Quantum Effects in the Thermoelectric Power Factor of Low-Dimensional Semiconductors

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Quantum Effects in the Thermoelectric Power Factor of Low-Dimensional Semiconductors

Nguyen T. Hung,1,* Eddwi H. Hasdeo,1 Ahmad R. T. Nugraha,1 Mildred S. Dresselhaus,2,3 and Riichiro Saito1

1Department of Physics, Tohoku University, Sendai 980-8578, Japan
2Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA
3Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

(Received 15 April 2016; revised manuscript received 26 May 2016; published 15 July 2016)

We theoretically investigate the interplay between the confinement length \( L \) and the thermal de Broglie wavelength \( \Lambda \) to optimize the thermoelectric power factor of semiconducting materials. An analytical formula for the power factor is derived based on the one-band model assuming nondegenerate semiconductors to describe quantum effects on the power factor of the low-dimensional semiconductors. The power factor is enhanced for one- and two-dimensional semiconductors when \( L \) is smaller than \( \Lambda \) of the semiconductors. In this case, the low-dimensional semiconductors having \( L \) smaller than their \( \Lambda \) will give a better thermoelectric performance compared to their bulk counterpart. On the other hand, when \( L \) is larger than \( \Lambda \), bulk semiconductors may give a higher power factor compared to the lower dimensional ones.

DOI: 10.1103/PhysRevLett.117.036602

Thermoelectricity is a promising technology to improve renewable energy performance through conversion of waste heat into electric energy [1,2]. The efficiency of a solid-state thermoelectric power generator is usually evaluated by the dimensionless figure of merit, \( ZT = S^2\sigma\kappa^{-1}T \), where \( S \) is the Seebeck coefficient, \( \sigma \) is the electrical conductivity, \( \kappa \) is the thermal conductivity, and \( T \) is the absolute temperature. A fundamental aspect in the research of thermoelectricity is the demand to maximize the \( ZT \) value by having large \( S \), high \( \sigma \), and low \( \kappa \). However, since \( S, \sigma, \) and \( \kappa \) are generally interdependent, it has always been challenging for researchers to find materials with \( ZT > 2 \) at room temperature [3]. Huge efforts have been dedicated to reduce \( \kappa \) using semiconducting materials with low-dimensional structures, in which \( \kappa \) is dominated by phonon heat transport. For example, recent experiments using Si nanowires have observed that \( \kappa \) can be reduced below the theoretical limit of bulk Si (0.99 W/mK) because the phonon mean free path is limited by boundary scattering in nanostructures [4,5]. In these experiments, the reduction of the semiconducting nanowire diameter is likely to achieve a large enhancement in thermoelectric efficiency with \( ZT > 1 \) at room temperature [4,5]. The success in reducing \( \kappa \) thus leads to the next challenge in increasing the thermoelectric power factor \( PF = S^2\sigma \).

The importance of maximizing the PF can be recognized from the fact that when the heat source is unlimited, the \( ZT \) value is no longer the only one parameter to evaluate the thermoelectric efficiency. In this case, it is also important to evaluate the output power density \( Q \) [6,7]. The PF term appears in the definition of \( Q \), particularly for its maximum value, \( Q_{\text{max}} = PF(T_h - T_c)^2/4h_\varepsilon \), where \( T_h \), \( T_c \), and \( h_\varepsilon \) are the hot side temperature, cold side temperature, and the length between the hot and the cold sides (called the leg length), respectively. Since the term \( (T_h - T_c)^2/4h_\varepsilon \) is given by the boundary condition, \( Q \) is mostly affected by the PF. Here we mention the definition of \( Q \) because some materials show high \( ZT \) but low thermoelectric performance due to their small \( Q \). For example, Liu et al. has compared two materials: PbSe (with maximum values of \( ZT = 1.3 \), \( PF = 21 \mu \text{W/cmK}^2 \)) and \( \text{Hf}_{0.25}\text{Zr}_{0.75}\text{NiSn} \ (ZT = 1, \ PF = 52 \mu \text{W/cmK}^2) \) at \( T_h = 500^\circ\text{C} \) and \( T_c = 50^\circ\text{C} \) with a leg length \( h_\varepsilon = 2 \text{mm} \) [7]. Their calculation showed that PbSe (\( \text{Hf}_{0.25}\text{Zr}_{0.75}\text{NiSn} \)) has a thermoelectric efficiency of about 11\% (10\%), while its output \( Q \) is about 5.4 W/cm\(^2\) (14.4 W/cm\(^2\)). From this information, we can see that although PbSe has a larger \( ZT \), its output power is smaller than \( \text{Hf}_{0.25}\text{Zr}_{0.75}\text{NiSn} \). Therefore, increasing the PF value is important to enhance not only \( ZT \) but also \( Q \) for power generation applications. We thus would like to consider the issue of maximizing the PF as the main topic of the present work.

Of several methods to increase the PF value, the reduction of the confinement length \( L \), which is defined by the effective size of the electron wave functions in the nonprincipal direction for low-dimensional materials, such as the thickness in thin films and the diameter in nanowires, might be the most straightforward technique, since it was proven to substantially increase \( ZT \) [5,8–10]. A groundbreaking theoretical study by Hicks and Dresselhaus in 1993 predicted that a decrease in \( L \) can increase the PF and \( ZT \) of low-dimensional structures [11,12]. However, if we look at some previous works more carefully regarding the subject of the effect of confinement on the PF, there were some experiments which showed that the PF of one-dimensional (1D) Si nanowires is still similar to that of the 3D bulk system [4,5], while other experiments on
Bi nanowires show an enhanced PF value compared to its bulk state [10]. These situations indicate that there is another parameter that should be compared with \( L \). We will show in this Letter that the thermal de Broglie wavelength \( \Lambda \) is a key parameter that defines quantum effects in thermoelectricity. In order to show these effects, we investigate the quantum confinement effects on the PF for typical low-dimensional semiconductors. By comparing \( L \) with \( \Lambda \), we discuss the quantum effects and the classical limit on the PF, from which we can obtain an appropriate condition to maximize the PF.

In this Letter, we give an analytical formula for the optimum PF value which can show the interplay between the quantum confinement length and the thermal de Broglie wavelength of semiconductors with different dimensionalities. We apply the one-band model with the relaxation time approximation (RTA) to derive the analytical formula for the PF of nondegenerate semiconductors. The justification for the one-band model with the RTA was already given in some earlier studies, which concluded that the model was accurate enough to predict the thermo-electric properties of low-dimensional semiconductors, such as semiconducting single wall carbon nanotubes (s-SWNTs) [13], \( \text{Bi}_2\text{Te}_3 \) thin films [11], and Bi nanowires [12,14]. To obtain the PF formula in this work, we use similar analytical expressions for the Seebeck coefficient \( S \) and the electrical conductivity \( \sigma \) which were derived in our previous paper [13]. However, compared with Ref. [13], there is a modification to the definition of the relaxation time \( \tau(E) \) that we adopt in the present work, i.e., \( \tau(E) = \tau_0(E/k_B T)^r \), where \( \tau_0 \) is the relaxation time coefficient, \( E \) is the carrier energy, \( k_B \) is the Boltzmann constant, \( T \) is the average absolute temperature, and \( r \) is a characteristic exponent determining the scattering mechanism. In Ref. [13], \( \tau(E) \) was defined by \( \tau(E) = \tau_0 E^r \) [15,16], where we considered only the case of \( r = 0 \) or constant relaxation time approximation (CRTA) for discussing the Seebeck coefficients of s-SWNTs. Redefinition of \( \tau(E) = \tau_0(E/k_B T)^p \) is, however, suitable for purposes of this work.

The Seebeck coefficient \( S \) and the electrical conductivity \( \sigma \) are given, respectively, by [13,17]

\[
S = -\frac{k_B}{q} \left( \eta - r - \frac{D}{2} - 1 \right),
\]

and

\[
\sigma = \frac{4q^2 \tau_0 (r + \frac{D}{2})(k_B T)^{D/2}}{DL^{3-D}(2\pi)^D/\hbar^D \Gamma(D/2)^2} (m^*)^{D/2-1} e^\eta,
\]

where \( D = 1,2, \) or \( 3 \) denotes the dimension of the material (1D, 2D, or 3D systems), \( q = \pm e \) is the unit carrier charge, \( m^* \) is the effective mass of electrons or holes, \( L \) is the confinement length for a particular material dimension, \( \Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx \) is the Gamma function, \( \eta = \zeta/k_B T \) is the reduced chemical potential (while \( \zeta \) is defined as the chemical potential measured from the top of the valence energy band in a \( p \)-type semiconductor), \( k_B \) is the Boltzmann constant, and \( \hbar \) is Planck’s constant. Note that for an \( n \)-type semiconductor, we can redefine \( \eta \) or \( \zeta \) to be measured from the bottom of the conduction band, while the formulas for \( S \) and \( \sigma \) remain the same. From Eqs. (1) and (2), the thermoelectric power factor can be written as

\[
\text{PF} = S^2 \sigma = A(\eta - C)^2 e^\eta,
\]

where \( A \) (in units of \( \text{W/mK}^2 \)) and \( C \) (dimensionless) are given by

\[
A = \frac{4\tau_0 k_B^2}{L^3 m^*} \left( \frac{L}{\Lambda} \right)^D \frac{(r + \frac{D}{2}) \Gamma(r + \frac{D}{2})}{D \Gamma(D/2)^2},
\]

and \( C = r + D/2 + 1 \), respectively. In Eq. (4), the thermal de Broglie wavelength is defined by

\[
\Lambda = (2\pi \hbar^2 / k_B T m^*)^{1/2}
\]

which is a measure of the thermodynamic uncertainty for the localization of a particle of mass \( m^* \) with the average thermal momentum \( h(2\pi/\Lambda) \) [18].

For a given \( E \), the carrier mobility is defined by

\[
\mu = \frac{qv \langle \tau(E) \rangle}{m^*},
\]

where

\[
\langle \tau(E) \rangle = \frac{\langle E \tau(E) \rangle}{\langle E \rangle} = \tau_0 \frac{\Gamma(\frac{D}{2} + r)}{\Gamma(r + D/2)},
\]

and \( \langle x \rangle = \int_0^\infty x e^{-x/k_B T} dx \) in Eq. (7) is a canonical average of \( x \). From Eqs. (4), (6), and (7), the term \( A \) of the power factor can be rewritten as

\[
A = \frac{4\mu k_B^2}{qL^3} \left( \frac{L}{\Lambda} \right)^D \frac{(r + \frac{D}{2}) B(r, \frac{D}{2})}{D B(r, \frac{D}{2})},
\]

where \( B(x, y) = \Gamma(x) \Gamma(y) / \Gamma(x + y) \) is the Beta function. We can now determine the optimum power factor as a function of \( \eta \) from Eq. (3) by solving \( d(\text{PF})/d\eta = 0 \). The optimum power factor, \( \text{PF}_{\text{opt}} \), is found to be

\[
\text{PF}_{\text{opt}} = \frac{16\mu k_B^2}{qL^3} \left( \frac{L}{\Lambda} \right)^D \frac{(r + \frac{D}{2}) B(r, \frac{D}{2})}{D B(r, \frac{D}{2})} e^{r+D/2-1},
\]

whereas the corresponding value for the reduced (dimensionless) chemical potential is \( \eta_{\text{opt}} = r + D/2 - 1 \).
Next, we discuss some cases where $P_{\text{Fopt}}$ may be enhanced significantly. Figure 1 shows $P_{\text{Fopt}}$ as a function of the characteristic exponent $r$ for the 1D, 2D, and 3D systems, in which the values of $r$ range from −0.5 to 1.5 for various scattering processes [15,16]. In these examples, we consider a typical semiconductor, $n$-type Si, at room temperature and high-doping concentrations on the order of $10^{18}$ cm$^{-3}$. The thermal de Broglie wavelength and the carrier mobility are set to be $\Lambda = 4.5$ nm and $\mu = 420$ cm$^2$/V·s, respectively. We note that the scattering time assumed under the CRTA corresponds to $r = 0$, and thus $t_0 = r_0$ [19]. As shown in Fig. 1, $P_{\text{Fopt}}$ increases with increasing $r$ for all the 1D, 2D, and 3D systems. The effect of the characteristic exponent $r$ on the 3D system is stronger than that of the 1D and 2D systems. Based on Eq. (9) and Fig. 1, $P_{\text{Fopt}}$ increases with decreasing $L$ corresponding to the confinement effect for the 1D and 2D systems. It is noted in Fig. 1 that $P_{\text{Fopt}}$ in the 3D system does not depend on $L$ as shown in Eq. (9) with $D = 3$. However, the qualitative behavior between $r$ and $P_{\text{Fopt}}$ is not much affected by changing $L$ since $r$ and $L$ are independent of each other in Eq. (9).

Figure 2 shows $P_{\text{Fopt}}$ as a function of confinement length $L$ and thermal de Broglie wavelength $\Lambda$ for the 1D, 2D, and 3D systems. The mobility is set to be $\mu = 420$ cm$^2$/V·s for each system and the scattering rate may be proportional to the density of final states (DOS). By assuming proportionality of the scattering rate with respect to the DOS, we obtain $r = 0.5$, $r = 0$ and $r = -0.5$ for 1D, 2D, and 3D systems, respectively [16]. Hereafter, we consider such different $r$ values for the different dimensions. The curves in Figs. 2(a) and 2(b) in particular show a $L^{-2}$ and $L^{-1}$ dependence of $P_{\text{Fopt}}$ for 1D and 2D systems, respectively [cf. Eq. (9)]. These results are consistent with the Hicks-Dresselhaus model [11,12]. In addition, in this Letter, we point out that it is important to consider the dependence of $P_{\text{Fopt}}$ on $\Lambda$. For an ideal electron gas under a trapping potential, the thermodynamic uncertainty principle may roughly be expressed as $\Delta P/P \times \Delta V/V \geq (D^{3/2}/\sqrt{2\pi}) \Lambda/L$, where $P$ and $V$ are the pressure and volume of the system, respectively [20]. The uncertainty principle ensures that when the confinement length is comparable with the thermal de Broglie wavelength, i.e., $L \leq (D^{3/2}/\sqrt{2\pi}) \Lambda$, the $P$ and $V$ cannot be treated as commuting observables. In this case, quantum effects play an important role in increasing $P_{\text{Fopt}}$ for nanostructures. For a 1D system [Fig. 2(a)] $P_{\text{Fopt}}$ starts to increase significantly when $L$ is much smaller than $\Lambda$, while for the 2D system [Fig. 2(b)] $P_{\text{Fopt}}$ starts to increase significantly when $L$ is comparable to $\Lambda$. As for the 3D system [Fig. 2(c)], $P_{\text{Fopt}}$ increases with decreasing $\Lambda$ for any $L$ values. Therefore, a nanostructure having both small $L$ and $\Lambda$ (while $L$ is also much smaller than its $\Lambda$) will be the most optimized structure to enhance the PF.

Now we can compare our model with various experimental data. In Fig. 3, we show $P_{\text{Fopt}}$ as a function of $L/\Lambda$ for different dimensions (1D, 2D, and 3D systems) following Eq. (9). The $P_{\text{Fopt}}$ values are scaled by the optimum power factor of a 3D system, $P_{\text{Fopt}}^{3D}$. From Eq. (9), we see that the ratio $P_{\text{Fopt}}^{1D}/P_{\text{Fopt}}^{3D}$ merely depends on $L/\Lambda$ and $D$. Hence, PF from various materials can be compared directly with the theoretical curves shown in Fig. 3. The experimental data in Fig. 3 are obtained from the PF values of 1D Bi nanowires [10], 1D Si nanowires [5], 2D Si quantum wells [21], and two different experiments on 2D PbTe quantum wells labeled by PbTe-1 and PbTe-2 [22,23]. Here we use fixed parameters for the thermal de Broglie wavelength of each material: $\Lambda_{\text{Bi}} = 32$ nm, $\Lambda_{\text{Si}} = 4.5$ nm, and $\Lambda_{\text{PbTe}} = 5$ nm. We also set some PF values for bulk systems: $P_{\text{Fopt}}^{3D} = 0.002$ W/mK$^2$ [10], $P_{\text{Fopt}}^{3D} = 0.004$ W/mK$^2$ [24], $P_{\text{Fopt}}^{3D}$ [PbTe-1] = 0.002 W/mK$^2$ [22], and $P_{\text{Fopt}}^{3D}$ [PbTe-2] = 0.003 W/mK$^2$ [23], which are necessary to put all the experimental results into Fig. 3.

We find that the curves in Fig. 3 demonstrate a strong enhancement of $P_{\text{Fopt}}$ in 1D and 2D systems when the ratio...
experiments \cite{4,5}, are still too large compared with nanowires, which were about 36 nm. In this case, the diameters (supposed to represent \( L \)) of the 1D Bi nanowires can be enhanced at \( L < \Lambda \), and thus the confinement length is mostly determined by the size of the material such as \( L \approx L_0 \). Creating a 1D channel from a 2D material by applying negative gate voltages on two sides of the 2D material can be an example of how to engineer the confinement length \cite{27}.

We already see that the thermal de Broglie wavelength \( \Lambda \) depends on the temperature and the effective mass for the material. As given in Eq. (5), \( \Lambda \) decreases (\( \propto T^{-1/2} \) or \( m^{-1/2} \)) with increasing temperature \( T \) or with increasing effective mass \( m^* \), which indicates that the PF of nondegenerate semiconductors would be enhanced at higher \( T \) or at larger \( m^* \) (smaller \( \Lambda \)). This result is consistent with the experimental observations for the PF values of Si and PbTe, which monotonically increase as a function of temperature \cite{5,24,28}. It should be noted that \( \Lambda \) is not necessarily independent of \( L \) and \( D \) because the term \( m^* \) may be altered by varying \( L \) or by changing \( D \). For example, based on the fitting in Ref. \cite{29}, the effective masses of 1D Si nanowires for \( L \) within the interval of 2–40 nm could change from 1.1 m\(_0\) to 0.8 m\(_0\), where m\(_0\) is the free electron mass. Meanwhile, Ref. \cite{30} reported that bulk 3D Si has an effective mass of about 1.09 m\(_0\) at room temperature. As a result, we estimate that the change of \( \Lambda \) is roughly about 5%-10% in this case. This fact might contribute to the small discrepancy between the PF values from our theory and those from experiments since we set \( \Lambda \) as a fixed quantity upon variation of \( L \) in 1D and 2D systems (see Fig. 3). For the 3D system, the theoretical values (PF\(_{3D}^{Bi} = 0.0019 \) W/mK\(^2\) and PF\(_{Si}^{3D} = 0.0044 \) W/mK\(^2\)) are in good agreement with the experimental data (PF\(_{Bi}^{3D} = 0.002 \) W/mK\(^2\) \cite{10} and PF\(_{Si}^{3D} = 0.004 \) W/mK\(^2\) \cite{24}).

In conclusion, we have shown that the largest power factor PF values might be obtained for low-dimensional systems by decreasing both the confinement length \( L \) and the thermal de Broglie wavelength \( \Lambda \) while keeping \( L < \Lambda \). Depending on the materials dimension, there is a different interplay between \( L \) and \( \Lambda \) to enhance the power factor. A simple analytical formula [Eq. (9)] based on the one-band model has been derived to describe the quantum effects on the PF in 1D, 2D, and 3D systems. We would suggest to experimentalists to be careful to check the trade-off between the confinement length and the thermal de Broglie wavelength.

FIG. 3. PF\(_{opt}^{3D} / PF\(_{opt}\) as a function of \( L/\Lambda \) for different dimensions. The \( L/\Lambda \) axis is given using a logarithmic scale. Theoretical results for 1D, 2D, and 3D systems are represented by dashed, dotted, and solid lines, respectively. Asterisks, pentagons, diamonds, circles, and triangles denote experimental results for 1D Si nanowires \cite{5}, 1D Bi nanowires \cite{10}, 2D Si quantum wells \cite{21}, 2D PbTe-1 quantum wells \cite{22}, and 2D PbTe-2 quantum wells \cite{23}, respectively. For the experimental results, we set the thermal de Broglie wavelength of each material as: \( \Lambda_{Bi} = 32 \) nm, \( \Lambda_{Si} = 4.5 \) nm, and \( \Lambda_{PbTe} = 5 \) nm. We also have the following PF values for 3D systems: PF\(_{Bi}^{3D} = 0.002 \) W/mK\(^2\) \cite{10}, PF\(_{Si}^{3D} = 0.004 \) W/mK\(^2\) \cite{24}, PF\(_{PbTe-1}^{3D} = 0.002 \) W/mK\(^2\) \cite{22}, and PF\(_{PbTe-2}^{3D} = 0.003 \) W/mK\(^2\) \cite{23}.

L/\( \Lambda \) is smaller than one (\( L < \Lambda \)). In contrast, if \( L \) is larger than \( \Lambda \), the bulk 3D semiconductors may give a larger PF\(_{opt} \) value than the lower dimensional semiconductors, as shown in Fig. 3 up to a limit of \( L/\Lambda \approx 2 \). We argue that such a condition is the main reason why an enhanced PF is not always observed in some materials although experimentalists have reduced the material dimensionality. For example, in the case of 1D Si nanowires, where we have \( \Lambda_{Si} \approx 4.5 \) nm, we can see that the experimental PF values in Fig. 3 are almost the same as the PF\(_{opt}^{3D} \). The reason is that the diameters (supposed to represent \( L \)) of the 1D Si nanowires, which were about 36–52 nm in the previous experiments \cite{4,5}, are still too large compared with \( \Lambda_{Si} \). It might be difficult for experimentalists to obtain a condition of \( L < \Lambda \) for the 1D Si nanowires. In the case of materials having larger \( \Lambda \), e.g., Bi with \( \Lambda_{Bi} \approx 32 \) nm, the PF values of the 1D Bi nanowires can be enhanced at \( L < \Lambda \), which is already possible to achieve experimentally \cite{10}. Furthermore, when \( L \gg \Lambda \), it is natural to expect that PF\(_{opt} \) of 1D and 2D semiconductors resemble PF\(_{opt}^{3D} \) as shown by some experimental data in Fig. 3. It should be noted that, within the one-band model, we do not obtain a smooth transition of PF\(_{opt} \) in Fig. 3 from the lower dimensional to the 3D characteristics for large \( L \) because we neglect contributions coming from many other subbands responsible for the appearance of the 3D density of states \cite{25}.

So far, we have used the confinement length \( L \) as an independent parameter in Eq. (9). It is actually possible to engineer the confinement length in the same material. For extremely thin films or nanowires, \( L \) is expressed by two components as \( L = L_0 + \Delta L \), where \( L_0 \) is the thickness of the material and \( \Delta L \) is the size of the evanescent electron wave function beyond the surface boundary. Within the range of \( L_0 \), the electron wave function is delocalized, approximated by the linear combination of plane waves, while within \( \Delta L \) the electron wave function is approximated by evanescent waves. For a single-layered material, e.g., a hexagonal boron nitride (h-BN) sheet, \( L_0 \approx 0 \) so that \( L \approx \Delta L = 0.333 \) nm \cite{26}. As for ultrathick 1D nanowires or 2D thin films, we have \( L \gg \Delta L \), and thus the confinement length is mostly determined by the size of the material such as \( L \approx L_0 \). Creating a 1D channel from a 2D material by applying negative gate voltages on two sides of the 2D material can be an example of how to engineer the confinement length \cite{27}.
between $L$ and $\Lambda$ in order to enhance the PF for different dimensions of their semiconductors.

N. T. H. and A. R. T. N. acknowledge the Interdepartmental Doctoral Degree Program for Multidimensional Materials Science Leaders in Tohoku University. R. S. acknowledges MEXT (Japan) Grants No. 25107005 and No. 25286005. M. S. D. acknowledges support from NSF (USA) Grant No. DMR-1507806.

*nguyen@flex.phys.tohoku.ac.jp

[17] There is a difference in the exponent of the $k_B T$ terms in Eq. (2) of this work and Eq. (B3) in Ref. [13], i.e., $D/2$ for this work and $D/2 + r$ for Ref. [13]. This is a consequence of the definition of $\tau(E) = \tau_0(E/k_B T)^r$ adopted in this work and $\tau(E) = \tau_0 E^r$ in Ref. [13].