Effect of quantum-well structures on the thermoelectric figure of merit

L. D. Hicks

Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

M. S. Dresselhaus

Department of Electrical Engineering and Computer Science and Department of Physics,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Currently the materials with the highest thermoelectric figure of merit Z are Bi_2Te_3 alloys. Therefore these compounds are the best thermoelectric refrigeration elements. However, since the 1960s only slow progress has been made in enhancing Z, either in Bi_2Te_3 alloys or in other thermoelectric materials. So far, the materials used in applications have all been in bulk form. In this paper, it is proposed that it may be possible to increase Z of certain materials by preparing them in quantum-well superlattice structures. Calculations have been done to investigate the potential for such an approach, and also to evaluate the effect of anisotropy on the figure of merit. The calculations show that layering has the potential to increase significantly the figure of merit of a highly anisotropic material such as Bi_2Te_3, provided that the superlattice multilayers are made in a particular orientation. This result opens the possibility of using quantum-well superlattice structures to enhance the performance of thermoelectric coolers.

I. INTRODUCTION

For a material to be a good thermoelectric cooler, it must have a high thermoelectric figure of merit Z. The figure of merit is defined by[1]

$$Z = \frac{S^2\sigma}{\kappa}, \quad (1)$$

where S is the thermoelectric power (Seebeck coefficient), σ is the electrical conductivity, and κ is the thermal conductivity.

Currently, the materials with the highest Z are Bi_2Te_3 alloys such as $Bi_{0.5}Sb_{1.5}Te_3$, with $ZT \simeq 1.0$ at 300 K.[1] Only small increases in Z have been achieved in the past two decades, so it is now felt that the Bi_2Te_3 compounds may be nearing the limit of their potential performance.[2]

For a material to have a high Z, one requires a high thermoelectric power S, a high electrical conductivity σ, and a low thermal conductivity κ. It is difficult to improve Z for the following reasons. Increasing the thermoelectric power S for simple materials also leads to a simultaneous decrease in the electrical conductivity. Also, an increase in the electrical conductivity leads to a comparable increase in the electronic contribution to the thermal conductivity. So with known conventional solids, a limit is rapidly obtained where a modification to any one of these parameters adversely affects the other transport coefficients so that the resulting Z does not vary significantly.

In this paper, it is proposed that it may be possible to increase Z of some materials by preparing them in the form of multilayered superlattices. These structures may significantly alter Z since the electrons are now confined to move in two dimensions. In addition, the layering may reduce the phonon thermal conductivity and therefore increase Z as phonons can now be scattered by the interfaces between layers. Calculations have been performed to investigate the potential of such an approach, and also to explore the effect of using the anisotropy of some materials to enhance Z.

II. CALCULATION

The figure of merit Z has been calculated for (i) a three-dimensional (3D) bulk material and (ii) a 2D multilayered superlattice. The calculations assume a one-band material. This is because one-band materials (such as heavily doped semiconductors) give the best Z. The reason for this is as follows. For a two-band material, the measured thermoelectric power is

$$S = \frac{\sigma_1 S_1 + \sigma_2 S_2}{\sigma_1 + \sigma_2}, \quad (2)$$

where the subscripts 1 and 2 denote the contributions from bands 1 and 2, respectively. For S_1 and S_2 of opposite sign, S and Z are greatly reduced from either of their one-band values. This is the case for an intrinsic semiconductor, where the two bands are the conduction and valence bands. If S_1 and S_2 are of the same sign, then S is still reduced from the better of the one-band values.

The calculations are for a general, anisotropic, one-band material (assumed to be the conduction band). The only other assumptions are that of a constant relaxation time τ and that of parabolic bands.

Note that the calculations are *not* restricted to semiconductors. The material can be a metal, semiconductor,

or semimetal, as long as it is effectively a one-band material.

A. Z for a 3D bulk material

The methods for calculating S, σ, κ, and hence Z have been described elsewhere.[3,4] Since parabolic bands are assumed, the electronic dispersion relation used is

$$\varepsilon(k_x, k_y, k_z) = \frac{\hbar^2 k_x^2}{2m_x} + \frac{\hbar^2 k_y^2}{2m_y} + \frac{\hbar^2 k_z^2}{2m_z}. \tag{3}$$

For conduction along the x direction, one obtains

$$\sigma = \frac{1}{2\pi^2}\left(\frac{2k_B T}{\hbar^2}\right)^{3/2}(m_x m_y m_z)^{1/2} F_{1/2} e\mu_x, \tag{4}$$

$$S = -\frac{k_B}{e}\left(\frac{5F_{3/2}}{3F_{1/2}} - \zeta^*\right), \tag{5}$$

$$\kappa_e = \frac{\tau\hbar^2}{6\pi^2}\left(\frac{2k_B T}{\hbar^2}\right)^{5/2}\left(\frac{m_y m_z}{m_x}\right)^{1/2}$$
$$\times k_B\left(\frac{7}{2}F_{5/2} - \frac{25F_{3/2}^2}{6F_{1/2}}\right), \tag{6}$$

where the Fermi-Dirac function F_i is given by

$$F_i = F_i(\zeta^*) = \int_0^\infty \frac{x^i dx}{e^{(x-\zeta^*)} + 1}, \tag{7}$$

$\zeta^* = \zeta/k_B T$ is the reduced chemical potential (relative to the edge of the conduction band), τ is the relaxation time, m_x, m_y, m_z are the effective-mass components, μ_x is the mobility in the x direction, and κ_e is the electronic thermal conductivity.

So, using

$$Z = \frac{S^2 \sigma}{\kappa_e + \kappa_{\text{ph}}}, \tag{8}$$

where κ_{ph} is the phonon thermal conductivity, gives

$$Z_{3D}T = \frac{\frac{3}{2}\left(\frac{5F_{3/2}}{3F_{1/2}} - \zeta^*\right)^2 F_{1/2}}{\frac{1}{B} + \frac{7}{2}F_{5/2} - \frac{25F_{3/2}^2}{6F_{1/2}}}, \tag{9}$$

where

$$B = \frac{1}{3\pi^2}\left(\frac{2k_B T}{\hbar^2}\right)^{3/2}(m_x m_y m_z)^{1/2}\frac{k_B^2 T\mu_x}{e\kappa_{\text{ph}}}. \tag{10}$$

For a given value of B, the reduced chemical potential $\zeta^* = \zeta/k_B T$ may be varied to change the value of $Z_{3D}T$. The maximum value of $Z_{3D}T$ occurs when ζ^* is equal to its optimal value ζ^*_{opt} (inset of Fig. 1). Figure 1 shows a graph of ζ^*_{opt} against B, and Fig. 2 a graph of $Z_{3D}T(\zeta^*_{\text{opt}})$ against B.

In 3D bulk materials, B is determined largely by the intrinsic properties of the material, but ζ^* may be varied by doping. To maximize $Z_{3D}T$ for a material, one first calculates B for the intrinsic material, then uses Fig. 1
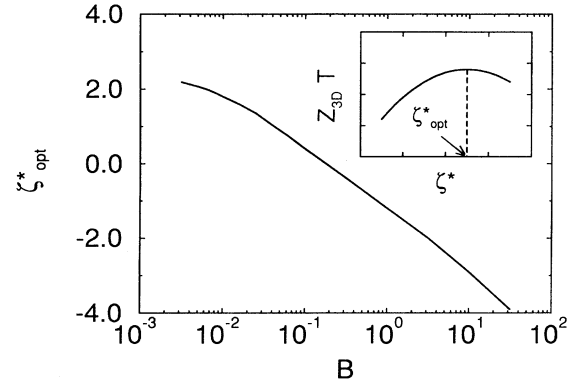

FIG. 1. Plot of ζ^*_{opt} vs B. The inset shows the variation of $Z_{3D}T$ with ζ^* at fixed B.

to determine the value of ζ^*_{opt} for this value of B. The next step is to adjust ζ^* so that $\zeta^* = \zeta^*_{\text{opt}}$: this may be achieved by doping with suitable impurities. The maximum $Z_{3D}T$ attainable for this value of B can be found from Fig. 2. Note that although doping changes ζ^* significantly, the corresponding change in B is negligible unless the material is very heavily doped, resulting in significant changes in the band structure or mobility.

From the graph of $Z_{3D}T$ against B in Fig. 2, it is clear that increasing B increases $Z_{3D}T$. With an anisotropic crystal in 3D, $Z_{3D}T$ varies with current direction and it is possible to increase B and therefore to maximize $Z_{3D}T$ by choosing the current to flow along the direction x of highest mobility μ_x.

B. Z for a 2D quantum well

For a suitably fabricated superlattice, the electrons are confined to move within the narrow-gap material in 2D quantum wells. Expressions for S, σ, κ, and Z are derived for transport in such quantum wells. The calculations assume that the electrons occupy only the lowest ($n=1$) subband of the quantum well (this is consistent with the optimization of Z arising from a one-band model), and also that there is no tunneling through the wide-gap semiconductor. The wide-gap semiconductor
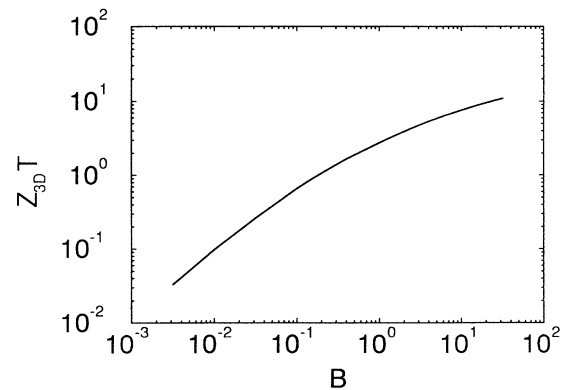

FIG. 2. Plot of $Z_{3D}T(\zeta^*_{\text{opt}})$ vs B.

does not contribute to the conduction since its carriers have found lower-energy states in the narrow-gap semiconductor.

Let the multilayers be parallel to the x-y plane and the currents flow in the x direction. The general expressions in Ref. 3 which were used to derive S, σ, κ, and Z in 3D were also used to derive the transport coefficients in a 2D quantum well. The electronic dispersion relation used is

$$\varepsilon(k_x, k_y) = \frac{\hbar^2 k_x^2}{2m_x} + \frac{\hbar^2 k_y^2}{2m_y} + \frac{\hbar^2 \pi^2}{2m_z a^2}, \tag{11}$$

indicating free-electron-like motion in the x-y plane and a bound state (for an infinite potential barrier) in the z direction. The results of the derivation are shown below:

$$\sigma = \frac{1}{2\pi a}\left(\frac{2k_B T}{\hbar^2}\right)(m_x m_y)^{1/2} F_0 e\mu_x, \tag{12}$$

$$S = -\frac{k_B}{e}\left(\frac{2F_1}{F_0} - \zeta^*\right), \tag{13}$$

$$\kappa_e = \frac{\tau\hbar^2}{4\pi a}\left(\frac{2k_B T}{\hbar^2}\right)^2\left(\frac{m_y}{m_x}\right)^{1/2} k_B \left(3F_2 - \frac{4F_1^2}{F_0}\right), \tag{14}$$

where F_i is defined in Eq. (7), a is the width of a quantum well, and

$$\zeta^* = \left(\zeta - \frac{\hbar^2 \pi^2}{2m_z a^2}\right)\Big/ k_B T. \tag{15}$$

So ZT for a 2D quantum well becomes

$$Z_{2D}T = \frac{\left(\dfrac{2F_1}{F_0} - \zeta^*\right)^2 F_0}{\dfrac{1}{B'} + 3F_2 - \dfrac{4F_1^2}{F_0}}, \tag{16}$$

where

$$B' = \frac{1}{2\pi a}\left(\frac{2k_B T}{\hbar^2}\right)(m_x m_y)^{1/2}\frac{k_B^2 T\mu_x}{e\kappa_{\rm ph}}. \tag{17}$$

For a given value of B', ζ^* in Eq. (15) may be varied to change the value of $Z_{2D}T$. The variation of $Z_{2D}T$ with ζ^* at fixed B' is similar to the variation of $Z_{3D}T$ with ζ^* at fixed B shown as an inset in Fig. 1. The maximum value of $Z_{2D}T$ occurs when $\zeta^* = \zeta^*_{\rm opt}$. Figure 3 shows a graph of $\zeta^*_{\rm opt}$ against B', and Fig. 4 a graph of $Z_{2D}T(\zeta^*_{\rm opt})$ against B'.

Note that for a 2D quantum well, the expression for ζ^* in Eq. (15) is different from the expression for the 3D case by the term $\frac{\hbar^2\pi^2}{2m_z a^2 k_B T}$, which is associated with confinement in the quantum well. Thus for the 2D case, ζ^* may be varied both by doping *and* by changing the layer thickness a. This extra degree of freedom allows an approach to increase $Z_{2D}T$ above the value characteristic of the bulk material as discussed below.

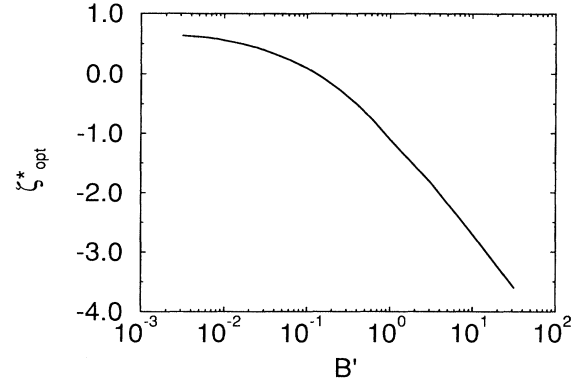To maximize $Z_{2D}T$ for a given quantum-well structure (and hence a fixed B'), one proceeds in a manner similar

FIG. 3.   Plot of $\zeta^*_{\rm opt}$ vs B' for a 2D quantum well.

to the 3D case. B' is calculated from Eq. (17) and $\zeta^*_{\rm opt}$ is determined from Fig. 3. $\zeta^* = \zeta^*_{\rm opt}$ is achieved by doping. Note that it is not possible to vary a to optimize ζ^*, as this will change the value of B' [Eq. (17)].

Recall that in 3D, the only means available to increase B in an anisotropic crystal is to choose the current direction to be the direction of highest mobility. For 2D quantum wells, the situation is more complex. There are more degrees of freedom available to increase B' and hence optimize $Z_{2D}T$. A higher mobility current direction will still give a higher B', but so will a narrower layer thickness a. It may also be possible to increase B' by a judicious choice of the crystallographic plane in which to make the layers. If the layers are made in the x-y plane and the currents flow along the x direction, then B' will be the expression in Eq. (17). If the layers are made in the x-z plane and the currents still flow along the x direction, then m_z will replace m_y in the expression for B'. In this orientation, B' and therefore $Z_{2D}T$ will be higher than the previous case if $m_z > m_y$. So one can increase $Z_{2D}T$ not only by choosing the optimum current direction, but also by using narrower layers and by choosing the best orientation in which to make the layers.
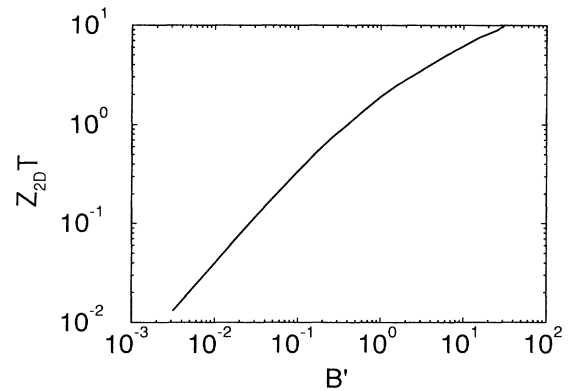

FIG. 4.   Plot of $Z_{2D}T(\zeta^*_{\rm opt})$ vs B' for a 2D quantum well.

III. EFFECT OF QUANTUM-WELL STRUCTURES ON Bi_2Te_3

One of the best materials for thermoelectric refrigeration is Bi_2Te_3, with a $Z_{3D}T = 0.67$ at 300 K.[1] The expressions derived previously are now used to calculate ZT for Bi_2Te_3 in both bulk form and in a quantum-well superlattice structure.

The equations for ZT derived so far assumed a single constant energy ellipsoid in the Brillouin zone. For multiple ellipsoids, the values of B and B' derived need to be multiplied by a number of the order of the number of ellipsoids. Bi_2Te_3 has six ellipsoids,[1] and multiplying B by a factor of 6 gives a value of $Z_{3D}T$ in good agreement with experiment, as shown below. The exact multiplicative factor will be slightly different from 6 because not all ellipsoids are oriented in the same direction, but this factor should not significantly affect the *comparison* between $Z_{3D}T$ and $Z_{2D}T$.

A. Calculated ZT for 3D bulk Bi_2Te_3

Bi_2Te_3 has a trigonal structure, which can be expressed in terms of a hexagonal unit cell of lattice parameters $a_0 = 4.3$ Å and $c_0 = 30.5$ Å.[1] The compound has a highly anisotropic effective-mass tensor, with effective mass components $m_x = 0.021m_0$, $m_y = 0.081m_0$, and $m_z = 0.32m_0$.[5] The phonon thermal conductivity is $\kappa_{ph} = 1.5$ W m^{-1} K^{-1} and the direction of highest mobility is the along the a_0 axis, with $\mu_{a_0} = 1200$ cm^2 V^{-1} s^{-1}.[1]

Substituting these values into Eq. (10), one obtains $B = 0.076$ at 300 K (after multiplying by 6). This gives a maximum $Z_{3D}T$ of $Z_{3D}T = 0.52$. This value is slightly different from the experimental result of $Z_{3D}T = 0.67$ (Ref. 1) because first the relaxation time τ may have a slight energy dependence, and second not all carrier pockets contribute equally, as mentioned previously. The maximum $Z_{3D}T$ occurs at $\zeta^*_{opt} = 0.9$, so the chemical potential is just above the conduction-band edge and the material is a partially degenerate n-type semiconductor.

B. Calculated ZT for Bi_2Te_3 layers in a quantum-well superlattice structure

In a quantum-well structure, since the electrons are confined to 2D motion parallel to the layers, there is no scattering off the interface between layers, so the carrier mobility in a direction parallel to the layers is unchanged. So if the a_0 axis is parallel to the layers, then $\mu_{a_0} = 1200$ cm^2 s^{-1} V^{-1} as before. However, phonons are not confined to move in 2D, so they can scatter off the interfaces. In thin layers this may reduce the phonon thermal conductivity κ_{ph}. Now

$$\kappa_{ph} = \tfrac{1}{3}C_v v l, \qquad (18)$$

where l is the phonon mean free path, C_v is the lattice heat capacity, and v is the velocity of sound in the material. For Bi_2Te_3, $C_v = 1.2 \times 10^6$ J K^{-1} m^{-3} and $v = 3 \times 10^3$ ms^{-1},[5] giving a value of $l = 10$ Å. If the layer

thickness a is greater than 10 Å, then layering does not seriously affect the mean free path l and κ_{ph} should then be the same as its bulk value. This is a conservative assumption used to make numerical estimates for $Z_{2D}T$, as interface scattering will still occur for $a > 10$ Å; this will cause a slight decrease in κ_{ph} and an increase in $Z_{2D}T$. However, if a is less than 10 Å, then l and κ_{ph} are limited by phonon scattering off the interfaces and a good estimate for κ_{ph} is obtained by setting $l = a$ and using Eq. (18). Again, this is a conservative estimate because the surface roughness and imperfections are expected to reduce κ_{ph} further.

From the expression for B' in Eq. (17), decreasing the layer thickness a increases B' and therefore increases $Z_{2D}T$. Also, if $a < 10$ Å, then κ_{ph} is reduced from its bulk value, resulting in an even greater increase in $Z_{2D}T$. So to achieve the best $Z_{2D}T$, it is necessary to make the layers as thin as possible.

When calculating ζ^*_{opt} for a material, one must use Eq. (15) to check the value of ζ to make sure that it does not lie above the energy $E_{n=2}$ of the next to lowest subband of the quantum well. If ζ does lie above $E_{n=2}$, then both the $n=1$ and $n=2$ subbands would contribute significantly to $Z_{2D}T$: this is inconsistent with the assumption of a one-band system and one would need to extend the model in order to get meaningful results. In practice, this is not a problem for the following reasons. In order to obtain a high $Z_{2D}T$, one requires a low a. This raises the energy of the $n=2$ subband because $E_{n=2} = 2\frac{\hbar^2\pi^2}{m_z a^2}$. At the same time, a lower a results in a higher B' and from Fig. 2 this means a lower ζ^*_{opt} and hence a lower ζ. So decreasing a moves $E_{n=2}$ up in energy and ζ down in energy. This means that for values of a below a certain characteristic thickness, ζ will always lie below $E_{n=2}$ and the model will be self-consistent. This characteristic thickness will usually be above the values of a which are of interest, as shown in the calculations for Bi_2Te_3 below.

$Z_{2D}T(\zeta^*_{opt})$ was calculated as a function of a for superlattice layers of Bi_2Te_3 in two distinct orientations: (i) layers parallel to the x-y plane (a_0-b_0 plane) and (ii) layers parallel to the x-z plane (a_0-c_0 plane). The current was assumed to flow along the high-mobility a_0 axis.

If the layers are made parallel to the x-y plane, the minimum possible layer thickness is 10.2 Å, as this is the length of the smallest repeating distance in the z direction.[1] Since this thickness is greater than 10 Å, $\kappa_{ph} = 1.5$ W m^{-1} K^{-1} as in 3D. If the layers are parallel to the x-z (a_0-c_0) plane, the minimum possible layer thickness is 3.8 Å. This is because the shape of the hexagonal unit cell is a parallelepiped, and the height of the cell perpendicular to the a_0-c_0 plane (i.e., in the y direction) is $a_0 \sin 60° = 3.8$ Å. So for $a > 10$ Å, $\kappa_{ph} = 1.5$ W m^{-1} K^{-1}, while for $a < 10$ Å, $\kappa_{ph} = \tfrac{1}{3}C_v v l$ where $l = a$. Figure 3 was used to calculate $Z_{2D}T(\zeta^*_{opt})$, after multiplying B' by a factor of 6. Note that layering lowers the symmetry of the material, so the effective number of carrier pockets contributing to $Z_{2D}T$ may be less than that for $Z_{3D}T$. Since there is no reliable way to estimate the actual number, the 3D value of 6 was assumed. The results are shown in Fig. 5, together with a line indicat-
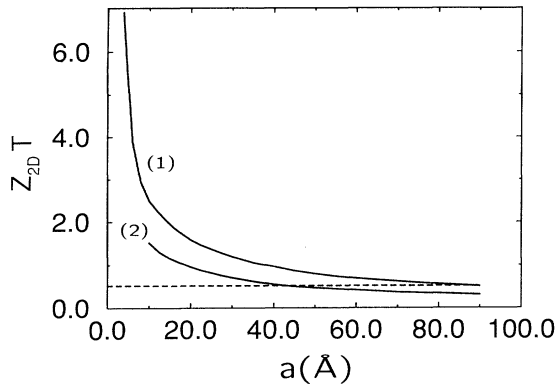
FIG. 5. Plot of $Z_{2D}T(\zeta_{opt}^*)$ vs layer thickness a for (1) a_0-b_0 plane layers and (2) a_0-c_0 plane layers. The dashed line indicates the best ZT for 3D bulk Bi_2Te_3.

ing the best 3D figure of merit of $Z_{3D}T = 0.52$.

Before any conclusions can be drawn, it is necessary to check the validity of the results by making sure that ζ does indeed lie below $E_{n=2}$ for the values of a in Fig. 5, i.e., for $a \leq 90$ Å. For the a_0-b_0 plane orientation, $B' = 0.085$ at $a = 90$ Å, giving $\zeta_{opt}^* = 0.2$ from Fig. 2. Using Eq. (15), this gives $\zeta = 0.7\frac{\hbar^2\pi^2}{m_x a^2}$, which is significantly less than $E_{n=2} = 2\frac{\hbar^2\pi^2}{m_x a^2}$, so the results for $a \leq 90$ Å are valid. Similarly, an a_0-c_0 plane layer of thickness $a = 90$ Å gives $B' = 0.17$ and $\zeta_{opt}^* = -0.05$. As ζ_{opt}^* is negative, ζ must be even lower than the $n=1$ subband, so the one-band model results are valid.

The results show that for layers parallel to the a_0-b_0 plane, $Z_{2D}T$ is higher than $Z_{3D}T$ for layers thinner than about 40 Å. The maximum $Z_{2D}T$ that can be obtained for this layer orientation is 1.5, which is about 3 times higher than the bulk value of 0.52. This value of $Z_{2D}T$ occurs at $a = 10.2$ Å, which is the minimum possible layer thickness.

For layers parallel to the a_0-c_0 plane, $Z_{2D}T$ is higher than $Z_{3D}T$ for layers thinner than 85 Å. As the layers are made even thinner, the increase in $Z_{2D}T$ becomes more significant, increasing sharply when $a < 10$ Å at which point phonon scattering off the interfaces becomes

important. To estimate a maximum $Z_{2D}T$ that can be obtained for this layer orientation, we assume $a = 3.8$ Å, a single-layer-thick quantum well, for which $Z_{2D}T = 6.9$, a 14-fold increase over the bulk value.

So, in principle, a huge increase in ZT can be achieved by using superlattices of Bi_2Te_3, provided the Bi_2Te_3 layers are oriented in the a_0-c_0 plane. Even if the layers are prepared in the a_0-b_0 plane, a factor of 3 increase over $Z_{3D}T$ is still possible, provided that very thin layers are used. To realize these increases, a number of experimental difficulties must be overcome. It may be difficult to produce a_0-c_0 plane layers, since thin films of Bi_2Te_3 grow predominantly in the a_0-b_0 plane.[6] It may also be difficult to prepare layers of uniform thickness if they are only a few unit cells thick. However, if these potential problems can be overcome, then a suitable lattice-matched wide-gap semiconductor must be found which will act as the quantum barrier in the superlattice. The band gap of the barrier material and the band offsets must be such that the electrons are confined to 2D motion in the Bi_2Te_3 layers.

IV. CONCLUSIONS

Using Bi_2Te_3 in a quantum-well structure has the potential to increase ZT by a factor of 13 over the bulk value. This large increase depends crucially on the material's highly anisotropic effective-mass tensor. In order to achieve this increase, the layers must be prepared in the a_0-c_0 plane and the current must flow along the high-mobility a_0 axis. If, however, the multilayers are prepared perpendicular to the c_0 axis along the a_0-b_0 plane, then there is still an increase of a factor of 3 over the bulk value, provided the layers can be made 10 Å thick.

The results of these calculations show that quantum-well structures can significantly modify and perhaps greatly improve the thermoelectric figure of merit of certain materials.

ACKNOWLEDGMENT

[1]H. J. Goldsmid, *Thermoelectric Refrigeration* (Plenum, New York, 1964).

[2]S. B. Horn, in *Proceedings of the 1st National Thermogenic Cooler Conference* (Center for Night Vision and Electro-Optics, Fort Belvoir, 1992).

[3]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), Chap. 13.

[4]E. S. Rittner, J. Appl. Phys. **30**, 702 (1959).

[5]H. H. Landolt and R. Börnstein, *Numerical Data and Functional Relationships in Science and Technology*, New Series, Vol. 17f (Springer-Verlag, Berlin, 1983), pp. 272–278.

[6]J. George and B. Pradeep, Solid State Commun. **56**, 117 (1985).