# Physics of Tunneling from a Macroscopic Perspective

M.G. Ancona
Naval Research Laboratory
Washington, DC, USA

A. Svizhenko
Silvaco Data Systems, Inc.
Santa Clara, CA, USA

*Abstract*—**Unlike other treatments of electron tunneling, density-gradient theory provides a macroscopic perspective on this quantum phenomenon that is grounded in Newton's quintessentially classical 2nd Law. We demonstrate the meaning and legitimacy of this unusual viewpoint through careful comparisons with non-equilibrium Green's function simulations. The main errors made by the density-gradient approach are shown to arise not from its description of the tunneling but rather from its representation of quantum confinement in the electrodes. A new physical understanding of the *ad hoc* tunneling boundary conditions used in previous work is also exhibited.**

***Keywords: Tunneling; macroscopic; density-gradient;; non-equilibrium Green's function***

## I. INTRODUCTION

It is generally regarded as self-evident that theories of electron tunneling must be explicitly microscopic, *i.e.*, quantum mechanical [1]. The density-gradient (DG) approach to tunneling [2,3] claims to contravene this conventional wisdom by considering electron tunneling as a *macroscopic* phenomenon, *i.e.,* one acting at the level of *populations* of electrons. For such an approach to be viable and more than curve-fitting, it needs to be founded on general physical principles, and when the electrons are incoherent, the only ones known are the classical conservation laws. This means that the DG theory of quantum mechanical tunneling is necessarily a classical field theory. The apparent contradiction in this statement naturally gives rise to questions about the legitimacy and meaning of the DG approach, and these serve as the main motivation for the present paper.

Two aspects of DG theory are most revealing of its macroscopic/classical character, namely, its inclusion of Newtonian inertia in the main governing equation, and its use of boundary conditions to represent rapid variation across material interfaces. Both of these features are examined closely in this work. To do so, we make careful comparisons between predictions of DG theory and quantum mechanics for transport through one-dimensional barriers in steady state. The comparisons clarify a number of foundational issues, including providing new perspective on the roles of electron inertia and of screening and scattering in the electrodes. The analysis also reveals the meaning of the *ad hoc* tunneling boundary conditions used in previous work [2,3]. Lastly, from a practical standpoint, the DG approach to tunneling is shown to provide good accuracy with *no* additional fitting parameters, and with the sources of error reasonably well understood.

## II. DG THEORY

### A. Foundations

DG theory is a classical field theory [4] that is based on the conservation laws of mass/charge and momentum and on electrostatics as applied to one or more electron gas continua. These basic principles are applied both in the interior of electronically conducting materials (resulting in differential equations) and across interfaces between such materials (resulting in boundary conditions). As usual, they are also supplemented by a set of constitutive equations that specify the material response and that make the system determinate. Among these equations, and crucial to DG theory, is the equation of state describing the electron gas response which depends not only on the electron density but also on the density-gradient. The usual form for these dependences is [2]

$$\varepsilon_n(n,\nabla n) = \varepsilon_n^0(n) - \frac{b_n}{2}\frac{\nabla n \cdot \nabla n}{n^2} \qquad (1)$$

The density-dependent term is the ordinary equation of state (usually given a Fermi-Dirac form) and the second term is the DG correction accounting for lowest-order quantum effects [5].

DG theory has been widely and successfully applied to a variety of quantum confinement problems [6]. The application of this same description to quantum tunneling, although first discussed nearly 20 years ago [2], has not been similarly accepted [7]. Two reasons for this are (i) the issues of principle addressed in this paper, and (ii) the fact that scattering tends to be unimportant in tunneling situations so that the DG tunneling equations are a generalization not of diffusion-drift theory but

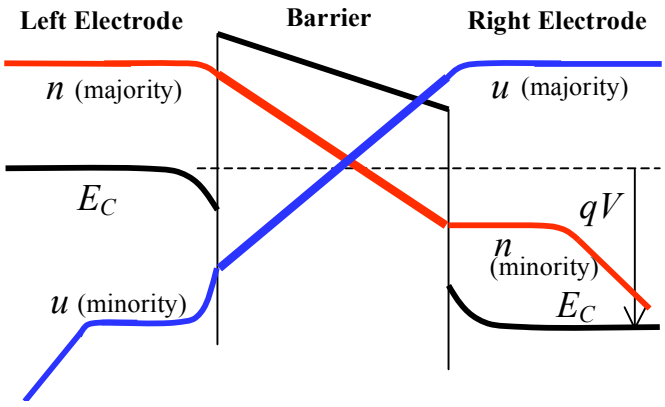


Fig. 1. Schematic depicting the macroscopic view of tunneling.

978-1-4244-1753-7/08/$25.00 ©2008 IEEE

rather of the less-familiar ballistic transport theory. It is hoped that the firmer foundation for DG tunneling theory provided by the present work will lead to its wider use in engineering-oriented device simulation.

*B. Differential Equations*

We consider in this paper only 1-D single barrier tunneling situations in steady-state. Because of the lack of scattering within the barrier the electron gas traveling left-to-right (with density $n$) does not interact with that traveling right-to-left (with density $u$) and so must be treated separately. For such a situation the governing equations of DG theory for $n$ reduce to

$$\left(n\mathrm{v}_n\right)_x = G_n - R_n \qquad \text{(mass/charge balance)} \tag{2}$$

$$\varepsilon_d \psi_{xx} = q\left(n + u - N_D\right) \quad \text{(electrostatics)} \tag{3}$$

$$2qb_n \frac{s_{xx}}{s} + q\psi - q\phi_n^0 - \frac{1}{2} m_n \mathrm{v}_n^2 = q\Phi_n^{DG} \quad \text{(momentum)} \tag{4}$$

where $s \equiv \sqrt{n}$. Inside the barrier there are assumed be no generation/recombination processes, and so (2) can be integrated; this allows $\mathrm{v}_n$ to be eliminated as a variable and constitutes the main simplification of the 1-D situation. Equation (4) is a so-called mechanical energy equation that results from a "Bernoulli" integration of Newton's 2nd law, and $\mathrm{v}_n$ is the (average) electron gas velocity, $\phi_n^0 \equiv \partial(n\varepsilon_n^0)/\partial n$ and $\Phi_n^{DG}$ is a generalized electrochemical potential that is constant when scattering is negligible. From a macroscopic standpoint the most interesting aspect of (4) is that the last term on the left is the macroscopic kinetic energy (neglected in previous work [2,3]), while the first term on the left, which arises from the gradient-dependent equation of state in (1), represents microscopic kinetic energy contributions [7]. The validity of this break-up of the kinetic energy (much like that in a classical context in the gas dynamics of a monatomic gas [8]) is a central question regarding the DG theory of tunneling and is a prime focus of the present paper.

*C. Boundary Conditions*

Because the electrons in the barrier split in two (in 1-D) and have *directionality*, the boundary conditions of DG tunneling also split into "upstream" and "downstream" conditions. The upstream conditions are simply continuity conditions, while the downstream conditions are a crucial part of the macroscopic description of tunneling. In past work, the first author proposed two types of "tunneling recombination velocity conditions" to be applied downstream:

$$\text{TRV1}: \quad s_x = 0 \quad \text{and} \quad J_n = \gamma_n s^2 \tag{5a}$$

$$\text{TRV2}: \quad s = 0 \quad \text{and} \quad J_n = \gamma_g s_x^2 \tag{5b}$$

The justification for these conditions was simply that their use gave qualitatively correct results [2,3]. But their origin and meaning were never clarified, and formulas for $\gamma_n$ and $\gamma_g$ were never provided. Addressing these issues is a second main motivation and result of the present paper.

In understanding the downstream BCs it is important to distinguish between "evanescence" and "tunneling". Evanescence (as we define it) is simply the phenomenon of barrier penetration that allows electrons to traverse the barrier and enter the downstream electrode as "minority" carriers (see Fig. 1). Evanescence produces no current. Tunneling instead is when the process of evanescence is completed by *capture* of the minority carriers by the downstream electrode and subsequently an subsequent *conversion* into "majority" carriers. Clearly, in order to model barrier tunneling our treatment of the downstream electrode — with BCs and possibly also differential equations — must represent this capture physics. Two versions are considered:

Elastic capture: Because the density of final states is generally much greater than the minority carrier density, for a simplest model it seems reasonable to suppose that all carriers that make it to the downstream electrode are immediately captured and converted to ballistic carriers obeying

$$\tfrac{1}{2} m_n \mathrm{v}_n^2 = qV \tag{6}$$

As depicted in Fig. 1 these ballistic carriers will ultimately thermalize and become majority carriers, however, this process is irrelevant to understand the tunneling. The TRV1 conditions in (5a) follow directly from (6) with

$$\gamma_n = \gamma_n^{ideal}\left(\text{ballistic}\right) \equiv \sqrt{2qV/m_n} \tag{7}$$

This derivation not only provides meaning and justification for (5a) but it also supplies an expression for $\gamma_n$ that contains no free parameters.

Inelastic capture: In this process, we suppose that scattering events cause the direct capture/conversion of minority carriers. Assuming the process inside the right electrode is described by

$$\left(n\mathrm{v}_n\right)_x = -\frac{n}{\tau_n} \quad \text{and} \quad \left(u\mathrm{v}_u\right)_x = \frac{n}{\tau_n} \tag{8}$$

where $\tau_n$ is the thermalization lifetime, and assuming $\mathrm{v}_n$ is slowly varying across the downstream interface, then $(8)_1$ can be integrated to obtain applied generalized TRV conditions:

$$\text{GTRV}: \qquad s_x = -\frac{s}{2\gamma_n \tau_n} \quad \text{and} \quad J_n = \gamma_n s^2 \tag{9}$$

When $\tau_n$ goes to infinite, the elastic limit (TRV1) is again obtained. And in the strong scattering limit ($\tau_n \to 0$), it is easily shown that GTRV→TRV2 with $\gamma_g = 4\gamma_n^3 \tau_n^2$. A meaning and justification for the TRV2 conditions is thereby provided.

In order to verify the foregoing theory, in this paper we compare its predictions with quantum mechanics. For specificity, we treat the problem of an SIS diode with varying barrier heights and thicknesses and different levels of doping in the semiconductors. The quantum mechanical simulations are performed using the standard approach of non-equilibrium Green's functions (NEGF) [9] with scattering assumed negligible, the effective mass approximation used, and self-consistent electrostatics included. For the comparisons to be most meaningful, we have strived to make the problems analyzed by DG and by NEGF as close to being identical as possible.

## III. Results and Discussion

### A. Current-Voltage Characteristics

In Figs. 2a-c we show comparisons between DG and NEGF calculations for specific SIS diodes with varying doping in the semiconductor contacts. The DG calculations were made using the TRV1 BCs in (5a) with (7). In general, the agreement is seen to be quite good over many orders of magnitude and with varying bias and barrier height. The agreements must be regarded as especially good when one considers that the DG calculations do not involve any fitting parameters. A similar plot but in log-log form in order to emphasize the low-voltage behavior is shown in Fig. 3. The excellent agreement is seen to hold even down to very low voltage.

As noted earlier, a key assertion of DG theory is that it is possible to split the kinetic energy of the tunneling electron gas into macroscopic and microscopic contributions. The plot in Fig. 4, again with no fitting parameters, explores the contribution of the macroscopic kinetic energy and clearly shows it to be important especially at higher voltages as one would expect.

### B. Solution Profiles

The largest error in Figs. 2a-c is one that is seen to grow systematically with increasing doping density. To understand the origin of this error, in Figs. 5 and 6 we plot the density and potential profiles across a diode with a 3eV, 3nm barrier under a bias of 1V and with varying doping. The DG density profiles (Fig. 5) are seen to become increasingly in error in describing the high-density screening layers, including in not capturing the Friedel oscillations clearly visible in the NEGF results. This error is not due to a lack of fidelity in the tunneling description, but rather results from a known inadequacy of the DG description of quantum *confinement* [10]. This error in turn impacts the exponentially sensitive tunneling currents through its effect on the electrostatic contribution to the tunnel barrier

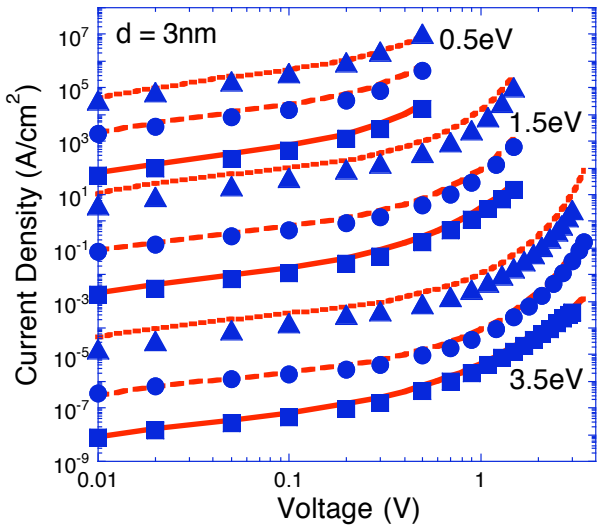


Figure 3. J-V curves as computed by NEGF (points) and DG (dashed lines) for a 3nm barrier with varying barrier heights and electrode dopings of $10^{19}$cm$^{-3}$ (squares), $10^{20}$cm$^{-3}$ (circles) and $10^{21}$cm$^{-3}$ (triangles).

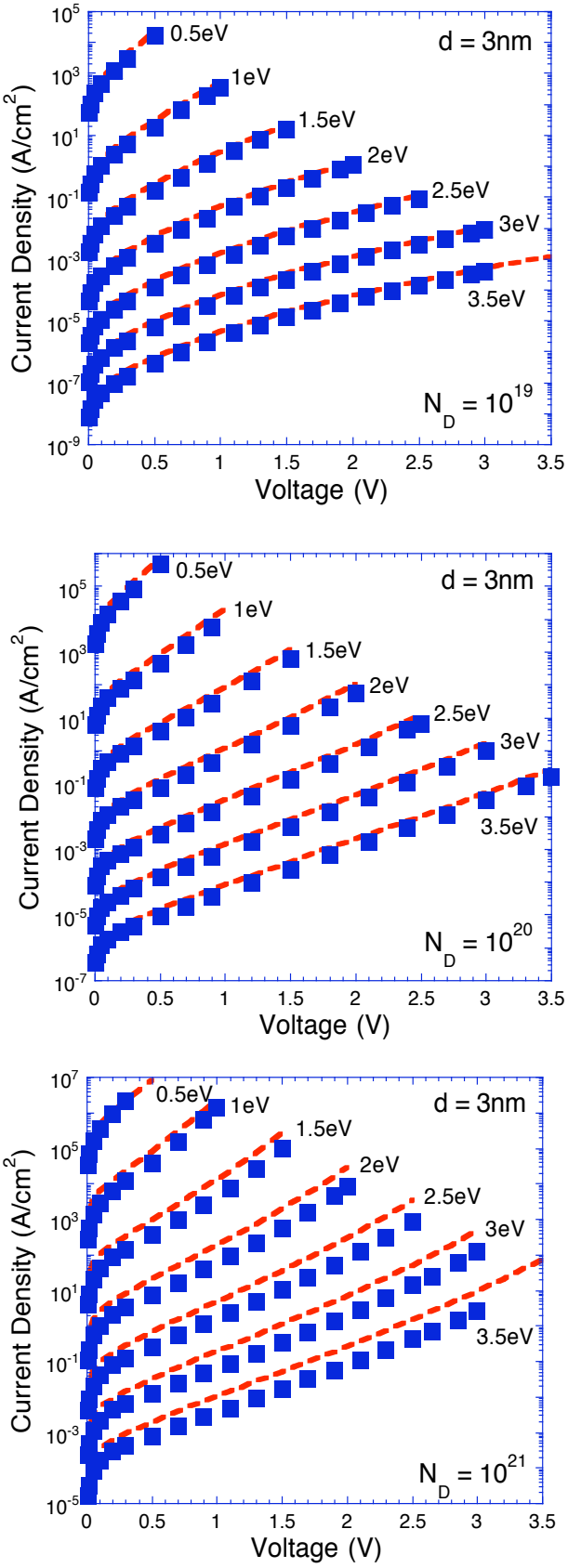


Figure 2. *J-V* characteristics of SIS diodes as computed by NEGF (squares) and DG (dashed lines) for a 3nm barrier with varying barrier heights and electrode dopings of (a) $10^{19}$cm$^{-3}$, (b) $10^{20}$cm$^{-3}$ and (c) $10^{21}$cm$^{-3}$.

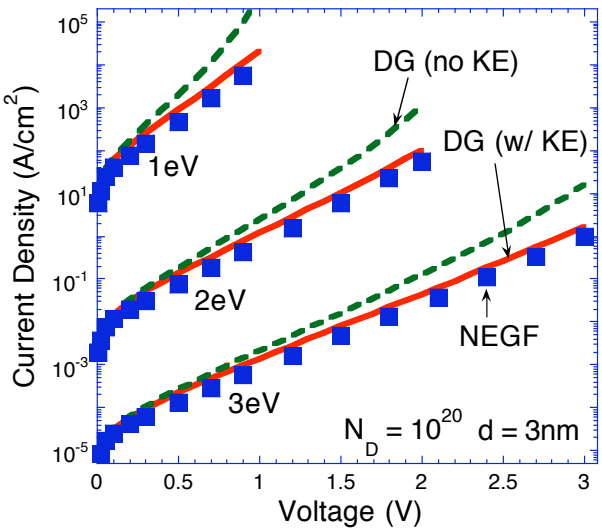


Figure 4. *J-V* curves for different barrier heights as computed by DG theory with and without the macroscopic kinetic energy term included. The squares show the NEGF results.

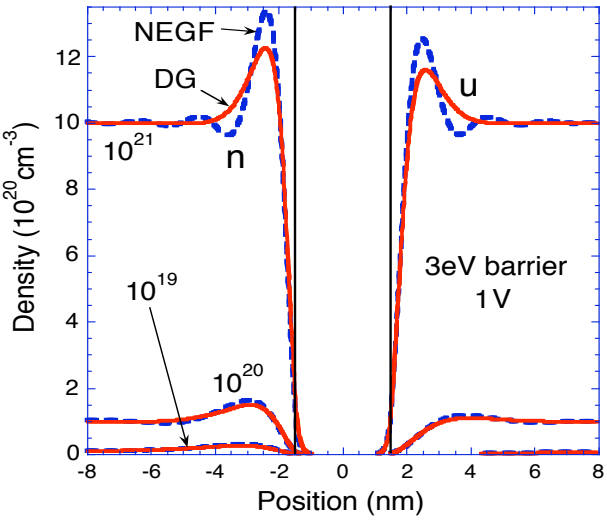


Figure 5. Electron density profiles for three doping levels at a bias of 1V.

("image force" effect) as depicted in Fig. 6.

## IV. Summary and Final Remarks

Through careful comparisons between DG tunneling theory and NEGF as applied to SIS diodes, in this paper we have clarified the meaning and strengthened the legitimacy of the DG approach. Special attention has been given to two aspects of the DG approach that are most associated with its macroscopic nature and which had not previously been well understood, namely, the role of the macroscopic kinetic energy, and the form and meaning of the boundary conditions. Excellent agreement of DG with NEGF has been obtained without fitting parameters. The largest error seen is due not to the tunneling description itself but instead arises from the description of the quantum confinement in the semiconductor contacts when the density is high. Overall, the outlook and results of this work are strong support for the view that DG theory can be a useful tool for analyzing many semiconductor devices that involve tunneling.

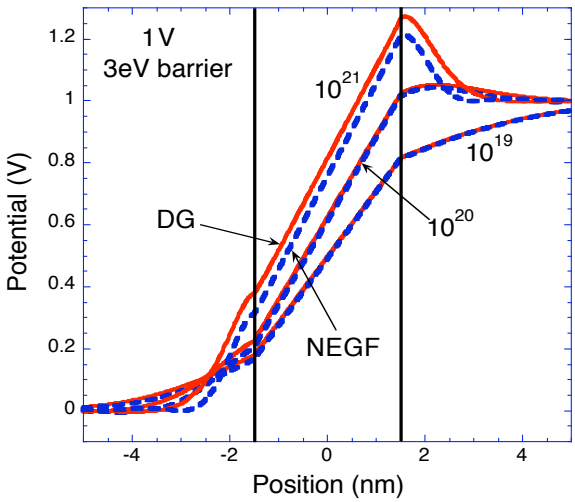


Figure. 6. Electric potential profiles across various SIS diodes and showing the barrier-lowering effect.


## Acknowledgment

MGA thanks Dr. Chagaan Baatar and the Office of Naval Research for funding support.


## References


[1] J.G. Simmons, Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film," *J. Appl. Phys.* **34**, 1793 (1963).

[2] M.G. Ancona, "Macroscopic description of quantum-mechanical tunneling," *Phys. Rev. B* **42**, 1222 (1990).

[3] M.G. Ancona et al, "Density-gradient analysis of MOS tunneling," *IEEE Trans. Elect. Dev.* **47**, 2310 (2000).

[4] C. Truesdell and R.A. Toupin, *The Classical Field Theories*, Handbuch der Physik, Vol. III/1. Springer-Verlag, Berlin (1960).

[5] M.G. Anona and G.J. Iafrate, "Quantum correction to the equation of state of an electron gas in a semiconductor," *Phys. Rev. B* **39**, 9536 (1989).

[6] M.G. Ancona, "Density-gradient theory analysis of electron distributions in heterostructures," *Superlatt. Microstruct.*, 119 (1990).

[7] T. Hohr, A. Schenk, A. Wettstein and W. Fichtner, "On density-gradient modeling of tunneling through insulators," *IEICE Trans. Elect.* **E86-C**, 379 (2003).

[8] W.G. Vincenti and C.H. Kruger, *Introduction to Physical Gas Dynamics*, Wiley, New York (1965).

[9] R. Lake, G. Klimeck, R. C. Bowen, D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.* **81**, 7845 (1997).

[10] M.G. Ancona, "Density-gradint analysis of field-emission from metals," *Phys. Rev. B* **46**, 4874 (1992).