Self-energy for semi-infinite electrodes modeled on a cubic tight-binding lattice

Branislav K. Nikolić

PHYS 824: Introduction to Nanophysics, `https://wiki.physics.udel.edu/phys824`
University of Delaware

1 Tight-binding model for two-terminal nanostructures

The so-called real-space modeling of ballistic or disordered nanostructures connected to two (or more) electrodes typically starts from a tight-binding Hamiltonian (TBH) defined on a finite cubic lattice[1] $N \times N_y \times N_z$ (lattice constant is denoted by a)

$$\hat{H} = \sum_{\mathbf{m}} \varepsilon_{\mathbf{m}} |\mathbf{m}\rangle\langle\mathbf{m}| + \sum_{\langle\mathbf{m},\mathbf{n}\rangle} t_{\mathbf{mn}} |\mathbf{m}\rangle\langle\mathbf{n}|. \tag{1}$$

Here $t_{\mathbf{mn}}$ are nearest-neighbor hopping matrix element between s-orbitals $\langle\mathbf{r}|\mathbf{m}\rangle = \psi(\mathbf{r}-\mathbf{m})$ on adjacent atoms located on sites $\mathbf{m}$ of the lattice. The disorder is simulated by taking either the on-site potential (diagonal elements in the Hamiltonian matrix) $\varepsilon_{\mathbf{m}}$ or the hopping (off-diagonal elements) $t_{\mathbf{mn}}$, or both, to be a random variable characterized by some probability distribution. The on-site energies $\varepsilon_{\mathbf{m}}$ correspond to the potential energy, while hopping matrix elements $t_{\mathbf{mn}}$ are the kinetic energy (and depend on the effective mass of an electron). The hopping defines the unit of energy.

The TBH is a matrix (in site-representation) of dimension $\sim (L/a)^d$, which is sparse since nearest-neighbor condition means that most of the elements are zero. It can be considered as a model of either a nanoscale conductor,[2] or a discretized version of a continuous one-particle hamiltonian $\hat{H} = -\hbar^2\nabla^2/2m + U(\mathbf{r})$. In a discretized interpretation the continuous position vector $\mathbf{r}$ is replaced by the position of a point $\mathbf{m}$ on a discrete lattice, and derivatives are approximated by finite differences.

The standard theoretical view of our two-probe measurement circuits is shown on Fig. 1. The sample is placed between two semi-infinite ideal leads. Each lead is modeled by the same clean TBH $\hat{H}_{\mathrm{L}}$, with $\varepsilon_{\mathbf{m}} = 0$ and $t_{\mathbf{mn}} = t_{\mathrm{L}}$, which is defined on an infinite Hilbert space of site states $|\mathbf{m}\rangle$. The coupling between the end layer sites in the lead and corresponding sites in the sample are taken into account through TBH, $\hat{H}_{\mathrm{C}}$, which describes only hopping $t_{\mathbf{mn}} = t_{\mathrm{C}}$ between these sites. The leads are connected at infinity to a particle reservoirs through smooth contacts. Left and right reservoirs (large conductors) are at a constant chemical potential μ_L and μ_R, respectively. Thus they are biased relative to each other by a battery of voltage $V = (\mu_L - \mu_R)/e$. Each reservoir injects the fully thermalized carriers into the lead. The distribution function of electrons to be injected is equilibrium Fermi-Dirac with chemical potential of the reservoir. It is assumed that reservoirs are large enough conductors such that passage of current does not disturb these equilibrium characteristics (i.e., chemical potential can be defined and stays the same as in the reservoir decoupled from the

---

[1]We simplify notation by using $N \equiv N_x$ for the number of sites along the x-axis.

[2]Lattices treated by present day computers cannot contain more than few tens of thousands of atoms. This is the limitation imposed by the available computer memory and computational complexity of matrix operations.
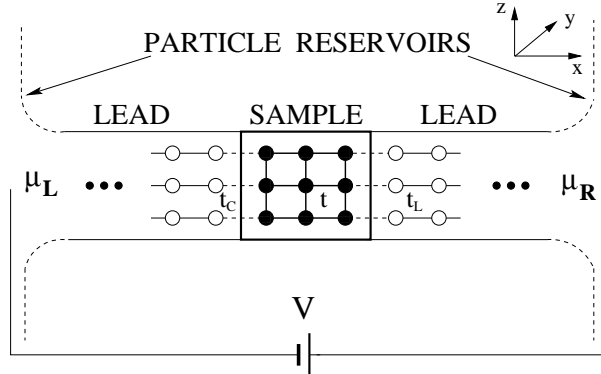
Figure 1: A two dimensional version of our actual 3D model of a two-probe measuring geometry. Each site hosts a single s-orbital which hops to six (or fewer for surface atoms) nearest neighbors. The hopping matrix element is t (within the sample), t_L (within the leads), and t_C (coupling of the sample to the leads). The leads are semi-infinite and connected smoothly at $\pm\infty$ to reservoirs biased by the chemical potential difference $\mu_L - \mu_R = eV$.

conductor). The transport in the central part is phase-coherent. Thus the reservoirs account for the dissipation necessary to establish the steady state. They accept non-equilibrium distribution of electrons from the non-dissipative conductor and provide the thermalization. Even though resistance is related to the dissipation, its value is determined solely by the momentum relaxation processes caused by the scatterers inside the disordered region. However, only the leads at a fixed potential are explicitly taken into account when calculating transport properties. The leads provide the boundary condition for the relevant equations. Since electron leaving the central mesoscopic sample looses the phase-coherence, leads, in a practical way for theoretical calculations, introduce the heuristic construction of the perfect macroscopic reservoirs.

1.1 Green function within the central region

The direct inversion of TBH for the whole system, consisting of semi-infinite leads and the sample,

$$\hat{G}^{r,a}(\mathbf{m}, \mathbf{n}) = (E - \hat{H}(\mathbf{m}, \mathbf{n}) \pm i\eta)^{-1}, \tag{2}$$

would lead into a trouble since $\hat{H}(\mathbf{m}, \mathbf{n})$ is an infinite matrix ($\hat{H} = \hat{H}_S + \hat{H}_L + \hat{H}_C$). The site representation of the Green operator $\hat{G}^{r,a}$ is a Green function matrix

$$\hat{G}^{r,a}(\mathbf{m}, \mathbf{n}) = \langle \mathbf{m} | \hat{G}^{r,a} | \mathbf{n} \rangle, \tag{3}$$

and the matrix of Hamiltonian in this representation is a band diagonal matrix of the bandwidth $2N_y N_z + 1$. The usual method in the literature to avoid this is to use the periodic boundary conditions. However, this would generate a discrete energy spectrum, instead of continuous one of our open system, and is plagued with problems when trying to use standard methods of nonequilibrium statistical mechanics (such as the Kubo formula) which are applicable to systems with continuous energy spectrum.

The correct handling of the leads and openness of the system was initiated by Caroli *et al.*.[3]. Instead of just truncating the matrix (2), which would lead to a conductor with reflecting boundaries instead of open one where electrons can enter and leave the conductor, the leads are taken

---

[3]C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, J. Phys. C: Solid State Phys. **4**, 916 (1971); see also Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992); S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995); and S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, Cambridge, 2005).

into account through the exact "self-energy" terms describing the "interaction" of the finite-size conductor with the leads.

If we consider just the sample and one lead[4] then Green function for this system can be written in the form of a block matrix

$$\hat{G}^r = \begin{pmatrix} \hat{G}^r_L & \hat{G}^r_{L-S} \\ \hat{G}^r_{S-L} & \hat{G}^r_S \end{pmatrix} = \begin{pmatrix} E + i\eta - \hat{H}_L & \hat{H}_C \\ \hat{H}^\dagger_C & E + i\eta + \hat{H}_S \end{pmatrix}^{-1}, \tag{4}$$

where we have shorten the notation by using operator labels without respective matrix indices. The partition above follows from the intrinsic separation of the Hilbert space of states, brought about by the physical separation of lead and the sample in the lattice space. The diagonal blocks are: infinite matrix $\hat{G}^r_L$, connecting the sites in the left lead; and finite $\hat{G}^r_S$ connecting the states on the lattice sites inside the conductor. The off-diagonal blocks, $\hat{G}^r_{L-S}$ and $\hat{G}^r_{S-L}$, connect the states in the lead and the sample. The matrix of the coupling Hamiltonian $\hat{H}_C(\mathbf{m}_L, \mathbf{m}_S) = t_C$ is non-zero only for the adjacent sites in the lead $\mathbf{m}_L$ and the sample $\mathbf{m}_S$. The set of matrix equations for $\hat{G}^r_S$ follows from $\hat{H} \cdot \hat{G}^r = \hat{I}$

$$[E + i\eta - \hat{H}_L] \cdot \hat{G}^r_{L-S} + \hat{H}_C \cdot \hat{G}^r_S = 0, \tag{5}$$

$$\hat{H}^\dagger_C \cdot \hat{G}^r_{L-S} + [E + i\eta - \hat{H}_S] \cdot \hat{G}^r_S = \hat{I}. \tag{6}$$

The Equation (5) can be solved for $\hat{G}_{L-S}$

$$\hat{G}^r_{L-S} = -\hat{g}^r_L \cdot \hat{H}_C \cdot \hat{G}^r_S, \tag{7}$$

$$\hat{g}^r_L = (E + i\eta - \hat{H}_L)^{-1}, \tag{8}$$

where we recognize $\hat{g}^r_L$ as a Green function of a bare semi-infinite lead. This is still an infinite matrix, but can be found exactly as demonstrated in the following Section. Using $\hat{G}^r_{L-S}$ (7) in Eq. (6) we get

$$\hat{G}^r_S = (E - \hat{H}_S - \hat{H}^\dagger_C \cdot \hat{g}^r_L \cdot \hat{H}_C)^{-1}. \tag{9}$$

The final result is a Green function inside a finite-size disordered region which "knows" about the semi-infinite leads, and relevant boundary conditions at infinity they provide, through the "self-energy" function[5]

$$\hat{\Sigma}^r(\mathbf{m}_S, \mathbf{n}_S) = t^2_C \hat{g}^r_L(\mathbf{m}_L, \mathbf{n}_L). \tag{10}$$

Since the self-energy provides a well defined imaginary part (which then "helps" the Green function to become retarded or advanced), we drop the small $i\eta$ in Eq. (9). The self-energy $\hat{\Sigma}^r(\mathbf{m}_S, \mathbf{n}_S)$ is non-zero only between the sites $(\mathbf{m}_S, \mathbf{n}_S)$ on the edge layer of the sample which are adjacent to the sites $(\mathbf{m}_L, \mathbf{n}_L)$ lying on the edge layer of the lead. This follows from the structure of lead-sample coupling matrix $\hat{H}_C$. If the sample is attached to many leads (multi-probe geometry) then one should add the self-energy terms generated by each lead, i.e., in our two-probe case

$$\hat{G}^r_S = (E - \hat{H}_S - \hat{\Sigma}^r)^{-1}, \tag{11}$$

where $\hat{\Sigma}^r = \hat{\Sigma}^r_L + \hat{\Sigma}^r_R$. Advanced functions are obtained in a standard way: $\hat{G}^a = [\hat{G}^r]^\dagger$, and $\hat{\Sigma}^a = [\hat{\Sigma}^r]^\dagger$. In the following Section we give a derivation of a Green function $\hat{g}^r_L(\mathbf{m}_L, \mathbf{n}_L)$ on the end layer of the lead.

---

[4]To clarify notation, we use the subscript L for a general lead and subscript L for the left lead or reservoir in a two-probe geometry.

[5]Analogous terms in Green functions appear when solving the Dyson equation in diagrammatic perturbation theory. Here we use the same labeling while keeping in mind that no approximation is taken for the lead self-energy (as is usually done when discussing self-energies in perturbation theory by summing only a specific set of diagrams).

The self-energies "measuring" the coupling of the sample to the leads can be related to the average time an electron spends inside the sample before escaping into the leads. This can be understood from the following simple arguments. The open system is surrounded by an ideal conducting medium. In that case we cannot talk about eigenstates. Nevertheless, we can formally use an effective Hamiltonian, which is inverted to get the Green function,

$$[\hat{H}_S + \hat{\Sigma}^r]|\alpha_{\text{eff}}\rangle = E_\alpha^{\text{eff}}|\alpha_{\text{eff}}\rangle. \tag{12}$$

This is not a Hermitian operator, and total probability is not conserved. If we write the formal eigenenergy using the eigenvalue of the corresponding isolated system E_α,

$$E_\alpha^{\text{eff}} = E_\alpha - \zeta_\alpha' - i\frac{\zeta_\alpha}{2}, \tag{13}$$

then its imaginary part ζ_α gives the "lifetime" of an electron in state α before escaping into the leads. The probability to stay in the state $|\alpha\rangle$ decays as $|\exp(-iE_\alpha^{\text{eff}}t/\hbar)|^2 = \exp(-2\zeta_\alpha t/\hbar)$, and the escape time into the leads is $\tau_{\text{esc}} = \hbar/2\zeta_\alpha$. The "loss" of electrons into the leads is also illustrated by the following identity

$$\int \nabla \cdot \mathbf{j}(\mathbf{r}) = \frac{1}{\hbar} \int d\mathbf{r}\, d\mathbf{r}'\ \Psi^*(\mathbf{r})\Gamma(\mathbf{r}, \mathbf{r}')\Psi(\mathbf{r}'), \tag{14}$$

where $\hat{\Gamma} = -2\,\text{Im}\,\hat{\Sigma} = i(\hat{\Sigma}^r - \hat{\Sigma}^a)$, and the evolution of wave functions $\Psi(\mathbf{r})$ is determined by the effective Hamiltonian $\hat{H}_S + \hat{\Sigma}^r$.

Even though the eigenstates are not defined in the standard quantum-mechanical sense, one can still use the local density of states (LDOS) given by the imaginary part of the Green function

$$\rho(\mathbf{m}, E) = -\frac{1}{\pi}\text{Im}\,\hat{G}_S^r(\mathbf{m}, \mathbf{m}; E). \tag{15}$$

It turns out that this LDOS is qualitatively similar to the LDOS of 2D and 3D closed system. However, in quasi-1D conductors LDOS computed from (15) is quite different from LDOS

$$\rho(\mathbf{r}, E) = \sum_\alpha |\Psi_\alpha(\mathbf{r})|^2 \delta(E - E_\alpha), \tag{16}$$

defined in terms of exact eigenstates of a closed system.

1.2   The Green function for an isolated semi-infinite ideal lead

In the previous Section we learned that the Green function matrix $\hat{G}_S^r(\mathbf{m}, \mathbf{n})$ (11) at a continuous energy E can be computed numerically by inverting the finite matrix $E - \hat{H}_S - \hat{\Sigma}^r$. This requires to know the matrix elements $\hat{g}^r(\mathbf{m}_B, \mathbf{n}_B)$ of the Green operator for (each) isolated semi-infinite lead between the states $|\mathbf{m}_B\rangle$ located on the sites $\mathbf{m}_B$ at the open boundary of the lead. The lead is modeled by TBH on a rectangular lattice $N_{\text{inf}} \times N_y \times N_z$, where $N_{\text{inf}} \to \infty$ to make the lead semi-infinite. The exact eigenstates of such lead (which has uniform cross section) are separable into a tensor product $|\mathbf{k}\rangle = |k_x\rangle \otimes |k_y, k_z\rangle$. Here $|k_y, k_z\rangle$ are transverse eigenstates (i.e., eigenstates of each isolated transverse layer)

$$|k_y, k_z\rangle = \sqrt{\frac{2}{N_y + 1}}\sqrt{\frac{2}{N_z + 1}} \sum_{n_y=1}^{N_y} \sum_{n_z=1}^{N_z} \sin(k_y n_y a)\, \sin(k_z n_z a)\, |n_y, n_z\rangle, \tag{17}$$

where $|n_y, n_z\rangle$ denotes the orbitals of the arbitrary 2D layer. We choose a hard wall boundary conditions in $\hat{y}$ and $\hat{z}$ direction, so the state $\langle \mathbf{m}|k_y, k_z\rangle$ vanishes at the sites $|\mathbf{m}\rangle$ belonging to the transverse boundary surfaces. This makes the transverse states $|k_y, k_z\rangle$ quantized with eigenvalues (dispersion relation)

$$\varepsilon(k_y, k_z) = 2t_{\mathrm{L}}[\cos(k_y a) + \cos(k_z a)], \tag{18}$$

defined by discrete $k_y(i) = i\pi/(N_y + 1)a$, and $k_z(j) = j\pi/(N_z + 1)a$. Here i runs from 1 to N_y and j runs from 1 to N_z. The longitudinal eigenstates $|k_x\rangle$ (i.e., on the 1D chains) are

$$\langle n_x|k_x\rangle = \sqrt{\frac{2}{N_{\mathrm{inf}}}} \sin(k_x n_x a), \tag{19}$$

with eigenvalues $\varepsilon(k_x) = 2t_{\mathrm{L}} \cos(k_x a)$. This states vanish at the open end on which $n_x = 0$.

The Green function $\hat{g}^r(\mathbf{m}_B, \mathbf{n}_B)$ can be expanded in terms of the exact eigenstates $|\mathbf{k}\rangle$,

$$
\begin{aligned}
\langle \mathbf{m}_B|\hat{g}^r|\mathbf{n}_B\rangle &= \sum_{\mathbf{k}} \frac{\langle \mathbf{m}_B|\mathbf{k}\rangle\langle \mathbf{k}|\mathbf{n}_B\rangle}{E - 2t_{\mathrm{L}}\cos(k_x a) - \varepsilon(k_y, k_z) + i\eta} \\
&= \sum_{k_y, k_z} \langle m_y, m_z|k_y, k_z\rangle\langle k_y, k_z|n_y, n_z\rangle \\
&\quad \times \frac{2}{N_{\mathrm{inf}}} \sum_{k_x} \frac{\sin^2 k_x a}{E - \varepsilon(k_y, k_z) + i\eta - 2t_{\mathrm{L}}\cos(k_x a)},
\end{aligned}
\tag{20}
$$

where only sites at the edge $n_x = 1$ are needed ($|\mathbf{n}_B\rangle \equiv |n_x = 1, n_y, n_z\rangle$). When $N_{\mathrm{inf}} \to \infty$, k_x is continuous and the sum $\sum_{k_x}$ can be replaced by the integral

$$
\begin{aligned}
J(k_y, k_z) &= \frac{2}{N_{\mathrm{inf}}} \sum_{k_x} \frac{\sin^2 k_x a}{E - \varepsilon(k_y, k_z) + i\eta - 2t_{\mathrm{L}}\cos(k_x a)} \\
&= \frac{a}{4\pi t_{\mathrm{L}}} \int_0^{\pi/a} dk_x \frac{2 - e^{2ik_x a} - e^{-2ik_x a}}{(E_J + i\eta)/2t_{\mathrm{L}} - \cos(k_x a)},
\end{aligned}
\tag{21}
$$

where we shorten the notation with $E_J = E - \varepsilon(k_y, k_z)$. This integral can be solved by converting it into a complex integral over the unit circle and finding the residues at the poles lying inside the circle

$$J(k_y, k_z) = -\frac{1}{4i\pi t} \oint_{|w|=1} \frac{1 - w^2}{w^2/2 + 1/2 - Yw}. \tag{22}$$

Here Y denotes the expression $Y = (E_J + i\eta)/2t_{\mathrm{L}}$. The poles of the integrand are at $w_{1,2} = Y \mp \sqrt{Y^2 - 1}$ and have the following properties: (a) $w_1 w_2 = 1$, for any $|Y|$; (b) $|w_1| < 1$, $|w_2| > 1$, for $|Y| > 1$; and (c) $|Y| \le 1$, both poles lie on the unit circle. If (c) is satisfied, then $+i\eta$ ($\eta \to 0^+$) is needed to define the retarded Green function

$$J(k_y, k_z) = -\frac{1}{t_{\mathrm{L}}} \mathrm{Res} \left(\frac{1 - w^2}{(w - w_1)(w - w_2)} \right)_{w=w_1} = \frac{1}{2t_{\mathrm{L}}^2} \left(E_J - i\sqrt{4t_{\mathrm{L}}^2 - E_J^2} \right). \tag{23}$$

If $|Y| > 1$, then

$$J(k_y, k_z) = \frac{1}{2t_{\mathrm{L}}^2} \left(E_J - \mathrm{sgn}\, E_J \sqrt{E_J^2 - 4t_{\mathrm{L}}^2} \right), \tag{24}$$

because one pole is always inside the circle, and the small imaginary term $i\eta$ is not required to define the Green function.

We summarize the results of this section by giving the complete expression for the self-energies introduced by each lead L (in a two-probe case left L and right R)

$$\hat{\Sigma}_L^r(\mathbf{m}, \mathbf{n}) = \frac{2}{N_y + 1} \frac{2}{N_z + 1} \sum_{k_y, k_z} \sin(k_y m_y a) \sin(k_z m_z a)$$
$$\times \frac{t_C^2}{2t_L^2} \left(E_J - i\sqrt{4t_L^2 - E_J^2} \right) \sin(k_y n_y a) \sin(k_z n_z a), \quad (25)$$

for $|E_J| < 2t_L$, and

$$\hat{\Sigma}_L^r(\mathbf{m}, \mathbf{n}) = \frac{2}{N_y + 1} \frac{2}{N_z + 1} \sum_{k_y, k_z} \sin(k_y m_y a) \sin(k_z m_z a)$$
$$\times \frac{t_C^2}{2t_L^2} \left(E_J - \operatorname{sgn} E_J \sqrt{E_J^2 - 4t_L^2} \right) \sin(k_y n_y a) \sin(k_z n_z a), \quad (26)$$

for $|E_J| > 2t_L$. In these expression it is assumed that $\mathbf{n}$ and $\mathbf{m}$ are the sites on the edge layers (first or Nth) of a conductor.

1.3 One-dimensional example: single impurity in a clean wire

To illustrate the power of concepts introduced above, we provide a "back of the envelope" calculation for the single impurity, modeled by the on-site potential ε, in a clean infinite 1D tight-binding chain ($\varepsilon_{\mathbf{m}} = 0$ on all other site). This toy example is illustrated in Fig. 2. Our derivation assumes
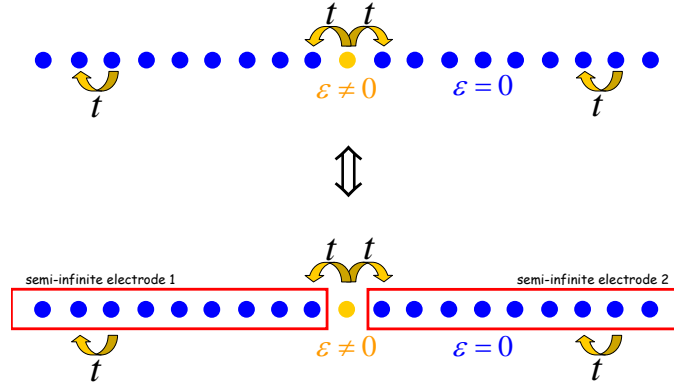

Figure 2: An infinite tight-binding chain with a single impurity atom (top) is equivalent to a "device" where the central region consisting of that atom is attached to two semi-infinite ideal electrodes. The Green function of such central region, as well as the self-energies introduced by the "interaction" with the leads, are 1×1 matrices (i.e., complex numbers).

that impurity is the "sample" from Fig 1 and the rest of the chain are the "leads" with hopping parameter t throughout the system. The Green function of the "sample" is just a number $G^r(E)$ (i.e., 1×1 matrix)

$$G_{\text{in}}^r(E) = [E - \varepsilon - (E - i\sqrt{4t^2 - E^2})]^{-1}, \quad (27)$$

for $|E| < 2t$. This gives the local density of states (15), which is independent of the lattice site, inside the band

$$\rho_{\text{in}}(E) = -\frac{1}{\pi} \operatorname{Im} G_{\text{in}}^r(E) = \frac{1}{\pi} \frac{\sqrt{4t^2 - E^2}}{\varepsilon^2 + 4t^2 - E^2}. \quad (28)$$

For energies outside the band, e.g., $E > 0 > 2t$ the Green function is

$$G_{\text{out}}^r(E) = [E - \varepsilon - (E - \sqrt{E^2 - 4t^2}) + i\eta]^{-1}, \tag{29}$$

where a small imaginary part is added to E because the "self-energy" generated by the "leads" is real. The corresponding LDOS is

$$\rho_{\text{out}}(E) = -\frac{1}{\pi}\text{Im}\,G_{\text{out}}^r(E) = \frac{1}{\pi}\frac{\eta}{(\sqrt{E^2 - 4t^2} - \varepsilon)^2 + \eta^2} \overset{\eta \to 0^+}{\to} \delta(-\varepsilon + \sqrt{E^2 - 4t^2}), \tag{30}$$

where delta function properties lead to the following simplification

$$\delta(-\varepsilon + \sqrt{E^2 - 4t^2}) = \frac{\sqrt{E_p^2 - 4t^2}}{E_p}\delta(E - E_p). \tag{31}$$

The delta function singularity in LDOS appears outside the band of a 1D chain. This is signaling the appearance of a bound state at the energy $E_p = \text{sgn}\,\varepsilon\sqrt{\varepsilon^2 + 4t^2}$. In a clean chain ($\varepsilon = 0$) LDOS is singular at the band edges (Fig. 3).
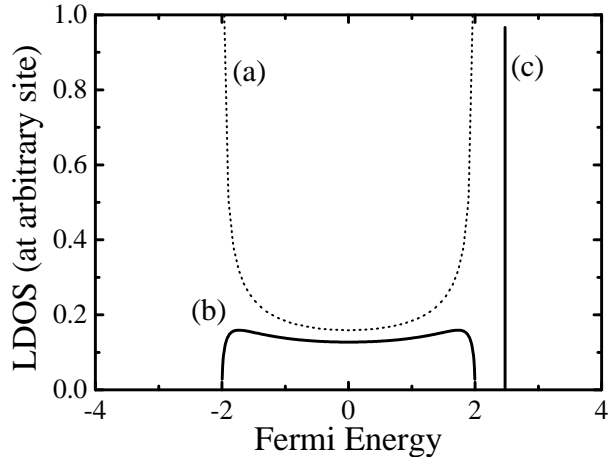

Figure 3: Local density of states (LDOS) at an arbitrary site of a 1D chain, described by a tight-binding Hamiltonian, for: (a) energies inside the band of a clean 1D chain, (b) energies inside the band of a 1D chain with one impurity of on-site energy $\varepsilon = 1$, and (c) outside the band of a 1D chain with one impurity of on-site energy $\varepsilon = 1$.

Thus, the introduction of a single impurity is enough to smooth out the band edge singularities in 1D. These proceeds in accordance with the sum rule: LDOS summed over all sites and energies is constant, meaning that weight is transferred from the continuous spectrum at each site $\mathbf{n}$ into the discrete level LDOS, proportional to the overlap of the discrete state with $|\mathbf{n}\rangle$ (Fig. 3).

7