Three-terminal semiconductor junction thermoelectric devices: improving performance

To cite this article: Jian-Hua Jiang et al 2013 New J. Phys. 15 075021

View the article online for updates and enhancements.

Related content
- Thermoelectric energy harvesting with quantum dots
  Björn Sothmann, Rafael Sánchez and Andrew N Jordan
- Thermoelectric effects in graphene nanostructures
  Philippe Dollfus, Viet Hung Nguyen and Jérôme Saint-Martin
- Tuning phonon properties in thermoelectric materials
  G P Srivastava

Recent citations
- Quantum-dot circuit-QED thermoelectric diodes and transistors
  Jincheng Lu et al
- Rafael Sánchez
- Near-field three-terminal thermoelectric heat engine
  Jian-Hua Jiang and Yoseph Imry

IOP eBooks
Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.
Start exploring the collection - download the first chapter of every title for free.
Three-terminal semiconductor junction thermoelectric devices: improving performance

Jian-Hua Jiang\textsuperscript{1,2,3}, Ora Entin-Wohlman\textsuperscript{4,5} and Yoseph Imry\textsuperscript{1}

\textsuperscript{1}Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel
\textsuperscript{2}Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, People’s Republic of China
\textsuperscript{3}Department of Physics, University of Toronto, 60 Saint George Street, Toronto, ON M5S 1A7, Canada
\textsuperscript{4}Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel
\textsuperscript{5}Department of Physics and the Ilse Katz Center for Meso- and Nano-Scale Science and Technology, Ben Gurion University, Beer Sheva 84105, Israel

\textit{New Journal of Physics} \textbf{15} (2013) 075021 (15pp)
Received 20 February 2013
Published 23 July 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/7/075021

\textbf{Abstract.} A three-terminal thermoelectric device based on a p–i–n semiconductor junction is proposed, where the intrinsic region is mounted onto what is typically a bosonic thermal terminal. Remarkably, the figure of merit of the device is governed also by the energy distribution of the bosons participating in the transport processes, in addition to the electronic. An enhanced figure of merit can be obtained when the relevant distribution is narrow and the electron–boson coupling is strong (such as for optical phonons). We study the conditions for which the figure of merit of the three-terminal junction can be greater than those of the usual thermoelectrical devices made of the same material. A possible setup with a high figure of merit, based on Bi\textsubscript{2}Te\textsubscript{3}/Si superlattices, is proposed.

\* Patent application pending.
1. Introduction

Thermoelectric energy conversion [1, 2] has for decades stimulated considerable research on fundamentals and applications. For a long time, people strove to find good thermoelectric materials with high thermal to electrical energy conversion efficiency. It was found that the optimal efficiency of a thermoelectric device in the linear-response regime is [1, 2]

\[
\eta_{\text{opt}} = \eta_c \frac{\sqrt{1 + ZT} - 1}{\sqrt{1 + ZT} + 1} \quad (1)
\]

with \(\eta_c\) being the Carnot efficiency. The optimal efficiency \(\eta_{\text{opt}}\) is an increasing function of the figure of merit \(ZT\). However, \(ZT = T\sigma S^2/(\kappa_e + \kappa_p)\) is limited by several competing transport coefficients, the conductivity \(\sigma\), Seebeck coefficient \(S\) and electronic (phononic) thermal conductivity \(\kappa_e (\kappa_p)\), making high values of \(ZT\) hard to achieve [3, 4]. Mahan and Sofo (henceforth ‘M–S’) [5] proposed analyzing and achieving high values of \(ZT\) by separating it into two factors: (A) \(T\sigma S^2/\kappa_e\) and (B) \(\kappa_e/(\kappa_e + \kappa_p)\). By recognizing that the electronic transport quantities, \(S\) and \(\kappa_e/\sigma\), are related to the mean and the variance of \(E - \mu\) (i.e. the heat transferred by an electron at energy \(E\) with \(\mu\) being the equilibrium value of the chemical potential [6]), over the transport distribution function, namely the energy-dependent conductivity \(\sigma(E)\), they were able to obtain [5]

\[
\frac{T\sigma S^2}{\kappa_e} = \frac{(E - \mu)^2}{(E - \mu)^2 - (E - \mu)^2} \quad (2)
\]

For a quantity \(O\) (which is a function of \(E\)) \(\langle O \rangle = \int dE \sigma(E)(-\partial f_0/\partial E)O(E)/\int dE(-\partial f_0/\partial E)\sigma(E)\), with \(f_0\) being the equilibrium Fermi distribution function. According to M–S [5], high \(ZT\) values can be achieved (i) by increasing the factor (A) through decreasing the variance of \(E - \mu\) via a sharp structure in \(\sigma(E)\), away from the chemical potential \(\mu\), and (ii) by reducing the ratio \(\kappa_p/\kappa_e\). Following this, there were many proposals to achieve effectively narrow electronic bands, especially in nanostructures with transmission resonances and where the enhanced scattering of phonons at interfaces also reduces the phononic heat conductivity \(\kappa_p\) [7–10]. However, narrow electronic bands do not necessarily lead to high \(ZT\) values. Specifically, when \(\kappa_e \ll \kappa_p\), \(ZT\) does not increase via reducing the variance of \(E - \mu\) when it is already limited by \(\kappa_p\), or if \(\sigma\) is concurrently decreased, it has been argued [11] to
even harm $ZT$ and reduce the power factor $\sigma S^2$. The best figure of merit can only be obtained by considering the competition of all those factors [11].

One should be aware that the energy-dependent conductivity $\sigma(E)$ is well defined only for elastic processes. In the direct generalization of the M–S results to include inelastic processes, $E$ in equation (2) effectively becomes the average, $\bar{E} = (E_i + E_f)/2$, of the initial and final energies ($E_i$ and $E_f$, respectively) of the transferred electron [12]. Non-trivial aspects of the inelastic processes are revealed in the ‘three-terminal thermoelectric devices’ proposed very recently [13–17]. By ‘three terminal’ we mean a setup with an additional thermal terminal supplying bosons (e.g. phonons, electron–hole excitations) involved in the inelastic processes, besides the two electronic terminals. In such devices, in addition to the normal thermoelectric effect in the two electronic terminals, there can be a thermoelastic effect due to the energy transfer between the thermal terminal and the electronic ones. Physically, this is because the energy exchange between the electronic and bosonic systems induces an electric current, or vice versa. The optimal efficiency of such a three-terminal thermoelectrical device in the linear-response regime was found in [16] to be the same as given by equation (1) but with $ZT$ replaced by the three-terminal figure of merit $\bar{Z}T$. By optimizing the efficiency of, say, a refrigerator working at equal temperatures of the two electronic reservoirs, it is found that the three-terminal figure of merit is

$$\bar{Z}T = \frac{\langle \omega \rangle^2}{\langle \omega^2 \rangle (1 + G_{el}/G_{in}) (1 + K_{pp}/K_{pe}) - \langle \omega \rangle^2}.$$  

The average in equation (3) is taken with respect to the conductance of each inelastic transport channel. Specifically for a quantity as a function of the initial and final energies $\mathcal{O}(E_i, E_f)$

$$\langle \mathcal{O} \rangle = \frac{\int dE_i dE_f g_m(E_i, E_f) \mathcal{O}(E_i, E_f)}{\int dE_i dE_f g_m(E_i, E_f)}$$  

with $g_m(E_i, E_f)$ being the conductance of the inelastic channel with given initial and final energies, and $\omega = E_f - E_i$ is the energy of the boson (also equal to the energy change of the carrier) in each inelastic process. This generalizes the results of [16] where only a single inelastic transport channel has been considered. In equation (3) $G_{el}$ ($G_{in}$) is the total elastic (inelastic) conductance, $K_{pp}$ is the purely boson-mediated thermal conductance between the thermal terminal and the other two terminals, and $K_{pe} = e^{-2}G_{in} \langle \omega^2 \rangle$ is the thermal conductance characterizing the heat transfer between the bosons and the electrons. This is the generalization of the theory of M–S to the three-terminal case where the principal quantity is now the energy change $\omega$. A direct consequence is that there is no cancelation of the electron and hole contributions to the three-terminal thermopower. We find that a high three-terminal figure of merit $\bar{Z}T$ requires (i) the dominance of the inelastic transport $G_{in} \gg G_{el}$, (ii) a small variance of the energy change $\langle \omega^2 \rangle - \langle \omega \rangle^2 \ll \langle \omega \rangle^2$ and (iii) a large ratio of $K_{pe}/K_{pp}$ that can be realized when $G_{in}$ and $\langle \omega^2 \rangle$ are large or $K_{pp}$ is small. Small $K_{pp}$ values should be achievable by e.g. engineering the interfaces between the central system and the two electronic terminals. Remarkably, the purely electronic heat conductance, $K_{ee}$, does not appear and does not need to be small!

In addition to the pursuit of a narrow distribution of $\sigma(E)$ in the M–S proposal, the three-terminal figure of merit may also benefit from a ‘selection’ of the energy change $\omega$ either via the electronic structure or via the bosonic spectrum so that the variance of $\omega$ can be small. This

---

6 There are some similar ideas and experimental studies in the literature, e.g. [18].
can be achieved also by a small bandwidth of the bosons involved in the inelastic transport. The merits of the three-terminal configuration are several. (i) There may be no restriction on the effective electronic bandwidth (or other parameters required for a small variance of $\sigma(E)$) as the electronic thermal conductivity $\kappa_e$ does not appear in the three-terminal figure of merit. (ii) Smaller effective boson bandwidths usually make the bosonic thermal conductance $K_{pp}$ smaller, which improves $\tilde{Z}_T$. (iii) In general if, e.g., due to momentum or energy conservation (a ‘selection’), only the bosons in a small energy range are involved in the transitions, the effective bandwidth can also be small. As a possible candidate, optical phonons have small bandwidths (see [19]) and their coupling with carriers is relatively strong. For the p–n configurations discussed here, this necessitates an electronic band gap smaller than or of the order of the phonon frequencies. This can happen e.g., in solid solutions of the HgCdTe family [19]. Further examples, such as superlattices, are mentioned below. For acoustic phonons the coupling to the carriers is usually stronger at large wavevectors/frequencies [20] (e.g., around the Debye frequency) where the density of states of phonons is also large. In [15, 17], the Coulomb interaction between the quantum-dot system and the lead(s) plays the same role as bosons to induce the inelastic transport.

The choice of the thermal terminal is very important. It should provide bosonic excitations with an energy matching the required electronic one, $E_g$. When $k_B T \ll E_g$, and for up going transitions the product of the effective $E_g$ and the exponentially small boson population has to be optimized (as done for example in [5]). In addition, the Debye energy is often not high enough, and multiple phonon processes are quite weak. This can be remedied by using optical phonons or by using a material with $\omega_D \gtrsim E_g$ with $\omega_D$ being the Debye energy (e.g., diamond, whose $\omega_D = 190$ meV) for the thermal terminal. An electronic thermal bath where its electron–hole pair excitations or plasmons interact with the electrons in the intrinsic region is another possibility [15]. The direct tunneling thermal conductance between the thermal terminal and the other two terminals can be made exponentially small by controlling the geometry of the contact. Here we suggest and advocate optical phonons. We will mainly consider the optical phonon bath without implying that it is the only possibility.

We thus propose a three-terminal device based on the p–i–n junction where the intrinsic region is contacted with e.g., a phonon source or a thermal terminal, as above. In section 2, we present the device structure and show how the proposal works for semiconductors and some of their superlattices where the band gap can be smaller than the phonon energy. We estimate the figure of merit and find that such a device can have better performance than the usual two-terminal device made of the same material. We conclude this study in section 3. The discussions throughout this paper are focused on the linear-response regime. Nevertheless, if the system is not far away from that regime, such treatment may still offer useful information [2].

2. Three-terminal p–i–n junction thermoelectric device

2.1. Device concept and structure

In the following discussions we specifically consider a p–i–n junction made of ‘extremely narrow-gap semiconductors’. The structure is depicted in figure 1(a), for the case of converting thermal to electrical energy. It can be viewed as an analogue of the p–i–n photodiode, where photons are replaced by phonons, i.e., the phonon-assisted inter-band transitions lead to current generation in the junction. The device can also be used to e.g., cool the thermal terminal.
Figure 1. (a) Schematic illustration of a possible three-terminal p–i–n junction thermoelectric device. The phonon source acts as the thermal terminal. The two electronic terminals are the p- and n- doped regions, respectively. As an example, we illustrate the situation where the device converts thermal energy from the thermal terminal to electrical energy. The arrows denote the direction of the electric current. (The device can also cool the thermal terminal by consuming the electrical energy, not directly indicated in the figure). (b) Band structure of the p–i–n junction. The dotted line is the chemical potential at equilibrium. The red arrow labels the phonon-assisted inter-band transition that generates electrons and holes. When drifting with the built-in electric field, the generated non-equilibrium carriers lead to current flow across the junction. (c) Schematic illustration of another possible setup of a three-terminal p–i–n junction thermoelectric device. The thermal terminal (labeled as ‘T’ in the figure) is a thermal finger, which can be either a phonon reservoir or an electronic one. In the latter case the electron–hole pair excitations or plasmons are coupled with the electrons in the intrinsic region via Coulomb interaction.

via the electric current between the electronic terminals. We focus on the situation where single-phonon-assisted inter-band transitions are allowed. (The electronic band gap is hence required to be smaller than the phonon energy.) Such processes may play a significant role in the
transport across the junction, whereas the transition rates of multi-phonon processes are much smaller. In semiconductors, the optical phonon energy is usually in the range of 20–100 meV. There are several candidates with such small band gaps: (i) Gapless semiconductors resulting from accidental band degeneracy in solid solutions, such as $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ and $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ [21]. At a certain mole fraction $x$ the band gap closes, around which it can be very small. (ii) Gapless semiconductors originating from band inversion, such as $\text{HgTe}$ and $\text{HgSe}$ [21]. The band gap can be tuned via the quantum-well effect in the superlattices composed of a gapless semiconductor and a normal semiconductor without band inversion [22]. An example is the $\text{HgTe}/\text{CdTe}$ superlattices with a tunable band gap [23]. (iii) Multilayers (superlattices) of ‘topological’ insulators with ordinary insulator layers sandwiched between the topological ones [24]. For example, in $\text{Bi}_2\text{Te}_3/\text{Si}$ superlattices the energy gap can be tuned by varying the thickness of the topological and ordinary layers [24]. A significant merit of superlattices is that the lattice thermal conductivity along the growth direction can be much smaller than both of the two bulk constituents. For example, the $\text{Si}/\text{Ge}$ superlattices have a thermal conductivity about two orders of magnitude smaller than the bulk values [25]. We also point out that the same idea can be applied to devices where the role of optical phonons is played by other bosons. If the energy of such bosons is higher, the requirement for a small band gap can be softened.

For simplicity, we consider a linear junction where the conduction band edge varies linearly in the intrinsic region with the coordinate along the junction $z$ from $z = -L/2$ to $L/2$ as

$$E_c(z) = E_g(1/2 - az/L),$$

with $E_g$ being the band gap [26]. We set $0 < a < 1$ so that the p-doped (n-doped) region is at the left (right) side of the junction (see figure 1(b)). These electronic terminals can have temperatures different from that of the thermal terminal. It is favorable to bend the two electronic terminals away from the thermal one so that they will be better thermally isolated from the thermal terminal and from each other (see figure 1(a)). Another possibility is a ‘thermal finger’ for the boson bath (figure 1(c)), which can be well isolated from the electronic leads.

2.2. ‘Ideal’ figure of merit

We start by considering only phonon-assisted transport, ignoring phononic thermal conduction and normal diode transport. The phonon-assisted inter-band transitions generate current flow in the junction. In the linear-response regime, the thermoelectric transport equations are written as [16]

$$
\begin{pmatrix}
I_e \\
I_{Q}^e \\
I_{Q}^{ec}
\end{pmatrix} =
\begin{pmatrix}
G_{in} & L_1 & L_2 \\
L_1 & K_e^0 & L_3 \\
L_2 & L_3 & K_{pe}
\end{pmatrix}
\begin{pmatrix}
\frac{\delta \mu}{e} \\
\frac{\delta T}{e} \\
\Delta T/T
\end{pmatrix}.
$$

Here $I$ and $I_{Q}^e$ are the electronic charge and heat currents flowing between the two electronic terminals, $e < 0$ is the electronic charge, $I_{Q}^{ec}$ is the heat current from the thermal terminal to the two electronic ones, $G_{in}$ is the conductance in the inelastic channels, $K_e^0$ is related to the electronic heat conductance between the electronic terminals and $K_{pe}$ is that between the thermal terminal and the electronic ones. $L_3$ is the off-diagonal heat conductance. $L_1$ and $L_2$ are related to the currents induced by the temperature differences (thermopower effect) and the current-induced temperature differences (refrigerator and heater effects). $\delta \mu = \mu_L - \mu_R$ ($\delta T = T_L - T_R$) is the chemical potential (temperature) difference between the two electronic terminals, and
\[ \Delta T = T_p - \frac{1}{2}(T_L + T_R) \] is the difference between the temperature of the thermal terminal and the average temperature of the two electric ones, with \( T_L, T_R \) and \( T_p \) being the temperatures of the left and right electronic terminals and the phonon terminal, respectively. With these definitions of \( \Delta T, \delta T \) and \( I_{Q}^p, I_{Q}^{pe} \), the Onsager reciprocal relationships are satisfied. In such a setup, as found in [16], the three-terminal Seebeck coefficient and figure of merit are, when \( G_{el} \) and \( K_{pp} \) are neglected,

\[ S_p = \frac{L_2}{T G_{in}}, \quad \bar{Z} T = \frac{L_2^3}{G_{pp} K_{pe} - L_2^2}, \] (7)

respectively. To obtain the transport coefficients and the figure of merit, we need to calculate the phonon-assisted currents through the system.

The Hamiltonian of the inter-band electron–phonon coupling is [27]\(^7\)

\[ H_{e-ph} = \frac{1}{\sqrt{V}} \sum_{kq\lambda,\nu,\rho} M_{kq,\lambda,\nu,\rho} c_{kq,\lambda,\nu}^{\dagger} c_{kq,\lambda,\nu} (a_{\lambda\nu} + a_{\lambda\nu}^{\dagger}) + H.c., \] (8)

where \( V \) is the volume of the system, \( \lambda \) is the phonon branch index, \( \nu (\rho) \) runs through the valence (conduction) band indices, \( M_{kq,\lambda,\nu,\rho} \) is the matrix element of the electron–phonon coupling and \( c^{\dagger} (a^{\dagger}) \) is the electron (phonon) creation operator. Due to momentum (when valid) and energy conservations, phonons involved in such processes will be in a small energy range. For indirect-band semiconductors and with momentum conservation, these phonons can be acoustic as well as optical ones. For simplicity and definiteness, we consider a direct-band semiconductor system and assume that the contribution from the optical phonons is the dominant one. From equation (8), the net electron–hole generation rate per unit volume, \( g_p \), is given for single-phonon transitions by the Fermi golden rule as

\[ g_p = \frac{2\pi}{\hbar} \sum_{kq,\lambda,\nu,\rho} |M_{kq,\lambda,\nu,\rho}|^2 [(1 - f_{k+q\nu}) f_{k\nu} N_{q\lambda} - f_{k+q\nu}(1 - f_{k\nu})(N_{q\lambda} + 1)] \delta (E_{k+q\nu} - E_{k\nu} - \omega_{q\lambda}) \]

\[ = \int dE_i dE_j \Gamma_p(E_i, E_j) [(1 - f(E_j)) f(E_i) N(\omega_{ji}) - f(E_j)(1 - f(E_i))(N(\omega_{ji}) + 1)], \] (9)

where

\[ \Gamma_p(E_i, E_j) = \frac{2\pi}{\hbar} \sum_{kq,\lambda,\nu,\rho} |M_{kq,\lambda,\nu,\rho}|^2 \delta (E_j - E_i - \omega_{q\lambda}) \delta (E_{k+q\nu} - E_j) \delta (E_{k\nu} - E_i). \] (10)

\(^7\) There are other possible inter-band transition processes, such as the defect trap state-mediated inter-band transitions (the Shockley–Read–Hall processes). These are higher-order processes, for which the rate should be smaller. There are also the Auger processes due to inter-band Coulomb interaction. The Auger processes are important for narrow-band semiconductors such as InSb. As far as we know, there is no theoretical work devoted to the comparison of the Auger processes and optical-phonon-induced inter-band transitions when the band gap is smaller than the optical-phonon energy. However, for the superlattice structures studied in this work, the transition should be very much like the inter-subband transitions in semiconductor quantum wells. In the latter, experimental and theoretical investigations (see [27]) revealed that the optical-phonon interaction (especially confined and interface phonons) dominates the inter-subband scattering. This supports our assumption that the inter-miniband transitions in the superlattice-based device are dominated by the optical-phonon processes. Moreover, as the Auger processes conserve the total energy of the electronic system, they will not affect the transfer of energy between the electronic system and the thermal bath. Hence, although we do not include the Auger processes in the theoretical treatment, they will not spoil the functionality and performance of the device.

New Journal of Physics 15 (2013) 075021 (http://www.njp.org/)
Here, $E_{k+q\ell}$, $E_{kv}$, and $\omega_{q\ell}$ are the electron and phonon energies when the two systems are uncoupled, and $f$ and $N$ are the non-equilibrium distributions of electrons and phonons in the intrinsic region, with $\omega_{ji} \equiv E_j - E_i > 0$ due to energy conservation. According to [28], when the length of the intrinsic region $L$ is sufficiently smaller than the carrier diffusion length, the electronic distribution in the conduction (valence) band in the intrinsic region can be well approximated as the distribution in the n-doped (p-doped) electronic terminal [28]. Similarly the phonon distribution is almost the same as that in the thermal terminal when the contact between the intrinsic region and the thermal terminal is good. Finally, a small amount of disorder that always exists in real systems and relaxes the momentum conservation can enhance the phonon-assisted inter-band transitions.

The transport coefficients are determined by studying the currents at a given bias and/or a temperature difference. The key relation is the continuity equation [28, 29]

$$\partial_t n_\alpha(z,t) = \frac{n_\alpha^eq(z) - n_\alpha(z,t)}{\tau_\alpha} - \frac{1}{q_\alpha} \partial_z I_\alpha(z,t) + g_p,$$

where $n_\alpha$ ($\alpha = e, h$) are the electron and the hole densities in the conduction and valence bands, and $n_\alpha^eq$ are the equilibrium values of those densities. $I_\alpha$ are the charge currents, $q_\alpha$ are the charges of the electron and the hole, $\tau_\alpha$ are the carrier lifetimes limited by the recombination processes other than the phonon-assisted ones that have already been taken into account in $g_p$, and $g_p$ is the net carrier density generation rate given in equation (9). The currents $I_\alpha$, which consist of diffusion and drift parts, are

$$I_\alpha(z) = -e \chi_\alpha n_\alpha(z) E - q_\alpha D_\alpha \partial_z n_\alpha(z),$$

where $\chi_\alpha$ and $D_\alpha$ are the mobilities and the diffusion constants, respectively. They are related by the Einstein relation, $D_\alpha = -(k_B T/e) \chi_\alpha$. $E = a E_b/(eL)$ is the built-in electric field in the intrinsic region.

If Boltzmann statistics for the electrons can be assumed everywhere, the net generation rate $g_p$ will depend on $z$ very weakly such that its spatial dependence can be ignored. In this situation, the total carrier densities can be divided into two parts, $n_\alpha = n_{\alpha, g} + n_{\alpha, n}$ where $n_{\alpha, g} = g_p \tau_\alpha$ are the spatially independent carrier densities generated by the phonon-assisted inter-band transitions and $n_{\alpha, n}$ are the ‘normal’ densities in the junction, determined by the continuity equation with $g_p = 0$. Similarly, the current is divided into two parts, $I_\alpha = I^a_\alpha + I^g_\alpha$. The current in the normal diode channel can be obtained from equations (11) and (12) with proper boundary conditions, yielding the celebrated rectification current–voltage relation

$$I^a_n = I^g_n (e^{-\delta \mu/k_BT} - 1),$$

where $I^a_n = -e D_\alpha L^{-1} n^m_\alpha$ being the saturated currents. Here $n^m_\alpha$ is the density of the minority carrier and $L_a$ is its diffusion length [28].

Inserting $n_{\alpha, g}$ into equation (12), one obtains the currents in the phonon-assisted channel\(^8\)

$$I^g_\alpha = -e \chi_\alpha g_p \tau_\alpha E = -\chi_\alpha \tau_\alpha E_b a L^{-1} g_p .$$

---

\(^8\) Here we have used certain simplifications. The consideration only holds for the central part of the intrinsic region, whereas the situation near the boundaries is more complicated. However, this is not supposed to change the results considerably [29].

New Journal of Physics 15 (2013) 075021 (http://www.njp.org/)
In the linear-response regime,

\[
g_\text{in} = \int dE_i dE_j \Gamma_p(E_i, E_j) f^0(E_i) f^0(E_j) N^0(\omega_{ji}) \left[ \frac{\delta \mu}{k_B T} + \frac{\overline{E}_{ij} - \mu}{k_B T} \frac{\delta T}{T} + \frac{\omega_{ji}}{k_B T} \frac{\Delta T}{T} \right]
\]

where \( f^0 \) and \( N^0 \) are the equilibrium distribution functions of the electrons and the phonons, respectively, and \( \overline{E}_{ij} \equiv (E_i + E_j)/2 \). Consequently

\[
G_\text{in} = -\frac{e E_g g_\text{in} \delta a}{k_B T L} \sum_a \chi_a \tau_a, \quad L_1 = \frac{G_\text{in}}{e} \langle \overline{E}_{ij} - \mu \rangle, \quad L_2 = \frac{G_\text{in}}{e^2} \langle (\overline{E}_{ij} - \mu)^2 \rangle, \quad L_3 = \frac{G_\text{in}}{e^2} \langle (\overline{E}_{ij} - \mu) \omega_{ji} \rangle, \quad K_{pe} = \frac{G_\text{in}}{e^2} \langle \omega_{ji}^2 \rangle,
\]

where \( g_\text{in} \) is the equilibrium transition rate (defined in equation (15)) and the average is defined in equation (4) with

\[
g_\text{in}(E_i, E_j) = -\frac{e E_g a}{k_B T L} \Gamma_p(E_i, E_j) f^0(E_i) f^0(E_j) N^0(\omega_{ji}) \sum_a \chi_a \tau_a.
\]

The above results are very similar to those obtained in [16]: due to the inelastic nature of the transport, the carrier energies at the p and n terminals are different; the heat transferred between the two terminals is the average one \( \overline{E}_{ij} - \mu \), whereas the energy difference \( \omega_{ji} \) is transferred from the thermal terminal to the two electronic ones. This is also manifested in the way the temperature differences are coupled to the heat flows [16] in equation (15), ensuring the Onsager relations. An important feature is that the three-terminal Seebeck coefficient, \( L_2/(T G_\text{in}) \), is negative definite because \( \omega_{ji} > 0 \). In contrast, the two-terminal Seebeck coefficient, \( L_1/(T G_\text{in}) \), does not possess this property. It can be positive or negative due to the partial cancelation of the contributions from electrons and holes, whereas there is no such cancelation for \( L_2 \) (see equation (16)). Equation (16) is a generalization of the results in [16] where there was only a single microscopic energy channel. When many inelastic processes coexist, the contribution of each process is weighed by its conductance. From equations (7) and (16), we find that the ‘ideal’ three-terminal figure of merit is

\[
\tilde{Z}T \bigg|_{\text{ideal}} = \frac{\langle \omega_{ji} \rangle^2}{\langle \omega_{ji}^2 \rangle - \langle \omega_{ji} \rangle^2}.
\]

For a single microscopic energy channel system where \( \omega_{ji} \) is fixed, this figure of merit goes to infinity [16]. When there are several such energy channels, it becomes finite due to the non-zero variance of \( \omega_{ji} \). We estimate the figure of merit when \( \langle \omega \rangle - E_g \simeq k_B T \) and \( \gamma, k_B T \ll \langle \omega \rangle \). Here \( \langle \omega \rangle \) is the average phonon energy and \( \gamma \) is the effective bandwidth (i.e. the variance of \( \omega_{ji} \) due to spectral dispersion) of the involved phonons. From equations (16) and (17) one finds that the variance of \( \omega_{ji} \) is limited by \( \gamma^2 \) or \( (k_B T)^2 \), whichever is smaller. For example, when the effective phonon bandwidth \( \gamma \) is much smaller than \( k_B T \), the variance is rather limited by \( \gamma^2 \). In the situations where \( \gamma, k_B T \ll \langle \omega_{ji} \rangle \), the numerator in equation (18) is much larger than

New Journal of Physics 15 (2013) 075021 (http://www.njp.org/)
the denominator. The ‘ideal’ figure of merit can be very high thanks to the electronic band gap when \( E_g \gg k_B T \) or the narrow bandwidth of the optical phonons \( \langle \omega \rangle \gg \gamma \). However, in realistic situations, as often happens, the parasitic heat conduction is another major obstacle to a high figure of merit. This will be analyzed in the next subsection.

2.3. Realistic figure of merit

Besides the phonon-assisted transport channel, there is the normal diode channel that is dominated by the (elastic) barrier transmission and the diffusion of minority carriers. It contributes to \( G, L_1 \) and \( K^0_e \) as well. In addition there are the ‘parasitic’ heat currents carried by phonons flowing between the two electronic terminals and those from the thermal terminal to the two electronic ones. Taking into account all these, the thermoelectric transport equations are written as

\[
\begin{pmatrix}
I_e \\
I_Q \\
I_Q^T
\end{pmatrix} = \begin{pmatrix}
G_{in} + G_{el} & L_1 + L_{1,el} & L_2 \\
L_1 + L_{1,el} & K_e^0 + K_{el}^0 + K_p & L_3 \\
\end{pmatrix} \begin{pmatrix}
\delta \mu / e \\
\delta T / T
\end{pmatrix},
\]

(19)

Here, \( I_Q = I_Q^a + I_Q^b \) is the total heat current between the two electronic terminals which consists of the electronic \( I_Q^a \) and the phononic \( I_Q^b \) contributions; \( I_Q^T = I_Q^{pe} + I_Q^{pp} \) is the total heat current flowing out of the thermal terminal to the two electronic ones with \( I_Q^{pe} \) being the purely phononic part. Finally, \( G_{el}, L_{1,el} \) and \( K^0_{el} \) are the contributions to the transport coefficients from the normal diode (elastic) channel, and \( K_p \) and \( K_{pp} \) are the heat conductances of phonons flowing between the two electronic terminals and those flowing from the thermal terminal to the two electronic ones, respectively. Note that the elastic channel does not contribute to \( L_2 \) and \( L_3 \) which are solely related to the inelastic processes. The three-terminal figure of merit can be obtained by optimizing the efficiency of a refrigerator working at \( \delta T = 0 \), but with finite \( \delta \mu \) and \( \Delta T \). This figure of merit is [16]

\[
\tilde{Z}T = \frac{L_2^2}{(G_{in} + G_{el})(K_{pe} + K_{pp}) - L_2^2},
\]

(20)

which is equivalent to equation (3). The diode conductance is \( G_{el} = -(e / k_B T) \sum_a I_{ns}^a \). A high figure of merit requires \( G_{in} \gg G_{el} \), which is not difficult to achieve according to the analysis in the next subsection (2.4). In such a situation and when \( \gamma \ll \langle \omega \rangle \) or \( k_B T \ll \langle \omega \rangle \), i.e. when the energy width due to \( k_B T \) or \( \gamma \) gives a much weaker limitation to \( \tilde{Z}T \) than \( K_{pp} \), one finds from equation (20)

\[
\tilde{Z}T \simeq \frac{K_{pe}}{K_{pp}}.
\]

(21)

Estimations carried out in section 2.4 indicate that there are parameter regimes where the figure of merit equation (20) can be greater than the usual two-terminal ones in the same material. We repeat that, unlike the two-terminal case, the thermal conductance \( K_p \) between the electronic terminals does not affect the three-terminal figure of merit.

9 Generally one can estimate the figure of merit as \( \tilde{Z}T \simeq \min \{ K_{pe} / K_{pp}, (E_\delta / k_B T)^2, (\langle \omega \rangle / \gamma)^2 \} \) when \( G_{el} \ll G_{in} \). We mostly assume that \( K_{pe} / K_{pp} \) is the smallest among the three.
Often $K_p$ and $K_{pp}$ are of the same order of magnitude. We note that $K_p$ and $K_{pp}$ are small in several gapless semiconductors such as PbSnTe, PbSnSe, BiSb, HgTe and HgCdTe. In addition, superlattice structures (and other planar composite structures) usually have much lower $K_p$ and $K_{pp}$ along the growth direction than the bulk materials [30]. The geometry where the electric current flows along that direction is promising for high figures of merit.

2.4. Estimation of the figure of merit

We defined $\tau_\alpha$ ($\alpha = e, h$) as the carrier lifetimes due to recombination processes other than the phonon-assisted ones that have already been taken into account in $g_p$ (see equation (9)). The carrier lifetimes due to the phonon-assisted processes, $\tau_{e,p}$ for the electrons and $\tau_{h,p}$ for the holes, satisfy the detailed balance relations $g_p^0 \tau_{e,p} = g_p^0 \tau_{h,p} = n_i$ in the intrinsic region. Here $n_i$ is the electron (hole) density in that region and $g_p^0$ is the equilibrium transition rate per unit volume defined in equation (15). The transition between minibands in semiconductor superlattices is usually dominated by the phonon-assisted processes in the dark limit when the miniband gap is smaller than the phonon energy [31]. We introduce the parameter $\zeta$ to write $g_p^0 \tau_\alpha \simeq g_p^0 \tau_\alpha = \zeta n_i$. This parameter is governed by the electron–phonon interaction strength, being of order unity when such a coupling is strong as is the case for some III–V (and other) semiconductors, or smaller (it will be taken as 1/4 below) [31]. Inserting this into equation (16), one finds

$$G_{in} \simeq e^2 (k_B T)^{-2} E_g \zeta n_i \sum_\alpha D_\alpha L^{-1},$$

where $L$ stands for the length of the intrinsic region, which will be taken to be much smaller than the diffusion length $L_\alpha$ [28]. In equation (22) we have chosen $a$ (defined in equation (5)) to be close to 1, which amounts to high doping density $N_d \lesssim N_0 \equiv n_i \exp[E_g/(2k_B T)]$ in the p- and n-doped regions. The elastic conductance is the slope of the $I$–$V$ rectification characteristics, equation (13), at $V = 0$,

$$G_{el} \simeq e^2 (k_B T)^{-1} n_i^2 N_d^{-1} \sum_\alpha D_\alpha L^{-1},$$

with $L_\alpha$ denoting the carrier diffusion length. The majority carrier densities (doping densities) in both the n and p regions are taken to be $N_d$. It follows then that

$$\frac{G_{el}}{G_{in}} \simeq \frac{k_B T}{E_g} \frac{n_i}{\zeta N_d} \frac{\sum_\alpha D_\alpha L_\alpha^{-1}}{\sum_\alpha D_\alpha L^{-1}}.$$

A small ratio can be easily achieved because $L < L_\alpha$, $k_B T < E_g$ and $n_i \ll N_d$. Therefore, the contribution of $G_{el}$ is not the main obstacle for a high figure of merit in this device.

For $G_{el} \ll G_{in}$ with $\gamma$ (the effective width of the relevant energy band) and $k_B T$ being much smaller than $\langle \omega \rangle$ (the typical phonon energy), the figure of merit $\tilde{Z} T$ is given by equation (21). The phononic heat conductance $K_{pp}$ is, as usual, the least known and an extremely important obstacle for increasing $\tilde{Z} T$. $K_{pp}$ can be reduced, in principle, by engineering the interfaces. Even without such improvement, according to the geometry $K_{pp} \simeq 4K_p$ where $K_p$ is the bulk phonon thermal conductance across the junction. In the relevant regime, $K_{pe} \simeq e^{-2} G_{in} \langle \omega \rangle^2 \gtrsim e^{-2} G_{in} E_g^2$. From equation (22) we then obtain

$$K_{pe} \simeq \left( \frac{E_g}{k_B T} \right)^2 e^{-\frac{E_g}{k_B T}} E_g \zeta N_d \sum_\alpha D_\alpha L^{-1}.$$

\[\text{New Journal of Physics 15 (2013) 075021 (http://www.njp.org/)}\]
We shall introduce yet another parameter to characterize the ratio \( K_e/K_p \) with \( K_e \) being the electron thermal conductance in a doped sample of the same geometry and size as the junction, with doping density \( N_a \). The temperature dependence of the ratio \( K_e/K_p \) varies with the doping, structure and temperature. The temperature dependence is assumed to be of the form \( K_e/K_p \propto T^\beta \) with \( \beta \) being a constant. A temperature-independent pre-factor parameter \( \xi = K_e E_g^\beta/(K_e k_B T^\beta) \) is then introduced. We further assume that the transport properties in the n and p regions are similar (up to signs). The Wiedemann–Franz law implies that \( K_e = \eta(k_B T)^2 e^{-2} G_D \), where \( G_D \) denotes the electrical conductance in the doped sample and \( \eta \approx 2 \) for Boltzmann statistics. According to [28], \( G_D \approx 0.25 G_{el} \exp[E_g/(k_B T)] \) when \( N_a \approx N_0 \) and \( L/L_a \) is small. Using equation (23) one can write \( K_{pp} \approx 4K_p \approx 4\xi^{-1} K_e \approx E_g^\beta (k_B T)^{1-\beta} \xi^{-1} \eta N_0 \sum_\alpha D_\alpha L_\alpha^{-1} \). The figure of merit is then estimated as

\[
\tilde{Z}T \approx \frac{K_{pp}}{K_{pe}} \approx \left( \frac{E_g}{k_B T} \right)^2 e^{-2\xi} E_g^\beta \xi N_0 \sum_\alpha D_\alpha L_\alpha^{-1} \approx \frac{E_g^\beta (k_B T)^{1-\beta} \xi^{-1} \eta N_0 \sum_\alpha D_\alpha L_\alpha^{-1}}{e^{-2\xi} \xi L_\alpha \eta L}. \tag{26}
\]

If the temperature dependence comes mainly from the first two factors, then a high figure of merit can be obtained for \( E_g/(2k_B T) = 3 - \beta \). For large \( L_a/L \), the three-terminal figure of merit can be larger than the two-terminal one. The power factor of the device, \( P = G_{in} S_{pp} \), can be estimated similarly.

An especially appealing setup exploiting the topological insulator–ordinary insulator–superlattices can be built as follows (taking the superlattice Bi₂Te₃/Si as an example). On the front and back surfaces of each thin Bi₂Te₃ layer, there are protected surface states with a gapless Dirac-cone like spectrum [32]. Tunneling between the two surfaces opens a gap in the spectrum of each surface band [24]. In superlattices these states form a pair of conduction and valence minibands where the band gap can be controlled via the thickness of the two types of layers (for details see [24] and footnote\(^\text{10}\)). As the energy of the optical phonon in Si is much larger than that in Bi₂Te₃, the optical phonons in Si layers are well localized in these layers. So are the optical phonons in the Bi₂Te₃ layers. Similarly, due to the significant mismatch of the mechanical properties of the two types of layers, the acoustic phonons also have difficulty in being transmitted across the interfaces. The phononic thermal conductivity along the growth direction is then considerably reduced. It should be smaller than the values for both materials in the bulk [25]. On the other hand, the phononic heat conductivity within each thin layer of Si is large. At \( T = 300 \) K, the former is 1.5 Wm\(^{-1}\)K\(^{-1}\) [33], whereas the latter is about 10\(^2\) Wm\(^{-1}\)K\(^{-1}\) [34]. When the electric current is along the growth direction, the small phononic heat conductance along this direction greatly reduces the parasitic heat conduction \( K_{pp} \) and benefits the figure of merit. On the other hand, within each Si layer phonons are transferred efficiently from the thermal terminal to the system, which is good for enhancing the output power of the device.

We now estimate the figure of merit of the device using the example of the semiconductor made of the Bi₂Te₃/Si superlattice. We shall use the transport parameters of bulk Bi₂Te₃ to do the

\(^\text{10}\) In [24] it is shown that the gap is \( E_g = 2|\Delta_S - \Delta_D| \) where \( \Delta_S \) and \( \Delta_D \) are the tunneling rates through the topological and the ordinary insulator layer, respectively. They (and the band gap) can be tuned in a large range via the thickness of these layers.
estimation, although the superlattice should have better thermoelectric performance [35]. First tune the superlattice structure to make \( E_g \lesssim \langle \omega \rangle \) with \( \langle \omega \rangle \) being the energy of the optical phonon in Si, which is about 60 meV (see [19]) (equivalent to about 700 K). We will choose \( E_g = 600 \) K. In Bi\(_2\)Te\(_3\) with \( N_d \simeq 10^{20} \text{cm}^{-3} \), one finds from [33] that \( \kappa_e \simeq \kappa_p \) at 300 K. This determines the parameter \( \xi \) to be \( 2^\beta \). Using these parameters we calculate and plot the figure of merit as a function of temperature in figure 2 for a practically achievable value \( L_u/L = 4 \) and a modest value of \( \xi = 1/4 \). In the calculation we ignored the temperature dependence of \( \zeta \) and \( L_u \). The results are computed for three situations with \( \beta = 1, 2 \) and 2.5. It is seen that the figure of merit with the several underestimations made can be larger than 1 around room temperature for all the three situations, indicating potential usefulness. Finally, we note that the same strategy and analysis can also be applied to the superlattice made of an inverted-band gapless semiconductor and a normal semiconductor, where the band gap can be tuned via the quantum-well effect [22] and the normal semiconductor can be chosen to have the proper optical phonon frequency and the high thermal conductivity to enable better thermoelectric performance.

3. Conclusions and discussions

We proposed and studied a thermoelectric device using the phonon bath as an example. The scheme is based on the ‘three-terminal’ geometry of thermoelectric applications [13–17], where inelastic processes play a crucial role. It has been shown that a high thermoelectric figure of merit can be achieved in this geometry in several nanosystems [16], where only one microscopic energy channel, in which the relevant electronic energy is fixed, is available. In this paper, we derived the figure of merit for the multiple energy channel case. We find that when only the inelastic processes are considered, the figure of merit is the ratio of the square of the mean value of energy difference between final and initial states to its variance with the average weighed by the conductance of each microscopic process. A small variance in the energy change is then favorable for a high figure of merit. To achieve such good energy selection, one can use either the electronic band gap, \( E_g \gg k_B T \), for electrons or the narrowness of the phonon band,
\( \gamma \ll \langle \omega \rangle \), for phonons. The realistic figure of merit including other processes, equation (3), is also discussed. It is found that a strong carrier–boson coupling as well as the dominance of inelastic transport and a small purely phononic thermal conductivity between the phonon terminal and the electronic ones are necessary for a high figure of merit. The suppression of elastic transport can be achieved with a semiconductor junction [28], while the coupling is strong when the bosons are e.g. phonons, electron–hole pair excitations, etc. Thanks to these, the proposed three-terminal device can have a figure of merit higher than that of the usual two-terminal device made of the same material.

In comparison with the existing literature, M–S suggested a narrow electronic band for elastic two-terminal transport to achieve high values of \( ZT \). When this scheme is generalized to inelastic processes where the initial \( E_i \) and final \( E_f \) energies are different, high values of \( ZT \) are possible when the distribution of the average energy \( \overline{E} = (E_i + E_f)/2 \) (measured from the common chemical potential) is narrow in the two-terminal geometry. In the three-terminal situation, a narrow distribution of \( \omega = E_f - E_i \) plays a crucial role. The latter can also be achieved by controlling the initial and final electronic states by a barrier enough higher than \( T \), or in a small system with a few initial and final states with fixed \( \omega = E_f - E_i \), as in [15–17], or for example, via the narrow bandwidth of optical phonons. Finally, we note the analogy of the suggested configuration, which may convert thermal to electrical energy, and a photovoltaic device.

Acknowledgments

OEW acknowledges support from the Albert Einstein Minerva Center for Theoretical Physics, Weizmann Institute of Science. This work was supported by the BMBF within the DIP program, BSF, by ISF and by its Converging Technologies Program.

References

[1] Harman T C and Honig J M 1967 Thermoelectric and Thermomagnetic Effects and Applications (New York: McGraw-Hill)

Goldsmid H J 2009 Introduction to Thermoelectricity (Heidelberg: Springer)

[2] Nolas G S, Sharp J and Goldsmid H J 2001 Thermoelectrics: Basic Principles and New Materials Development (Berlin: Springer)

[3] Snyder G J and Toberer E S 2008 Nature Mater. 7 105

[4] Dresselhaus M S, Chen G, Tang M Y, Yang R, Lee H, Wang D, Ren Z, Fleurial J P and Gogna P 2007 Adv. Mater. 19 1043

Shakouri A 2011 Annu. Rev. Mater. Res. 41 399

[5] Mahan G D and Sofo J O 1996 Proc. Natl Acad. Sci. USA 93 7436

[6] Ziman J M 1995 Principles of the Theory of Solids 2nd edn (Cambridge: Cambridge University Press) sections 7.7–7.9

[7] Hicks L D and Dresselhaus M S 1993 Phys. Rev. B 47 12727

Hicks L D and Dresselhaus M S 1993 Phys. Rev. B 47 16631

[8] Venkatasubramanian R 2000 Phys. Rev. B 61 3091

Yu J-K, Mitrovic S, Than D, Varghese J and Heath J R 2010 Nature Nanotechnol. 5 718

[9] See e.g. Cai J and Mahan G D 2008 Phys. Rev. B 78 035115

Zhou J and Yang R 2010 Phys. Rev. B 82 075324

Kratzer P, Fomin V M, Hülsen B and Scheffler M 2010 arXiv:1011.0324

New Journal of Physics 15 (2013) 075021 (http://www.njp.org/)
[10] Humphrey T E and Linke H 2005 Phys. Rev. Lett. 94 096601
[11] Zhou J, Yang R, Chen G and Dresselhaus M S 2011 Phys. Rev. Lett. 107 226601
   Jeong C, Kim R and Lundstrom M 2012 J. Appl. Phys. 111 113707
[12] Zvyagin I P 1973 Phys. Status Solidi b 58 443
[13] Rutten B, Esposito M and Cleuren B 2009 Phys. Rev. B 80 235122
   Cleuren B, Rutten B and Van den Broeck C 2012 Phys. Rev. Lett. 108 120603
[14] Entin-Wohlman O, Imry Y and Aharony A 2010 Phys. Rev. B 82 115314
[15] Sánchez R and Büttiker M 2011 Phys. Rev. B 83 085428
[16] Jiang J H, Entin-Wohlman O and Imry Y 2012 Phys. Rev. B 85 075412
[17] Sothmann B, Sánchez R, Jordan A N and Büttiker M 2012 Phys. Rev. B 85 205301
[18] Edwards H L, Niu Q and de Lozanne A L 1993 Appl. Phys. Lett. 63 1815
   Edwards H L, Niu Q, Georgakis G A and de Lozanne A L 1995 Phys. Rev. B 52 5714
   Kharpai V S, Ludwig S, Kotthaus J P, Tranitz H P and Wegscheider W 2006 Phys. Rev. Lett. 97 176803
   Zippilli S, Morigi G and Bachtold A 2009 Phys. Rev. Lett. 102 096804
   Prance J R, Smith C G, Griffiths J P, chorley S J, Anderson D, Jones G A C, Farrer I and Ritchie D A 2009
   Phys. Rev. Lett. 102 146602
   Jordan A N, Sothmann B, Sánchez R and Büttiker M 2013 Phys. Rev. B 87 075312
[19] Madelung O 2004 Semiconductors: Data Handbook 3rd edn (Berlin: Springer) pp 11–241
[20] See e.g. Mickevičius R and Reklaitis A 1990 J. Phys.: Condens. Matter 2 7883
[21] Tsidilkovski I M 1997 Electron Spectrum of Gapless Semiconductors (Berlin: Springer)
[22] Bernevig B A, Hughes T A and Zhang S C 2006 Science 314 1757
[23] Halász G B and Balents L 2012 Phys. Rev. B 85 035103
[24] Burkov A A and Balents L 2011 Phys. Rev. Lett. 107 127205
[25] Cahill D G, Ford W K, Goodson K E, Mahan G D, Majumdar A, Maris H J, Merlin R and Phillpot S R 2003
   J. Appl. Phys. 93 793
[26] Sze S M and Ng K K 2007 Physics of Semiconductor Devices (New York: Wiley)
[27] Woerner M and Elsaesser T 2001 Ultrafast nonequilibrium dynamics of intersubband excitations Ultrafast
   Phenomena in Semiconductors ed K-T Tsan (New York: Springer) p 127
[28] Shockley W 1950 Electrons and Holes in Semiconductors (New York: D Van Nostrand Company, Inc.)
[29] Tarettou K, Rau U and Werner J H 2003 Appl. Phys. A 77 865
[30] Sofo J O and Mahan G D 1994 Appl. Phys. Lett. 65 2690
   Broido D A and Reinecke T L 1995 Phys. Rev. B 51 13797
   Hylgaard P and Mahan G D 1997 Phys. Rev. B 56 10754
[31] Ridley B K 1993 Quantum Processes in Semiconductors (Oxford: Oxford University Press)
[32] Hasan M Z and Kane C L 2010 Rev. Mod. Phys. 82 3045
   Qi X L and Zhang S C 2011 Rev. Mod. Phys. 83 1057
[33] Satterthwaite C B and Ure R W Jr 1957 Phys. Rev. 108 1164
[34] Slack A 1964 J. Appl. Phys. 35 3460
[35] Touzelbaev M N, Zhou P, Venkatasubramanian R and Goodson K E 2001 J. Appl. Phys. 90 763