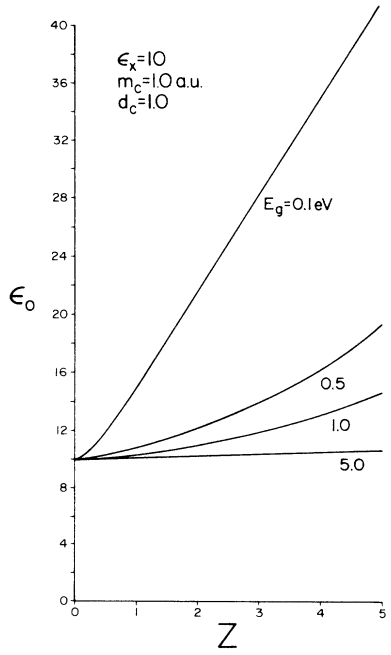

FIG. 5. Results for ϵ_0 vs Z for various E_g within present TF insulator model.

Figure 6 contains results for ϵ_0 versus Z for various ϵ_x

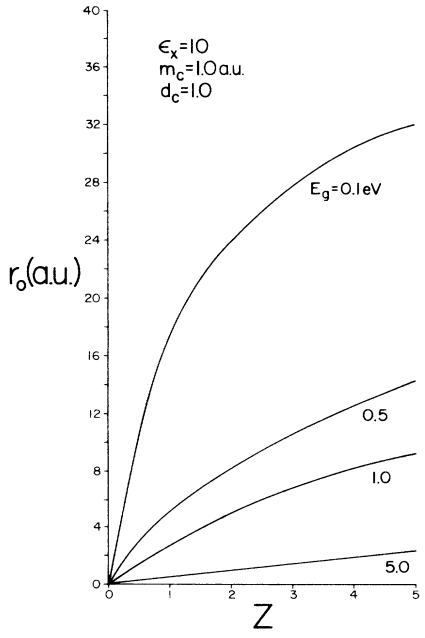

FIG. 4. Results for r_0 vs Z for various E_g within present TF insulator model.

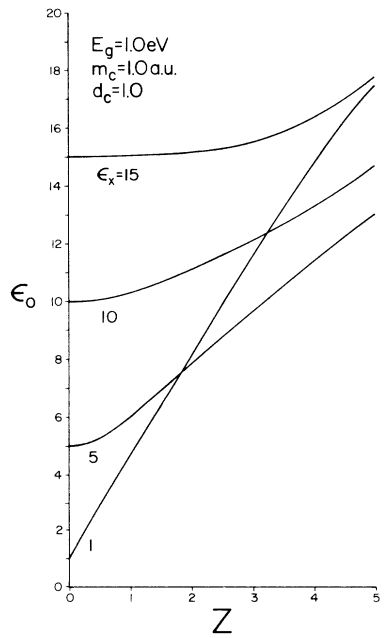

FIG. 6. Results for ϵ_0 vs Z for various ϵ_x within present TF insulator model.

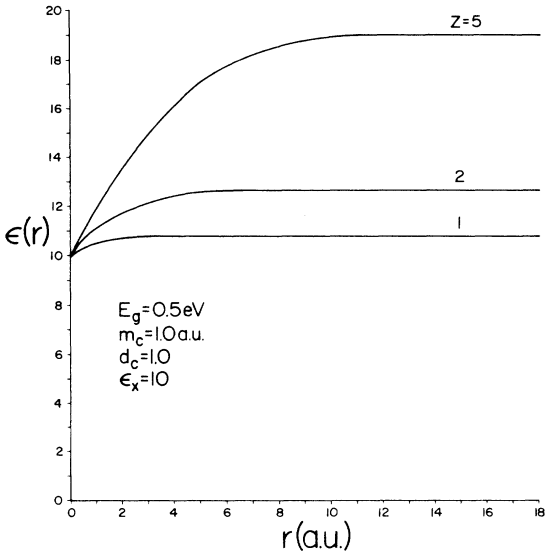

FIG. 7. Results for $\epsilon(r)$ for various Z within present TF insulator model.

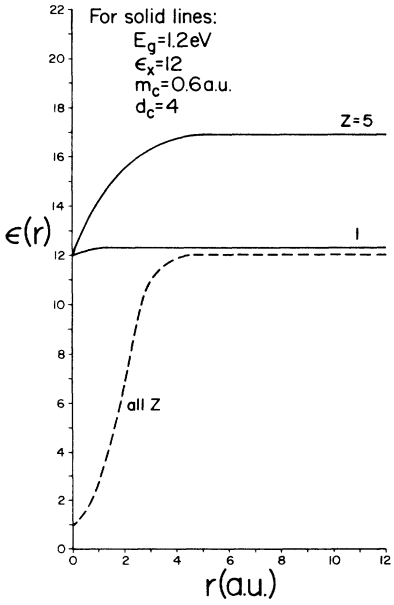

FIG. 8. Results for $\epsilon(r)$ for model Si. Solid lines: results for present model (relevant parameter values indicated on plot). Dashed line is Resta's result.

and for the values $E_g=1$ eV, $m_c=1$ a.u., and $d_c=1$. For all ϵ_x, ϵ_0 tends to ϵ_x as Z tends to 0; this again is a consequence of the "rapid" vanishing of the conduction electron cloud for vanishing Z. As Z increases, the conduction electron screening increases and is manifested as an increase in ϵ_0. We note that the small Z plateau region where $\epsilon_0 \approx \epsilon_x$ is larger (i.e., extends up to a larger value of Z) for larger ϵ_x. This follows since a larger ϵ_x implies a larger dipolar background screening of the point charge (for a given Z) and thus a point-charge field which is effectively smaller and less able to draw in conduction electrons as Z then increases.

Results for the real space dielectric function $\epsilon(r)=Ze/r\phi(r)$ for various Z and for the values $E_g=0.5$ eV, $m_c=1.0$ a.u., $d_c=1.0$, and $\epsilon_x=10$, appear in Fig. 7. We note that $\epsilon(r)$ tends to ϵ_x as r tends to 0 for all Z. This follows since the induced conduction-electron density enclosed in the sphere of radius r tends to 0 as r tends to zero; in our model the dipolar background contribution dominates $\epsilon(r)$ for small r. For large r, $\epsilon(r)$ saturates to ϵ_0; as noted earlier, larger Z implies a larger ϵ_0. The crossover region in r from $\epsilon \approx \epsilon_x$ to $\epsilon \approx \epsilon_0$ gets larger for increasing Z; this is associated with larger conduction electron cloud radius for larger Z.

Finally, calculations were made for a simple model of crystalline Si within the framework of the present approach. In this model we use for the gap the nominal value: $E_g=1.17$ eV and we use for ϵ_x the empirical value of the usual dielectric constant $\epsilon_x=12$. Now the actual conduction bands in Si are rather more complicated than the conduction band of our model. Thus we only very crudely model the parameters m_c and d_c; we take $m_c=0.6$ a.u. (for the actual conduction band minima we have for the longitudinal and transverse effective masses $m_l=0.98$ a.u. and $m_t=0.19$ a.u., respectively) and we take $d_c=4$.

Results for $\epsilon(r)$ for this model of Si together with Resta's results[4] appear in Fig. 8. For our model we have results for both $Z=1$ and $Z=5$. We first note that for r tending to zero, Resta's result tends to the free-electron value unity, whereas, our results tend to ϵ_x. Resta does not include a dipolar background and, in fact, in his model the electron response near $r=0$ is metallic-like. For large r our results tend to $\epsilon_0 \neq \epsilon_x$, whereas Resta's results tend to ϵ_x for all Z. We, of course, again note the significant nonlinear response aspect in our model as Z is increased. For our model, the crossover range in r in going from $\epsilon(0)$ to the far field value of $\epsilon(r)$ is slightly larger (somewhat smaller) for $Z=5(1)$ than in Resta's model.

VI. CONCLUSION

We have discussed a TF treatment of the problem of screening of a positive point charge Ze in an insulator. The approach has paralleled the ordinary treatment of the same problem for a metal. The key difference is in the use in the case of the insulator of a suitable *insulator* band structure (and corresponding density of states), with, in the present model, two bands and a gap. For this case a TF equation with appropriate boundary conditions (tied into the insulator character) was derived. Account of the induced dipole background of the insulator was phenomenologically made. We found that the point charge induces an additional nonlinear conduction band screening charge of finite radius. For very small r this screening cloud has no effect on the dielectric response $\epsilon(r)$ and $\epsilon(0) \rightarrow \epsilon_x$, the background dielectric constant. For very large r we found that $\epsilon(r)$ had the correct $\approx \epsilon_0$ behavior apropos incomplete screening but with an (asymptotic) dielectric constant ϵ_0 which is *different* from

ϵ_x and which is Z dependent; this is due to the presence of the induced conduction-electron cloud. General numerical results were presented as were model results for Si; for Si it was shown that the long-range nonlinear effect was not important for $Z=1$ but clearly important for $Z=5$.

Further work based on this model is planned. Issues to be addressed are finite temperature effects, effects of exchange, the question of negatively charged impurities, more realistic modeling of the conduction band(s) and implications for shallow impurity levels in semiconductors.

ACKNOWLEDGMENT

The author would like to thank Professor Daniel Schechter for many useful discussions during the course of this work.

APPENDIX

We discuss the series solution of the TF Eq. (6) for small x for the nontypical case where the boundary condition at the origin is $v_0 \equiv v(0) \neq 1$ (usually $v_0=1$, e.g., in an atom, or at a charged impurity in a metal). The power series for $v(x)$ is of the form[11]

$$v(x)=v_0+b_1x^{1/2}+b_2x+b_3x^{3/2}+b_4x^2+\ldots.+b_nx^{n/2}+\cdots .$$

For us, $v_0=1/\epsilon_x$. The initial slope b_2 is adjustable here. We find the following expressions for the first several coefficients:

$$b_1=0,\quad b_3=\tfrac{4}{3}v_0^{3/2},\quad b_4=0,\quad b_5=\tfrac{2}{5}v_0^{1/2}b_2,\quad b_6=\tfrac{1}{3}v_0^2 ,$$

$$b_7=\frac{3}{70}\frac{b_2^2}{v_0^{1/2}},\quad b_8=\tfrac{2}{15}v_0b_2,\quad b_9=\tfrac{2}{27}v_0^{5/2}-\frac{b_2^3}{252}\frac{1}{v_0^{3/2}} ,$$

$$b_{10}=\tfrac{1}{175}b_2^2,\quad b_{11}=\frac{31}{1485}v_0^{3/2}b_2+\frac{1}{1056}\frac{1}{v_0^{5/2}}b_2^4 ,$$

$$b_{12}=\frac{4}{405}v_0^3+\frac{4}{1575}\frac{b_2^3}{v_0} .$$

[1]L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).
[2]E. Fermi. Z. Phys. **48**, 73 (1928).
[3]See, e.g., C. Kittel, *Introduction to Solid State Physics,* 4th Ed. (Wiley, New York, 1971), Chap. 8.
[4]R. Resta, Phys. Rev. B **16**, 2717 (1977).
[5]F. Cornolti and R. Resta, Phys. Rev. B **17**, 3239 (1978).
[6]P. Csavinsky and K. R. Brownstein, Phys. Rev. B **24**, 4566 (1981).
[7]P. Csavinsky and K. R. Brownstein, Phys. Rev. B **25**, 1362 (1982).
[8]K. R. Brownstein, Phys. Rev. B **26**, 2253 (1982).
[9]This is done except that we allow the above type of conduction-band filling; this is a type of breakdown effect.
[10]C. Kittel, *Introduction to Solid State Physics,* Ref. 3, Chap. 11.
[11]H. Bethe and R. Jackiw, *Intermediate Quantum Mechanics,* 2nd Ed. (Benjamin, Reading, Mass., 1968), Chap. 5.