First-principles electronic structure and its relation to thermoelectric properties of Bi_2Te_3

S. J. Youn and A. J. Freeman
Department of Physics and Astronomy Northwestern University, Evanston, Illinois, 60208-3112
(Received 29 June 2000; published 7 February 2001)

The electronic structure of Bi_2Te_3, which is a major constituent of the best thermoelectric material operating at room temperature, was calculated by using the first principles full-potential linearized-augmented-plane-wave method with spin-orbit interaction included by a second variational method. A search of the whole Brillouin zone shows that the band edges are located off the symmetry lines, with locations that are in accord with the phenomenological six-band model. In doped Bi_2Te_3, Fermi surfaces near the band edges show a nonparabolic behavior. At a high doping concentration, the Fermi surfaces display elongated features, i.e., a knifelike Fermi surface for the valence band and spoonlike Fermi surfaces for the conduction band, which can be attributed to the layered structure of Bi_2Te_3. The effect of the anisotropic electronic structure combined with a low lattice thermal conductivity of Bi_2Te_3 gives a large figure of merit.

DOI: 10.1103/PhysRevB.63.085112 PACS number(s): 71.18.+y, 72.15.Jf, 71.20.−b

Bi_2Te_3 is a narrow-gap semiconductor which attracted considerable attention in the 1950s and 1960s due to its practical applications as a thermoelectric devices.[1] Today, it is one of the major constituents in the best thermoelectric material with the highest figure of merit for use near room temperature. A dimensionless figure of merit for thermoelectric materials is defined as

$$ZT = S^2 \sigma T / (\kappa_e + \kappa_L), \tag{1}$$

where S is the thermopower, σ is the electrical conductivity, T is the temperature of the sample, and κ_e and κ_L are the thermal conductivities due to the electron and lattice, respectively. Whereas much effort has been made to improve the ZT of materials based on Bi_2Te_3 by doping or alloying other materials, ZT is not much improved, and is not much more than $ZT \sim 1$, which is still not high enough compared to compressor-based refrigerators. These days, the demand for an efficient thermoelectric material is increasing due to its need in electronics and out of environmental concern.[2] As a result, thermoelectric materials have begun to receive renewed attention. Several theoretical concepts were proposed to increase the efficiency of the materials.[3] One idea to improve the figure of merit is to use superlattice or quantum-well structures;[4] a preliminary report indicated that a ZT of about 3 can be obtained at 300 K in a Bi_2Te_3/Sb_2Te_3 superlattice,[5] which is comparable to that of compressor-based refrigerators.

In developing materials with a higher figure of merit, it is crucial to understand their electronic structure, especially around the energy gap, as this determines the performance of the material. Several phenomenological models were proposed to explain experimental observations involving the Fermi surface of doped Bi_2Te_3. Six valley models for both conduction and valence bands with high anisotropy in transport properties are generally accepted in which both lowest conduction band and highest valence band have six valleys in the mirror plane of the Brillouin zone (BZ).[6] In addition to these six-valley models for band edges, second band edges are needed to explain the experiments.[7,8] Further, whereas most phenomenological models assume quadratic bands, i.e., $E \sim k^2$, nonquadratic behavior was observed.[9,10]

Whereas there are several electronic structure calculations for Bi_2Te_3 by various methods like pseudopotential,[11–13] tight binding,[14] linear augmented plane wave,[15] linear muffin-tin orbital,[16] and very recently by full-potential linearized augmented plane wave,[17] phenomenological six-band model have not yet been confirmed. Recent first-principles studies of the electronic structure by Mishra *et al.*[16] and Larson *et al.*[17] gave only two conduction-band valleys, while both gave six valleys for the valence band. Mishra *et al.* ascribed their failure in finding the six valleys to their atomic sphere approximation. However, Larson *et al.* could not find six valleys for the conduction band, although they used a full-potential method. This may be due to different approximations for the exchange-correlation potential or by different linearizations of the wave function. They utilized the generalized gradient approximation, while our uses are the local-density approximation (LDA). Also, they set the energy cutoff between the core and valence states at -6 Ry where Te $4p$ electrons are treated as the valence states while ours is set at -1.8 Ry, where Te $4p$ electrons are treated as the core states.

In this paper, we confirm the six-band model for both conduction and valence bands from a search for global extrema through the whole Brillouin zone. We find that the band edges are located at points off the symmetry lines. In previous calculations, electronic energy bands were studied along symmetry lines connecting two symmetry points in the mirror plane. Since the energy differences between band edges and other local extrema are very small, it is important to use accurate potentials. For this reason, we have obtained the electronic structure by employing the highly precise first-principles local-density FLAPW method,[18] which makes no shape approximations to the charge density or the potential. From the calculated electronic band structure, we give several reasons for the large figure of merit of Bi_2Te_3.

Bi_2Te_3, a naturally layered material, has a trigonal crystal structure with space group D_{3d}^5. The crystal structure can be given in a hexagonal unit cell as three sets of hexagonal layers where each set consists of five atoms,

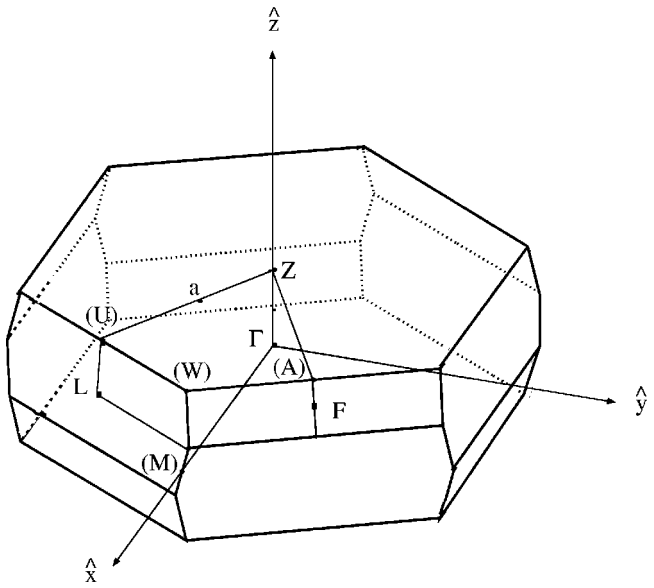

FIG. 1. Brillouin zone of Bi_2Te_3 according to Ref. 16. The points in parentheses are not points of symmetry because whose symmetry is not higher than the neighbor.

Te^1-Bi-Te^2-Bi-Te^1. The bonding within the Te^1-Bi-Te^2-Bi-Te^1 layer is considered to be covalent, while the bonding between layers (Te^1-Te^1) is of weak van der Waals type. We have used a trigonal lattice where only five atoms are contained in a primitive unit cell. Hereafter all coordinates refer to the trigonal unit cell. An experimental lattice constant[19] is used in the calculation. The spin-orbit interaction, which plays an important role in the electronic structure of Bi_2Te_3, was included self-consistently by a second variational method.[20] The Hedin-Lundqvist form of the exchange-correlation potential was employed for the exchange-correlation energy.[21] In all, 60 special **k** points[22] are sampled for the **k**-space integration, which is sufficient for an electronic structure calculation for a semiconductor.

Figure 1 shows the Brillouin zone of a trigonal crystal structure.[16] Some points are not symmetric points, since the symmetry of the point is not higher than its neighbor's. Three mirror planes containing the *L*, Γ, and *Z* points are parallel to the *z* axis, whereas the basal plane passing through the Γ point is perpendicular to the *z* axis. Figure 2(a) shows the energy bands plotted along the symmetry lines. Overall, the band structure shows similarity to that of Mishra *et al.*,[16] except for small differences caused by the different potentials employed; Mishra *et al.* used the atomic sphere approximation for the potential in their linear muffin-tin orbital calculation. If we see the band structure along the symmetry lines, the conduction-band edge (CBE) and valence-band edge (VBE) are located between *Z* and *F*, with an energy gap of 106 meV. However, since it is necessary to check if the CBE and VBE on the symmetry line are true extrema, we sampled 1000 **k** points randomly in the irreducible Brillouin zone. As a result, we were able to confirm that the CBE and VBE lie in the mirror plane. However, their locations are off the symmetry line. Thus if one looks only at the band along the symmetry lines connecting symmetry points, one can

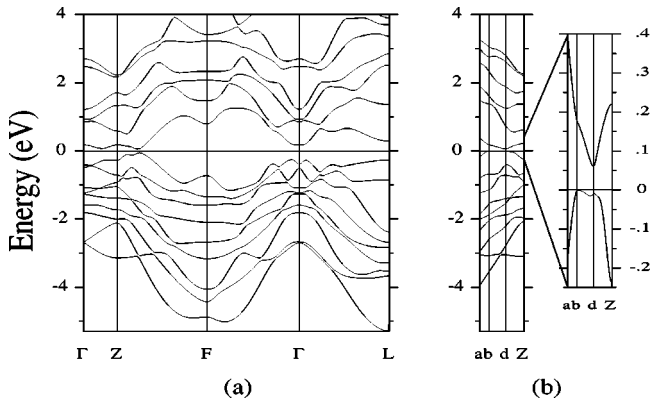

FIG. 2. Band structure of Bi_2Te_3 (a) along the high-symmetry lines and (b) along the lines connecting band extrema.

miss the band edges. Figure 2(b) shows the energy-band structure of Bi_2Te_3 along the lines connecting the band edges. The energy gap (of 61 meV) is smaller than the 120 meV found in experiment. This is in keeping with the well-known fact that the LDA usually gives smaller energy gaps than experiment.

Figure 3(a) shows a contour plot of the eigenvalues of the highest valence band in the mirror plane where $\mathbf{z}_1 = \frac{1}{2}Z$, $\mathbf{z}_1 = \frac{3}{2}Z$, and **a** is a point located at 0.435 of the distance from *Z* to *U*. The contour lines in Fig. 3(a) can be regarded as showing the cross section of the hole Fermi surface with various doping concentration since the Fermi level will be located at an energy level of the contour lines for a given doping concentration. In Fig. 3(a), the VBE is located at **b**(0.546,0.383,0.383), which is off the symmetry lines. The threefold rotation and inversion symmetry of the crystal gives six valence-band valleys in agreement with experiment. However, there is a second-highest valence-band edge (SVBE) at **c**(0.665,0.586,0.586) which differs in energy by only 3.8 meV (or 46 K). If we count these two Fermi surfaces together, 12 Fermi surface pockets can exist at a certain hole doping concentration, which may give rise to a better

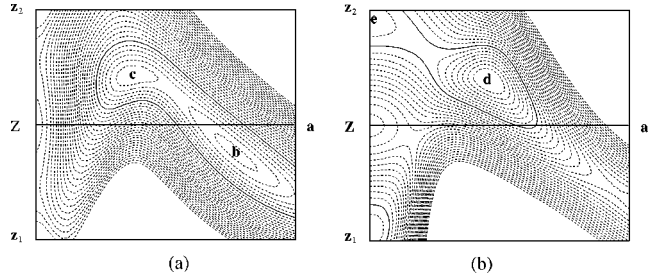

FIG. 3. (a) Contour plot of eigenvalues of the highest valence (24th) band in the mirror plane. **b** and **c** represent the location of the valence-band edge and the second-highest valence-band edge, respectively. The solid line represents a cross section of the Fermi surface at a hole doping concentration of 10^{19} cm^{-3}. (b) Contour plot of eigenvalues of the lowest conduction (25th) band in the mirror plane. **d** and **e** represent the location of the conduction-band edge and the second-lowest valence-band edge, respectively. The solid line represents a cross section of the Fermi surface at an electron doping concentration of 10^{19} cm^{-3}.

efficiency of Bi_2Te_3 since the figure of merit is a monotonic increasing function of the number of Fermi surface, as will be shown in Eq. (2). The SVBE and VBE can be connected to give a combined Fermi surface when it is doped with more than 1.0×10^{15} cm^{-3} acceptors. Since Bi_2Te_3 is used as a thermoelectric device with heavy doping like 10^{19} cm^{-3}, one would observe only one Fermi surface comprised of the two Fermi surfaces at **b** and **c**. The solid line with the shape of a knife in Fig. 3(a) represents the cross section of the Fermi surface at a doping of 10^{19} cm^{-3}. Thus at high doping, the Fermi surface cannot be described as an ellipsoidal surface. The knifelike Fermi surface is tilted by about 42° with respect to the basal plane, which is larger than the experimental value of 31.5°.[9] This difference between theory and experiment can be ascribed to the simplifying assumption in experiment that the energy changes quadratically, like $E\sim k^2$, near the band edge. However, this quadratic assumption cannot be applied at doping concentrations greater than 1.0×10^{15} cm^{-3}, as we can see from the combined two Fermi surfaces in Fig. 3(a). Thus it follows that more sophisticated experimental analyses are required to check the topology of the Fermi surface.

This knifelike Fermi surface gives rise to a large anisotropy in the electronic transport properties. If we denote the most elongated direction of the Fermi surface as z', the elongated Fermi surface is due to a very flat band around the z' axis which can be seen in the highest valence band in Fig. 2(b) between **b** and **d**. The elongated Fermi surface also gives a large effective mass for electronic transport in the z' direction, which makes the mobility in this direction very small. The large anisotropy in the effective mass can be ascribed to the layered structure of Bi_2Te_3, since cylinderlike Fermi surfaces elongated in a direction are usually observed in layered materials like graphite or layered superconductors. However, the length of the knifelike Fermi surface in the z' direction is finite, which is different from other layered materials with infinite length. A finite length means that the two-dimensional layered structure is not complete, and has a moderate bonding between layers. If we describe the bonding between layers as covalent bonding, we can see that the bonding between layers has a mixed covalent and weak van der Waals character. In order to increase the mobility of holes, charge transport should be made within the basal plane or along trigonal axes in order to avoid the large effective mass in the z' direction.

We can see a third VBE in Fig. 3(a) between Γ and Z. However, the energy difference between the third VBE and the VBE is 228 meV, which corresponds to a very high hole doping of 6.5×10^{20} cm^{-3} or 0.11 electrons/f.u., which is higher than the usual doping concentration. Hence the Fermi surfaces of hole-doped Bi_2Te_3 can be effectively modeled by combining the two Fermi surfaces at **b** and **c**. This is similar to the phenomenological two-band model for the valence band,[6] but the energy separation between the VBE and SVBE is found to be very small in the calculation.

Figure 3(b) shows a contour plot of the energy eigenvalues of the lowest conduction band. The CBE is located at **d**(0.663,0.568,0.568), and the second-lowest conduction-band edge can be found at **e**(0.273,0.273,0.273), which is about halfway between Γ and Z. If the electron doping is low, only the Fermi surface at **d** is observed. The Fermi surface at **d** gives six valleys by crystal symmetry, which confirms the six-valley model for the conduction band. The shape of the Fermi surface at **d** is like a lens, and does not change much from the shape of low doping up to a moderate doping concentration, which means that the effective mass is nearly constant with respect to doping. However, if it is doped heavily (like 10^{19} cm^{-3}) the two Fermi surfaces at **d** and **e** become connected to give one Fermi surface with a complicated effective mass, as seen from the solid line in Fig. 3(b), which is the Fermi-surface cross section at an electron doping of 10^{19} cm^{-3}. This Fermi-surface cross section has the shape of a spoon which is tilted 31° with respect to the basal plane, in excellent agreement with the experimental value of 33.5°.[10]

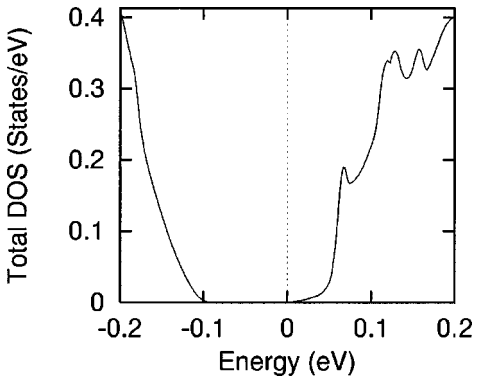

FIG. 4. Density of states near the energy gap.

Again, at high doping an elongated Fermi surface in one direction can be ascribed to the quasi-two-dimensional nature of the crystal. The six Fermi surfaces are not independent, since each group of three Fermi surfaces meets at a point **e** or $-$**e**. The combined Fermi surfaces at **e** and $-$**e** have the shapes of tripods, whose arms are open toward Z. Due to the crystal symmetry, the tripodlike Fermi surfaces at $-$**e** and **e** are point-inverted images of each other with respect to Z. In this heavy doping case (10^{19} cm^{-3}), a simple six-valley model may not work for the conduction band. The energy difference between **d** and **e** is 53 meV, which corresponds to a doping of 4.8×10^{18} cm^{-3} or 8.2×10^{-4} electrons/f.u. Thus the conduction band can be modeled as the six-band model at low doping concentration ($<4.8\times10^{19}$ cm^{-3}), while they become connected to give two tripodlike Fermi surfaces at a high doping concentration.

Figure 4 shows the density of states (DOS) around the energy gap. We can see a steep increase in the DOS at around 50 meV, which is due to the second conduction-band edge. Hence if one were to measure the DOS by a suitable experiment like inverse photoemission spectroscopy, the peak structure in the DOS would be observed.

From our results we can think of several reasons why Bi_2Te_3 has a large thermoelectric figure of merit: Mahan[23] showed that ZT is a monotonic increasing function of the B coefficient, where

$$B_j = Te\mu n_j(T)(k_B/e)^2/(4\kappa_L), \quad (2)$$

$$n_j(T) = N_j(2\pi m_j^* k_B T/h^2)^{3/2}. \quad (3)$$

Here e is the charge, μ is the mobility, κ_L is the lattice thermal conductivity, and N_j is the degeneracy of the band. The density of states effective mass of the jth band is given by $m_j^* = (m_1 m_2 m_3)^{1/3}$, where m_1, m_2, and m_3 are the effective masses along the principal axes of the constant energy surface of the jth band. Although Mahan's calculation assumes a parabolic band, his results suggest important qualitative features for improving ZT. Obviously, to increase ZT, we should increase B_j. Thus a high carrier mobility, a large density-of-states effective mass, a large multiplicity N_j, and a low lattice thermal conductivity are required for a good thermoelectric material, which generally agrees with the observed properties of good thermoelectric materials. As for Bi_2Te_3, there are several factors that can contribute to an increase of B. First the band edges are located at **k** points with very low symmetry, which gives a high multiplicity N_j of the equivalent Fermi surface; these **k** points only have mirror reflection symmetry, so that as many as six equivalent Fermi surfaces can be generated. Second, the quasi-two-dimensional character of Bi_2Te_3 gives a large anisotropy in the effective mass, and hence a large density of states mass, m_j. Third, while m_j is large, we can choose a direction with a small effective mass due to the large anisotropy of the Fermi surface, and a good mobility μ can be obtained. Aside from these electronic properties, Bi_2Te_3 has a low thermal conductivity due to the large atomic mass of its constituent elements. The combined effect of electronic structure and lattice effects can give rise to a large figure of merit.

In conclusion, we have calculated the electronic structure using the first-principles local-density FLAPW method. By globally searching throughout the whole BZ, we have found that the valence- and conduction-band edges are located off the symmetry lines. The Fermi surfaces show high anisotropy in the effective mass, which is due to the quasi-two-dimensional nature of the layered structure. The incomplete two-dimensional nature of the Fermi surface also reveals that the bonding between layers is not only of van der Waals type but is also covalent. The valence-band Fermi surface of heavily hole doped Bi_2Te_3 can be modeled by six knifelike Fermi surfaces, where each Fermi surface is made of two merged Fermi surfaces. The conduction band of Bi_2Te_3 with low electron doping has six lenslike Fermi surfaces. However, if it is heavily doped with electrons, the six-valley model does not work, since two tripodlike Fermi surfaces are formed. From an analysis of the electronic structure, we have found that Bi_2Te_3 has several good reasons to have a high figure of merit, since it has many Fermi surfaces due to the low symmetry of the position of the band edges, a large density-of-states effective mass due to the quasi-two-dimensional layered structure, and a low mobility due to the large anisotropy of the Fermi surface. In addition to these electronic properties, the combined effect with a low lattice thermal conductivity can yield a high figure of merit.

ACKNOWLEDGMENTS

This work was supported by DARPA (DAAG55-97-0130) and by a grant of computer time at the San Diego Supercomputer Center supported by the National Science Foundation.

[1] D.R. Lovett, *Semimetals and Narrow-Bandgap Semiconductors* (Pion Limited, London, 1977), p. 181.
[2] T.M. Tritt, Science **283**, 804 (1999).
[3] G. Mahan, B. Sales, and J. Sharp, Phys. Today (3) **50**, 42 (1997).
[4] L.D. Hicks and M.S. Dresselhaus, Phys. Rev. B **47**, 12 727 (1993).
[5] F.J. DiSalvo, Science **285**, 30 (1999).
[6] J.R. Drabble, R. Groves, and R. Wolfe, Proc. Phys. Soc. London **71**, 430 (1958).
[7] R.B. Mallinson, J.A. Rayne, and R.W. Ure, Jr., Phys. Rev. **175**, 1049 (1968).
[8] V.V. Sologub, A.D. Goletskaya, and R.V. Parfen'ev, Fiz. Tverd. Tela (Leningrad) **14**, 915 (1972) [Sov. Phys. Solid State **14**, 783 (1972)].
[9] H. Köhler, Phys. Status Solidi **74**, 591 (1976).
[10] H. Köhler, Phys. Status Solidi **73**, 95 (1976).
[11] S. Katsuki, J. Phys. Soc. Jpn. **26**, 58 (1969).
[12] R. Togei and G.R. Miller, in *The Physics of Semimetals and Narrow-Gap Semiconductors*, edited by D.L. Carter and R.T. Bate (Pergamon Press, Oxford, 1971), p. 349.
[13] E.V. Oleshko and V.N. Korolyshin, Fiz. Tverd Tela (Leningrad) **27**, 2856 (1985) [Sov. Phys. Solid State **27**, 1723 (1985)].
[14] P. Pecheur and G. Toussaint, Phys. Lett. A **135**, 223 (1989).
[15] G.A. Thomas, D.H. Rapkine, R.B. Van Dover, L.F. Mattheiss, W.A. Sunder, L.F. Schneemeyer, and J.V. Waszczak, Phys. Rev. B **46**, 1553 (1992).
[16] S.K. Mishra, S. Satpathy, and O. Jepsen, J. Phys.: Condens. Matter **9**, 461 (1997).
[17] P. Larson, S. Mahanti, and M.G. Kanatzidis, Phys. Rev. B **61**, 8162 (2000).
[18] E. Wimmer, H. Krakauer, M. Weinert, and A.J. Freeman, Phys. Rev. B **24**, 864 (1981); M. Weinert, E. Wimmer, and A.J. Freeman, *ibid.* **26**, 4571 (1982).
[19] S. Nakajima, J. Phys. Chem. Solids **24**, 479 (1963).
[20] A.H. MacDonald, W.E. Pickett, and D.D. Koelling, J. Phys. C **13**, 2675 (1980).
[21] L. Hedin and B.I. Lundqvist, J. Phys. C **4**, 2064 (1971).
[22] H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976).
[23] G.D. Mahan, J. Appl. Phys. **65**, 1578 (1989).