Causal self-energies for NEGF modelling of quantum nanowires

J Barker¹, A Martinez², M Aldegunde² and R Valin²

¹School of Engineering, University of Glasgow, G12 8LT, UK,
²College of Engineering, University of Swansea, UK.

Email: John.Barker@glasgow.ac.uk

Abstract. Electron-phonon scattering is shown to increase dramatically at small nanowire cross-sections so that the transport is not ballistic. Non-ballistic dissipative device modelling requires the full complexity of the non-equilibrium Green Function (NEGF) method. The role of causality in obtaining spectral sum rule-conserving approximations to the electron-phonon self-energies is demonstrated and applications given for wrap-round gate silicon nanowire field effect transistors. Causality violations give erroneous density of states and current densities.

1. Introduction

Non-Equilibrium Green Function (NEGF) modelling is important for studying future limiting scale transistors such as the gate-all-around nanowire field effect transistor (GAA NWFET [1]). Such devices (Fig.1) were once thought to exhibit ballistic transport [2-3] that requires only a simple NEGF formalism. However, from our NEGF simulation studies of phonon-limited mobility shown in Fig. 2, including other data, it is clear that for small NW cross-sections (1-3 nm) the inelastic electron-phonon interactions increase significantly (due primarily to the dependence on the quasi-1D density of states). The transport becomes dissipative with a significant degradation in effective mobility [3-5].

Computing the self-consistent self-energies for inelastic scattering is problematic, particularly for non-local self-energies. Traditionally, the computation may be lessened by first computing $\Gamma(E)$, the imaginary part of the retarded self-energy $\Sigma(E) = \Delta(E) - i\Gamma(E)$ and using a Kramers-Kronig dispersion relation to compute the real part $\Delta(E)$. Often $\Delta(E)$ is neglected: we show here that this strategy may lead to serious errors in the local density of states (DOS) and estimates of dissipation.
2. Theory

The NEGF method computes the $G^<(r_1, r_2; E)$ Green function which determines the carrier density and the current density as functions of position. For a stationary system $G^<$ is determined from two coupled equations: Dyson’s equation for the retarded Green function $G^R$

$$\{E - i\eta - H_0 - \Sigma^R\}^{-1} G^R = 1$$  \hspace{1cm} (1)

(where the causal factor $\eta \rightarrow 0^+$) and the quantum kinetic equation for $G^<$:

$$G^< = G^R \Sigma^A G^R$$  \hspace{1cm} (2)

The total self-energy $\Sigma^\kappa$ ($\kappa = <, R, A = \text{advanced}$) comprises terms due to the contact regions ($S=\text{source}$ and $D=\text{drain}$) and a term due to the electron-phonon interactions ($I$):

$$\Sigma^\kappa = \Sigma^\kappa_s + \Sigma^\kappa_R + \Sigma^\kappa_I$$  \hspace{1cm} (3)

The Ward identities that enforce the local conservation laws (including current conservation) require at least that $\Sigma^\kappa$ satisfies the self-consistent Born approximation (SCBA). For inelastic inter-valley optical phonon scattering with phonon energy $\hbar \omega$ the SCBA has the form (dropping the label $I$):

$$\Sigma^\kappa(E) = C_0 [N_0 G^\kappa(E + \hbar \omega) + (1 + N_0) G^\kappa(E - \hbar \omega)]$$  \hspace{1cm} (4)

$$\Sigma^R(E) = C_0 \{N_0 G^R(E + \hbar \omega) + (1 + N_0) G^R(E - \hbar \omega)\} + \Lambda(E)$$  \hspace{1cm} (5)

$$\Lambda(E) = \frac{P}{\pi} \int d\omega \text{Im} \frac{G^R(E - \omega) - G^R(E + \omega)}{E + \hbar \omega}$$  \hspace{1cm} (6)

The coupling term $\Lambda(E)$ vanishes for elastic scattering. Typically, the self-energies are initialised to zero and the SCBA equations are iterated until convergence, usually ~ 100 iterations.

The spectral density operator and retarded self-energy may be expressed as:

$$A(E) = \delta(E - H) \rightarrow -\frac{1}{\pi} \text{Im} G^R(E); \quad \Sigma^R(E) = \Delta^R(E) - i\Gamma^R(E)$$  \hspace{1cm} (7)

$\Delta^R(E), \Gamma^R(E)$ are the level shift and level broadening operators respectively. It is well-known that $A(E)$ satisfies the Spectral Density Sum Rule and its trace gives $\rho(E)$ the density of states in energy:

$$\int A(p, E) dp = 1 = \int A(r, E) dr$$  \hspace{1cm} (8)

$$\rho(E) = \text{trace} A(E)$$  \hspace{1cm} (9)

$A(p, E)$ and $A(r, E)$ are diagonal matrix elements in the momentum and position representations.

In the time domain, the retarded Green function $G^R(t, t')$ has the “causal” response property:

$$G^R(t, t') = 0 \quad \text{for} \quad t < t'$$  \hspace{1cm} (10)

Using the transformation: $\tau = (t - t')$; $T = \frac{1}{2}(t + t')$ the Green function $G^R(E)$ in the energy domain $E$ is recovered by Fourier transforming over $\tau$ and noting that $G^R$ is independent of $T$ in the stationary limit. The Titchmarsh theorem shows that $G^R(E)$ is a causal function i.e an analytic function in the upper half complex energy plane satisfying a Kramers-Kronig relation or Hilbert Transform. Similarly, $\Sigma^R(E)$ is often assumed to be a causal function satisfying the dispersion relation:

$$\Delta^R(E) = \frac{P}{\pi} \int_{-\infty}^{\infty} \frac{\Gamma^R(E')}{E - E'} dE'$$  \hspace{1cm} (11)

$\Sigma^R(E)$ need not be a simple causal transform if there are bound states or resonances in the electron spectrum. It may be shown [6] that if $\Sigma^R(E)$ or $\Sigma^R(E)G^R(E)$ is a causal transform then the spectral sum rule (8) is automatically satisfied. From eqns (8-9) it follows that any violations of causality in $\Sigma^R(E)$ or $\Sigma^R(E)G^R(E)$ gives rise to errors in the DOS.
3. Results
We first consider a simple 1D model for the causal retarded self-energy for scattering via optical phonon emission using an exaggerated optical deformation potential $C$ to enhance the details. We choose $\Gamma^R(E) = \Gamma_1(E) \propto \rho_0(E - \omega) \propto (E - \omega)^{1/2}$. The inset to Fig. 3 shows $\Gamma_1(E)$ and $\Delta_1(E)$ obtained from the Hilbert transform (11): $\Sigma_1(E) = \Delta_1(E) - i\Gamma_1(E)$ is fully causal by construction: sum rule satisfied; it is the first Born approximation. Fig. 3 also compares the free DOS $\rho_0(E)$ (purple line) with the perturbed DOS $\rho_1(E)$ (red dashed curve) derived from (7) and (9). The effect of neglecting $\Delta_1(E)$ i.e setting $\Sigma_1(E) = -i\Gamma_1(E)$ is shown as the blue curve: this DOS has serious errors due causality violation and sum rule failure. Continuing the iteration leads to the second Born approximation shown in Fig. 4. The fully causal DOS (dashed curve) shows a threshold shift (spectral re-normalisation) compared to the acausal approximation (blue curve).

These features are borne out by full NEGF simulations for a long (38 nm) narrow (2.2 nm wide) 3D GAA NW FET silicon device at low gate voltage using the same exaggerated optical deformation potential $C=930$ eV/nm for just one phonon mode ($g_3$ phonon: see fig. 5; energy 62 meV) but including absorption processes and using a converged SCBA (~ 100 iterations). Fig. 6 shows results for the DOS in the centre of the channel at low field corresponding closely to the conditions of the simple analytical model. The causal DOS (dashed red curve) and acausal DOS (blue curve) are very similar to the simple model of fig. 4: the kink in the causal DOS shows a threshold shift and the sum rule is satisfied in contrast to the acausal approximation.

Finally, Figures (7 - 8) show the energy-resolved current density in the middle of the channel computed from a full 3D self-consistent NEGF simulation for the realistic silicon GAA NW FET device of Figure 1. Acoustic phonon scattering and all the $f$ and $g$ inelastic inter-valley phonon scattering (Figure 5) are included via the Self-Consistent Born Approximation (convergent at around 100 iterations). Deformation potentials and other transport data are from [5]: the optical deformation potential for $g$ phonons is $C = 310$ eV/nm. Results are shown for fully causal self-energies and for comparison we show results for non-causal self-energies obtained by neglecting the real part of the retarded self-energy. The non-causal results show strong violation of the spectral sum rules and significant departures from the causal spectral current density of states especially for the low gate voltage case (Fig 7: $V_G=0.3$ V); the effects persist to high gate voltages as shown in Fig. 8 ($V_G=0.7$ V).
4. Conclusions

Electron-phonon scattering increases with inverse cross-sectional area for silicon nanowire devices with channel diameters < 4 nm. Causally valid inelastic self-energy models are required for the NEGF modelling of nanowires. The spectral density sum rule holds automatically if $\Sigma^R$ or $\Sigma^R G^R$ are strict causal transforms. A simple 1D model for optical phonon emission shows that neglecting the real part of the self-energy leads to: causality violations, failure of the spectral sum rule and erroneous DOS. The quasi-1D nature of transport in narrow nanowires enhances these effects. The results are confirmed for NEGF computations of realistic dissipative silicon nanowire devices where non-causal approximations are shown to give large errors in the current spectral density.

References

[1] H. Yan, et al, *Nature* **470** 240 (2011).
[2] A. Martinez, J.R. Barker, et al, *J. Computational and Theoretical Nanoscience* **5**, 2289 (2008).
[3] S. Jin, Y. J. Park and H.S. Min, *J. Appl. Phys.* **99** 123719 (2006).
[4] M. Luisier, *App. Phys. Lett.* **98** 032111 (2011).
[5] M. Aldegunde, A. Martinez and J. R. Barker, *J. Appl. Phys.* **113** 014501 (2013).
[6] J. R. Barker, unpublished.