RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 32, NUMBER 10 15 NOVEMBER 1985

Schottky barriers and semiconductor band structures

J. Tersoff
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 8 August 1985)

Various models of Schottky-barrier formation suggest Fermi-level pinning in midgap. Elementary band-structure considerations indicate that, for diamond-structure semiconductors, the physically relevant gap is the *indirect* gap, corrected for spin-orbit splitting. Schottky-barrier heights for elemental and III-V compound semiconductors can be predicted to 0.1 eV from *measured* indirect gaps and splittings. The dimensionless pinning strength $\bar{S}$ is given by the optical dielectric constant. Chemical trends are thus simply explained.

After decades of intensive study, even the most elementary aspects of Schottky-barrier formation at metal-semiconductor interfaces remain controversial. Many different theoretical models have been proposed.[1-8] Attempts to correlate barrier heights with other measurable quantities have revealed many suggestive qualitative regularities,[9-12] without elucidating the underlying mechanisms.

The Schottky-barrier height is determined by the position of the Fermi level (E_F) relative to the local semiconductor band gap at the interface. Experimentally, barrier heights are known often to depend only weakly on the metal used, i.e., E_F is "pinned" relative to the semiconductor. A previous paper[8] suggested a specific microscopic model of barrier formation, which led to quantitative predictions of barrier heights in the limit of strong pinning (barrier independent of metal). That work, however, did not address the issue of how strong the proposed pinning mechanism was, a crucial and controversial point, and it related the barrier height to the semiconductor band structure in a complicated and nonintuitive way.

Ideally, one would like to identify a few material parameters, all experimentally measureable, which are sufficient both to predict the barrier height and to quantify the degree of pinning. This goal has, however, proven elusive, and in recent years has essentially been abandoned. The purpose of this paper is to propose specific *quantitative* relationships between barrier heights and *measured* bulk semiconductor band-structure properties, motivated by the most elementary theoretical considerations. These semiempirical relationships not only predict Schottky-barrier heights to 0.1 eV with a single fitted parameter; they also yield new insight into the crucial role of direct versus indirect gaps in determining barrier heights. The degree of "pinning" can be simply related to the screening of a (pseudo)dipole at the interface, given directly by the optical dielectric constant. The success of this approach appears to confirm that barrier formation is associated with states intrinsic to the interface. An understanding of chemical trends in barrier height follows naturally from these results.

Surprisingly, the only attempt until recently[8] to correlate barrier heights directly with bulk band structure was that of Mead and Spitzer.[9] They proposed, on strictly empirical grounds, that the p-type barrier ϕ_{bp} was given by

$$\phi_{bp} \equiv E_F - E_V = E_g/3 \ , \qquad (1)$$

where E_g is the semiconductor band gap, and E_V is the valence-band maximum.

Unfortunately, this simple relationship is not borne out in detail by experimental data. Figure 1(a) shows the level of agreement between (1) and experiment. For simplicity and consistency, we use barrier heights for Au on each semiconductor, taken from a standard reference.[13] (All covalent and III-V semiconductors for which data are given in Ref. 13 are included.) While (1) correctly describes the overall trend, it has a large rms error of 0.20 eV, with errors up to 0.35 eV in individual cases. More important, it fails *qualitatively* in many cases. For InAs, E_F falls *above* the top of the gap, while for GaSb and InSb it falls at or near the bottom. Similar, though less drastic, deviations occur for other semiconductors as well. To understand the qualitative success and quantitative failure of (1), one must consider possible mechanisms of barrier formation.

With a few exceptions,[5,6] theoretical models of Schottky-barrier formation have been based on the original suggestion of Bardeen,[3] that E_F is pinned at the interface by states in the semiconductor band gap. The nature of these states

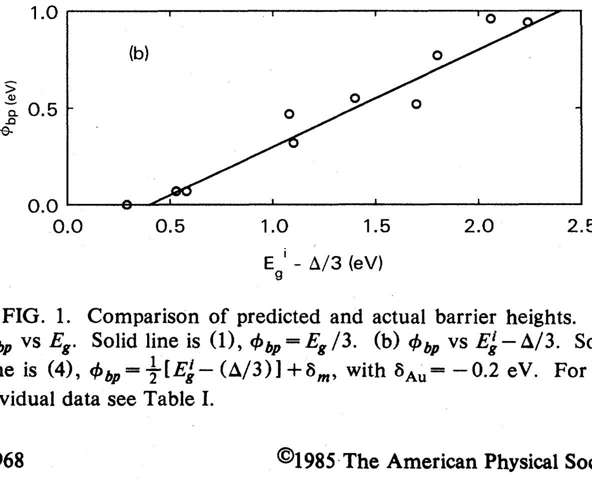

FIG. 1. Comparison of predicted and actual barrier heights. (a) ϕ_{bp} vs E_g. Solid line is (1), $\phi_{bp}=E_g/3$. (b) ϕ_{bp} vs $E_g^i-\Delta/3$. Solid line is (4), $\phi_{bp}=\frac{1}{2}[E_g^i-(\Delta/3)]+\delta_m$, with $\delta_{\mathrm{Au}}=-0.2$ eV. For individual data see Table I.

remains the subject of conjecture, since they are experimentally very difficult to measure. Fermi-level pinning has been variously attributed to surface states,[3] intrinsic interface states,[4,8] or any of a variety of defects[7] in the semiconductor.

Appelbaum and Hamann[14] have pointed out that, in one dimension, surface states which pin E_F (i.e., which are neutral when half filled) tend to fall near the branch point in the complex band structure, because of simple electrostatic and band-structure considerations. The branch point, in turn, tends to fall near the center of the band gap (in one dimension).[15] This energy similarly plays a crucial role in a recent model of the metal-semiconductor interface,[8] and perhaps in the theory of defect levels. It is therefore natural to speculate that the Fermi level should be pinned in the center of the band gap. This suggestion closely resembles (1) and, taken naively, is no more successful.

In fact, in three dimensions the center of the gap does not appear to have any special properties. Because of the overwhelming importance of shallow dopant levels in semiconductors, attention has naturally focused on the band edges. However, states which lie deep in the gap are spatially rather localized, and therefore sample a substantial portion of the Brillouin zone, not just the band edge. The states which are believed[3,4,7,8] to determine barrier heights are certainly deep levels in this sense, whatever their detailed nature and origin. These levels cannot be described by an effective-mass approximation, and need bear no simple relationship to the band edges. The task then is to identify some relevant *average* gap center.[8]

An examination of semiconductor band structures[16,17] reveals some important points. The Γ point at the zone center determines the gap in many III-V semiconductors, yet its energy bears little relation to the conduction band as a whole, because it is associated with a symmetry-induced decoupling of s and p states. As a result, it has a very small effective mass, so there is little **k** space associated with this minimum. Moreover, its strict s symmetry further excludes it from contributing to the gap states, which are largely p-like.[18] It is therefore probably safe to ignore the Γ minimum in discussing deep levels, at least as a first approximation. The indirect minima on or near the zone faces are far more characteristic of the conduction band, and, because of their degeneracy and large effective masses, they describe a relatively large region of **k** space. In suggesting that E_F should fall in the middle of the gap, one should clearly refer to the *indirect* gap, regardless of the energy of the Γ minimum.

In the absence of splitting, the threefold degenerate valence maximum adequately characterizes the valence bands. However, the spin-orbit interaction splits these states by an amount Δ at Γ, while leaving the valence band as a whole relatively unaffected. Thus, relative to the valence band, the twofold degenerate valence-band maximum is pushed up in energy by an amount $\frac{1}{3}\Delta$, while the split-off state is pushed down by $\frac{2}{3}\Delta$. It is therefore most consistent in this context to consider not the actual valence band maximum (E_V), but its position in the absence of spin-orbit splitting,

$$\bar{E}_V = E_V - \tfrac{1}{3}\Delta \ .$$

Thus, by considering (diamond-structure) semiconductor band structures in a general way, one arrives as the conclusion that the most physical simple definition of the gap center E_0 is

$$E_0 = \tfrac{1}{2}(\bar{E}_V + \bar{E}_c) \ , \tag{2}$$

where $\bar{E}_c$ is the *indirect* conduction minimum, i.e., excluding Γ. A better criterion would probably require more detailed knowledge of the band structure.[8]

It is tempting to equate this gap center with E_F, as discussed above. However, two factors have so far been neglected. First, barrier heights are known to depend somewhat on the metal, primarily through its electronegativity.[10,19] One ought therefore to allow for some shift from the gap center, depending upon the metal. The simplest assumption is $E_F = E_0 + \delta_m$, where δ_m depends only upon the metal. Second, the definition of the effective gap here is necessarily somewhat arbitrary. In one dimension, the branch point in the complex band structure generally falls slightly *below* the center of the gap, owing to the underlying parabolic dispersion.[15] Moreover, (2) is obviously rather crude. It is therefore convenient to treat δ_m as an adjustable parameter, which can absorb some of the arbitrariness of the definition of E_0 above. At least δ_{Au} should be negative, both because Au is highly electronegative, and because the branch point falls below midgap, and δ_m should be small compared to E_g; otherwise the present treatment is not well justified.

We are therefore lead naturally to the equation

$$E_F = \tfrac{1}{2}(\bar{E}_V + \bar{E}_c) + \delta_m \ . \tag{3}$$

This may be related directly to the barrier height,

$$\phi_{bp} = \frac{1}{2}\left[E_g^i - \frac{\Delta}{3}\right] + \delta_m \ , \tag{4}$$

where $E_g^i \equiv \bar{E}_c - E_V$ is the minimum *indirect* gap. Experimental values of ϕ_{bp}, E_g^i, and Δ at room temperature are given in Table I.

A comparison of (4) with experiments is shown in Fig. 1(b) and Table I. The effective gap center E_B defined in Ref. 8 is also given for comparison. The choice of δ_m which best fits the data for Au is $\delta_{\text{Au}} = -0.2$ eV. With this single parameter, the rms error is remarkably small: 0.07 eV. The *maximum* error is only 0.13 eV. In view of the uncertainties

TABLE I. Semiconductor properties: band gap[a] E_g, minimum indirect gap[a,b] E_g^i, spin-orbit splitting[a] Δ, predicted barrier height $\phi_{bp}^{\text{theory}} = (E_g^i - \Delta/3)/2 + \delta_{\text{Au}}$, experimental barrier[c] $\phi_{bp}^{\text{expt}}(\text{Au})$, and calculated effective midgap energy E_B of Ref. 8.

	E_g	E_g^i	Δ	$\phi_{bp}^{\text{theory}}$	ϕ_{bp}^{expt}	E_B
Si	1.11	1.11	0.04	0.35	0.32	0.36
Ge	0.66	0.66	0.29	0.08	0.07	0.18
GaP	2.27	2.27	0.08	0.92	0.94	0.81
InP	1.34	1.84	0.11	0.70	0.77	0.76
AlAs	2.15	2.15	0.28	0.83	0.96	1.05
GaAs	1.43	1.81	0.34	0.65	0.52	0.70
InAs	0.36	1.21	0.39	0.34	0.47	0.50
AlSb	1.63	1.63	0.70	0.50	0.55	· · ·
GaSb	0.70	0.80	0.75	0.07	0.07	0.07
InSb	0.18	0.62	0.98	0.00	0.00	· · ·

[a] Reference 17. [b] Reference 20. [c] Reference 13.

both in the barrier heights and in the indirect minima, better agreement woud probably be fortuitous.

Equation (4) should be viewed as predictive in that, given any large subset of the data, one could determine δ_m and thereby predict the remaining data to within 0.1 eV. Despite its phenomenological nature, (4) is more accurate than any theory to date, except perhaps that of Ref. 8 (which had no adjustable parameters). Moreover, it makes testable predictions about the dependence of barrier height on alloy composition, temperature, hydrostatic pressure, etc.

Equally important, however, is the insight into trends in barrier height which (4) provides. The Γ minimum is particularly sensitive to the cation, and decreases sharply in the sequence Al$\rightarrow$Ga$\rightarrow$In. If Γ dips very low in energy relative to the rest of the conduction band, E_F will fall high in the gap. This is the case for InAs and, to a lesser extent, InP. The point is that E_F is not anomalously high in the gap, but rather the conduction minimum at Γ is anomalously low relative to the physically relevant indirect minimum and to the conduction band overall. On the other hand, if the spin-orbit splitting is very large, as for the antimonides, then the valence-band maximum is pushed up in energy relative to the valence band as a whole, so E_F appears to fall low in the gap. This helps explain why GaSb and InSb have E_F pinned very low in the gap, compared with other semiconductors.

So far only one metal has been considered, in order to emphasize variations in barrier height with semiconductor. A great deal of effort, however, has gone into understanding the variation of barrier height with metal, for a given semiconductor. The barrier has been correlated with many properties of the metal,[2,6,10,12] but attempts at a *predictive* correlation have had limited success.[19]

It is common to analyze the degree of pinning of E_F for a semiconductor in terms of a parameter

$$S \equiv -\frac{d\phi_{bp}}{d\chi_m} , \tag{5}$$

where χ_m is the metal electronegativity. This equation provides at best an incomplete description of the effect of metal on barrier height,[19] but is nevertheless useful in organizing a large body of data. The basic idea[19,21] is that as the metal Fermi level is changed, this change is screened by states in the semiconductor, reducing the effect on the barrier by a factor $\bar{S} = S/A$, where $A \simeq 2.8$ eV serves to convert conventional electronegativity units into absolute energy.[19] $\bar{S}$ is thus dimensionless, and approaches unity in the limit of noninteracting metal and semiconductor (the unpinned "Schottky limit").

In reality, increasing the metal valence density lowers the potential in the metal while increasing E_F relative to the average potential. The net effect is a modest lowering of E_F, which is reflected in the increased electronegativity. Also, S-d overlap plays an important role in transition and noble metals. Equation (5) implicitly treats these various factors as merely a shift in the potential in the metal, in the spirit of a pseudopotential approximation.

Taking this simplistic view to its logical conclusion, one may view an increase in metal electronegativity as corresponding to an extra constant attractive potential in the metal, i.e., to a step potential with the step falling at the metal-semiconductor interface. This is formally identical to a sheet dipole at the interface. The parameter $\bar{S}$ is then simply and exactly the degree to which this dipole sheet is screened. If the dipole fell in the metal, one would have $\bar{S} = 0$. If the dipole fell in the semiconductor, $\bar{S} = \epsilon_\infty^{-1}$ would result, where ϵ_∞ is the optical dielectric constant. For a sufficiently "intimate" interface, with good metal-semiconductor bonding, it seems reasonable to assume that the total screening is a monotonic function of dipole position. This implies

$$0 \leq \bar{S} \leq \epsilon_\infty^{-1} . \tag{6}$$

The obvious estimate is, then, the mean value

$$\bar{S} \simeq \tfrac{1}{2}\epsilon_\infty^{-1} . \tag{7}$$

It is important to use ϵ_∞ instead of the static dielectric constant ϵ_0, which includes an ionic contribution. The electronegativity is determined by the kinetic energy as much as by the potential. The ions will therefore not feel the same effective (pseudo)potential step at the interface as will the electrons. It is worth stressing that ϵ_∞, like the indirect gap, is strictly a band-structure property,[22] and is closely related to the average gap.

The available data for S have been analyzed by Schluter,[19] with estimated accuracies typically ±20%. In Fig. 2 $\bar{S}$ is plotted versus ϵ_∞^{-1}. All data for diamond-structure semiconductors in Ref. 19 are included. The result is striking. All of the data points which show strong pinning ($\bar{S} < 0.1$) fall very near the simple estimate (7). This suggests that the pinning is associated with states *intrinsic* to the interface,[4,8,21] since other mechanisms should not show this specific correlation with ϵ_∞.

Equally important, all of the data points with weak pinning ($\bar{S} > 0.2$) lie near or *above* the upper bound (6). This indicates an incipient breakdown in the assumption that the dielectric response varies smoothly and monotonically across the interface, and suggests that poor interfacial bonding may play a crucial role in determining the weak pinning in these interfaces. [It is, however, puzzling that GaP and CdS do not obey (7) well, while similar compounds do.]

It should be emphasized that the association of a unique number S with each semiconductor has only semiquantitative validity, and there is some risk in too detailed an analysis. Nevertheless, a simple analysis in terms of bulk dielectric properties predicts specific values for $\bar{S}$, which are

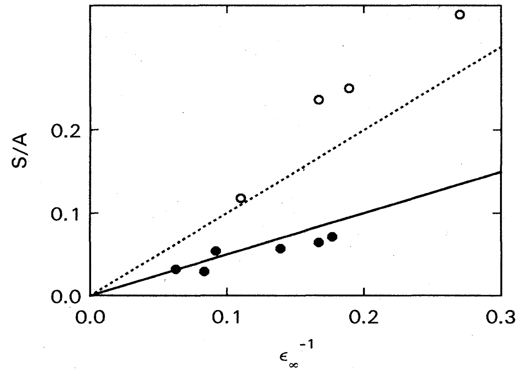

FIG. 2. Pinning strength $\bar{S} = S/A$ (Ref. 19) vs ϵ_∞^{-1}. Solid line is theoretical estimate (7); dashed line is upper bound from (6). Solid circles are cases with strong pinning ($\bar{S} < 0.1$) (left to right): Ge, Si, GaAs, CdTe, CdSe, C; open circles are GaP, ZnSe, CdS, ZnO.

in good agreement with experiment for all strongly pinned interfaces. The trivial upper bound $\bar{S} = \epsilon_\infty^{-1}$ describes well the more weakly pinned interfaces. The analysis also suggests that the qualitative separation into two classes of semiconductors, emphasized by Kurtin, McGill, and Mead,[10] may have its physical origin in the poor interfacial bonding between metals and the more ionic semiconductors and insulators.

Schottky-barrier heights, and their trends with semiconductors and metals, have long been the subject of both fundamental and phenomenological studies. However, the potential of detailed theoretical analysis to predict specific quantitative correlations has not yet been exhausted. Barrier heights and pinning strengths show extremely simple trends with semiconductor band structure, once the physically relevant factors are identified. This simplicity should prove of critical importance in both the fundamental and phenomenological understanding of interface behavior.

It is a pleasure to thank D. E. Aspnes for suggesting the possible role of indirect gaps, and S. M. Kelso for helpful information on indirect gaps.

---

[1] For a brief review, see, for example, M. Schluter, Thin Solid Films **93**, 3 (1982).

[2] W. Schottky, Phys. Z. **41**, 570 (1940).

[3] J. Bardeen, Phys. Rev. **71**, 717 (1947).

[4] V. Heine, Phys. Rev. **138**, A1689 (1965); F. Yndurain, J. Phys. C **4**, 2849 (1971); S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976); C. Tejedor, F. Flores, and E. Louis, J. Phys. C **10**, 2163 (1977).

[5] J. C. Inkson, J. Vac. Sci. Technol. **11**, 943 (1974).

[6] J. L. Freeouf and J. M. Woodall, Appl. Phys. Lett. **39**, 727 (1981).

[7] W. E. Spicer, I. Lindau, P. R. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Lett. **44**, 420 (1980); W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979).

[8] J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984). For a more complete tabulation of results, see also Phys. Rev. B **30**, 4874 (1984).

[9] C. A. Mead and W. G. Spitzer, Phys. Rev. **134**, A713 (1964).

[10] S. Kurtin, T. C. McGill, and C. A. Mead, Phys. Rev. Lett. **22**, 1433 (1969).

[11] J. O. McCaldin, T. C. McGill, and C. A. Mead, Phys. Rev. Lett. **36**, 56 (1976); J. Vac. Sci. Technol. **13**, 802 (1976).

[12] L. J. Brillson, Phys. Rev. Lett. **40**, 260 (1978).

[13] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1969). The relationship $\phi_{bp} = E_g - \phi_{bn}$ is used where appropriate in extracting p_{bp} from this tabulation.

[14] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B **10**, 4973 (1974).

[15] E. T. Goodwin, Proc. Cambridge Philos. Soc. **35**, 205 (1935).

[16] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).

[17] *Landolt-Bornstein: Numerical Data and Functional Relationships in Science and Technology,* edited by O. Madelung (Springer-Verlag, Berlin, 1982), Group 3, Vol. 17, Part a.

[18] N. V. Dandekar, A. Madhukar, and D. N. Lowy, Phys. Rev. B **21**, 5687 (1980).

[19] M. Schluter, Phys. Rev. B **17**, 5044 (1978).

[20] S. M. Kelso, D. E. Aspnes, C. G. Olson, D. W. Lynch, and K. J. Bachmann, Proc. Soc. Photoopt. Instrum. **452**, 100 (1984), and references therein; S. M. Kelso (private communication).

[21] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B **15**, 2154 (1977).

[22] H. Ehrenreich and M. H. Cohen, Phys. Rev. **115**, 786 (1959).