Semi-empirical descriptions of temperature dependences of band gaps in semiconductors

Roland Pässler

Institut für Physik, Technische Universität Chemnitz, 09107 Chemnitz, Germany

Received 22 May 2002, revised 2 January 2003, accepted 8 January 2003
Published online 3 April 2003

PACS 71.20.Mq, 71.20.Nr, 71.35.–y

Starting from general foundations of the semi-empirical theory of monotonic temperature dependences of band gaps in semiconductors, we devote ourselves in this article to further analytical development and applications of the power function model. On the basis of a detailed analytical study of asymptotic low and high temperature behaviors, including numerical comparisons for intermediate temperatures, we devise a novel analytical (higher order root) expression. Applications to least-mean-square fittings of fundamental band gap data available for a variety of group IV, III–V, and II–VI materials are shown to provide, both for regimes of moderate and large phonon dispersion, self-consistent values for basic dispersion-related parameters. This is essential for detailed comparisons of the power function model with other elaborate semi-empirical models developed in last years, which differ above all by qualitatively different types of low temperature asymptotes. An exemplification for bulk samples of Si and GaAs shows that it is rather difficult to distinguish clearly between \(T \to 0\) asymptotes, \([E(0) - E(T)] \propto T^p\), that are controlled by system-specific fractional exponents \((2 < p < 3)\) and other dispersion-related descriptions that admit exclusively quadratic asymptotes \((p = 2)\). Novel high-precision data for single quantum dot luminescence peak positions are found to give us better chances for distinguishing between qualitatively different types of low temperature asymptotes.

1 Introduction

The temperature dependence of the fundamental band gap widths, \(E_g(T)\), of semiconductors and insulators (wide band gap materials) represents a basic material-specific property which is of considerable theoretical as well as practical interest. It is known from a wealth of experimental results [1] that the band gap widths of many materials decrease monotonically with increasing temperature. These dependences are usually found to be relatively weak (quadratic or cubic) in the cryogenic region and strong (approaching linear asymptotes) above Debye temperature, \(T > \Theta_D\) [1]. Various formulas have been proposed and repeatedly used during the last three decades for approximate numerical simulations of such \(E_g(T)\) dependences. Frequently used in experimental papers has been Varshni’s formula [2, 3] (cf. Sec. 5), which represents a special type of combination of a linear high-temperature dependence with a certain quadratic low-temperature asymptote. Many least-mean-square fittings based on this formula, however, revealed fundamental deficiencies. These are manifested above all as (1) systematic deviations between measured and calculated \(T\)-dependences, particularly in the cryogenic region [4–9], and (2) a high degree of instability (arbitrariness) of Varshni’s parameter sets [8–10]. Various analytical aspects of these characteristic deficiencies have already been discussed in considerable detail in prior papers [7,
The general inadequacy of Varshni’s model shows up particularly in large differences between numerical estimates for $T = 0$ renormalization energies (see Sec. 2), that are resulting from alternative least-mean-square fitting procedures using Varshni’s formula [2, 3] (see Sec. 5), on the one hand, and physically reasonable $E(T)$ models (see Sections 2 to 4 and cf. [11]) or the isotopic substitution effect [12], on the other hand. We find here, among other things, that the notorious overestimation of this fundamental quantity by Varshni’s model amounts e. g. for silicon to a factor larger than 5. In the case of a high-precision data set obtained recently for an In$_{0.4}$Ga$_{0.6}$As/GaAs quantum dot (QD) system [13] via single QD photoluminescence spectroscopy measurements [14] (limited to $2 \leq T \leq 100$ K), the overestimation of the $T = 0$ renormalization energy by Varshni’s formula increases dramatically up to a factor of order 10$^3$. This is not surprising, however, because the eventuality even of such huge numerical discrepancies is anticipated as a casual consequence of the fact [11] that Varshni’s model [2, 3] refers, implicitly, to a hypothetical regime of excessively large dispersion of phonon energies (see Sec. 5 and cf. [11]), which has never been found in experiment. Discrepancies of this type provide the general motivation for devising physically more reasonable (theoretically justifiable) alternative models.

A basic alternative to Varshni’s ad hoc model had been invoked, already years ago, in the form of model equations of Bose–Einstein type [15–24] (see also [5–12]). Such an ansatz is physically more reasonable than Varshni’s formula because it accounts at least for the well known proportionality between the average number of phonons in a given mode, on the one hand, and its contribution to the observable $E(T)$ dependence, on the other hand. However, all formulas of Bose–Einstein type correspond to the limiting regime of vanishing dispersion. Among the multitude of $E(T)$ dependences found in experimental papers one can find special cases (e. g. certain ternary compounds [25] or MQW structures [26, 27]) where the dispersion is found to be in fact rather small. In such cases one can get reasonable least-mean-square fits even with a limiting expression of Bose–Einstein type. For most materials, however, the actual degree of phonon dispersion is found to range fairly between the hypothetical regime of extremely large dispersion associated with Varshni’s formula [2, 3] and the limiting regime of completely vanishing dispersion represented by expressions of Bose–Einstein type [5–12 and 15–24]. A proper interpretation of the corresponding $E(T)$ data sets requires, consequently, more general and flexible analytical models which account in particular for significant variations of the degree of phonon dispersion from material to material.

We have presented in recent years several dispersion-related models [4, 6, 10, 26, 28, 29] that use to provide acceptable least-mean-square fittings of monotonic $E(T)$ dependences in a variety of semiconductor materials. Among the qualitatively different analytical approaches published, the structurally simplest $E(T)$ formula is based on a power function ansatz [6] (see Sec. 3) for the electron–phonon spectral function. It has the form [4, 9]

$$E(T) \approx E(0) - \left(\frac{\alpha \Theta_p}{2}\right) \left(1+\frac{2T}{\Theta_p}\right)^{\frac{1}{p}} - 1$$

In this expression the parameter $\alpha = S(\infty)$ represents the $T \to \infty$ limit of the slope, $S(T) = -dE(T)/dT$ (=entropy [3, 5]), $\Theta_p$ is comparable with (roughly equal to) the average phonon temperature [6, 9, 28], $\Theta_p \approx \Theta = \langle \hbar \omega \rangle / k$, and the empirical (fractional) exponent $p$ decreases monotonically [9] with increasing dispersion. It has been found that the model function (1) provides best fits of measured $E(T)$ data sets for a variety of elemental and binary materials [9] as well as ternary compounds [8, 25] and MQW structures [7, 26, 27]. However, due to the approximate nature of this simple model function (1), the fitted $\Theta_p$ values [9] may differ significantly in various cases (up to about 20% in the regime of small dispersion, $p > 3$) from the actual magnitudes of average phonon temperatures, $\Theta = \langle \hbar \omega \rangle / k$ (that can be determined by means of a more elaborate, self-consistent $E(T)$ expression; cf. [11, 28, 29] and see Secs. 3 and 4). Unsatisfactory, particularly from basic theoretical points of view, is the circumstance that the empirical $\Theta_p$ values [9] due to Eq. (1) are not accurate enough for reliable determinations of zero-point ($T = 0$) renormalization energies [11, 12].
It is possible, of course, to come to theoretically consistent $\Theta$ values when we a priori avoid the use of an approximate analytical $E(T)$ expression (like (1)) and perform, instead, numerical calculations of the original $E(T)$ integrals (cf. Sec. 2). Examples of successful least-mean-square fittings involving such numerical integration procedures have been presented in Refs. [26] and [30] for exciton peak energy shifts in a CdTe/ZnTe multiple quantum well structure and a CuI quantum dot system, respectively. However, the corresponding numerical analysis procedures are computationally relatively expensive and can thus hardly be considered as being ready for current analyses of $E(T)$ data sets within the frame of experimental studies. The existing computational problems and/or numerical uncertainties encountered hitherto in applications of the power function model motivate us to device in this study an analytical $E(T)$ expression more elaborate than Eq. (1).

An additional motivation for the present study was provided by the successful construction of a duly elaborate analytical (sixth-order root) $E(T)$ expression in a recent paper [11], which refers to a more easily tractable spectral function model (envisaged for the first time in Ref. [6]). This model involves, for all cases of non-vanishing dispersion, a strictly quadratic $T \to 0$ asymptote whose actual weight is a function of the material-specific degree of dispersion. This model is qualitatively in accordance with corresponding theoretical results presented many years ago in several theoretical papers [31–33], where the occurrence of certain quadratic low temperature asymptotes has been predicted. In contrast to this, however, Cardona [12] has given recently some global theoretical arguments in favor of cubic $T \to 0$ asymptotes (in conjunction with a denial of the notion of quadratic asymptotes as seemingly incompatible with first principles considerations). Unfortunately, sufficiently detailed analytical descriptions of electron-phonon spectral functions in the immediate vicinity of the zero phonon energy limit, $\hbar \omega_{\nu,\alpha,\alpha} \to 0$, do not seem to be available hitherto. The theoretical problem of an unambiguous determination of the exponent $p$ governing the actual $T \to 0$ power dependences, $[E(0) - E(T)] \propto T^p$, can thus not yet been considered as definitively solved, neither theoretically nor experimentally. In view of the still existing ambiguities it might thus be rather useful to dispose of a theoretical model which offers a certain chance to distinguish, via numerical fittings of high-precision $E(T)$ data, between the traditional notion of quadratic asymptotes ($p = 2$; cf. [6, 10, 11, 26, 28, and 31–33]) and the possible alternative of higher order power dependences ($p > 2$; cf. [6, 9, 12]). Among the semi-empirical, dispersion-related $E(T)$ models considered hitherto, the power-function model is obviously the only one that gives us a chance to distinguish in one or the other case between qualitatively different $T \to 0$ asymptotes. Under this perspective it is timely now to tackle the task of advising (in analogy to [11]) a sufficiently elaborate, self-consistent analytical $E(T)$ expression even for the power function model.

The analytical foundations of the semi-empirical, dispersion-related theory [11, 28] of monotonic $E(T)$ dependences are briefly sketched in Sec. 2. We specialize in Sec. 3 the general equations for a continuous weighting function given as a power dependence [6, 26] with a variable (fractional) exponent. On the basis of a detailed analytical and numerical study of the corresponding dispersion-related shape functions (in appendix A) we are able to construct a sufficiently elaborate $E(T)$ expression suitable for accurate determinations of the parameter sets. Corresponding numerical results following from least-mean-square fittings of experimental $E_g(T)$ and/or $E_o(T)$ data sets available for group IV, III–V, and II–VI materials, including high-precision data obtained recently by Ortner et al. [14] for single quantum dot luminescence peak positions for an In$_{0.5}$Ga$_{0.5}$As/GaAs self-assembled QD system [13], are presented and discussed in Sec. 4.

2 Basic equations of the semi-empirical theory of monotonic $E(T)$ dependences

The temperature dependences of energy gaps in semiconductors are well known to be the cumulative effect of electron–phonon interaction (EPI) and thermal lattice expansion (TLE) mechanisms (see [28] and articles quoted therein). Of considerable importance for a general formulation of a semi-empirical theory is the observation [21, 28] that, for both types of qualitatively different gap shrinkage mecha-
nisms, the contributions of the individual lattice oscillators with phonon energies $\varepsilon = \varepsilon_j = \hbar \omega_j$ (j = branch, $q$ = wave vector) are proportional to the thermal average values of phonon occupation numbers, $\bar{\pi}(\varepsilon; T) = \left[ \exp \left( \varepsilon / kT \right) - 1 \right]^{-1}$. Consequently, in confining ourselves to the experimentally most important class of materials with monotonically decreasing energy gaps, we can represent the corresponding $E(T)$ dependences by an integral of the general form [28, 29, 11]

$$E(T) = E(0) - \frac{\alpha}{k} \int d\varepsilon \frac{\varepsilon w(\varepsilon)}{\exp(\varepsilon/kT) - 1}. \quad (2)$$

Here we have denoted as usually by $\alpha$ the limiting magnitude of the slope, $\alpha = -dE(T)/dT|_{T \to \infty}$, and $w(\varepsilon)$ represents a positive semi-definite weighting function, $w(\varepsilon) \geq 0$. The latter is generally normalized to unity, $\int d\varepsilon \ w(\varepsilon) = 1$. (Note that this normalization is due to its definition, $w(\varepsilon) = (k/\alpha) \cdot f(\varepsilon)/\varepsilon$, on the basis of the underlying overall spectral function [28, 11], $f(\varepsilon) = f^{\text{EPI}}(\varepsilon) + f^{\text{TEL}}(\varepsilon) \geq 0$). The actual $\varepsilon$-dependences of weighting functions, $w(\varepsilon)$, are throughout highly complicated and change significantly from material to material. Detailed numerical information on the fine structure of weighting functions, $w(\varepsilon)$, does not use to be available.

Fortunately it possible in many cases to reliably determine at least the first and second order moments of weighting functions,

$$\langle \varepsilon^m \rangle = \int d\varepsilon \ \varepsilon^m w(\varepsilon), \quad m = 1 \text{ and } 2, \quad (3)$$

via numerical fittings of measured $E(T)$ dependences. For this purpose it is convenient to approximate the integral representation (2) by a practicable analytical expression involving parameters that are connected in unambiguous way with these two dispersion-related moments (3), $\langle \varepsilon^1 \rangle$ and $\langle \varepsilon^2 \rangle$, of the phonon energy spectrum.

It is usual to represent the position of the center of gravity, $\langle \varepsilon \rangle$, on a temperature scale by an effective (average) phonon temperature [6, 28],

$$\Theta = \langle \varepsilon \rangle / k, \quad (4)$$

and the material-specific degree of phonon dispersion by the dispersion coefficient [28],

$$\Delta = \sqrt{\langle \varepsilon^2 \rangle - \langle \varepsilon \rangle^2} / \langle \varepsilon \rangle. \quad (5)$$

One can easily see that a total set of four empirical parameters, which includes the effective phonon temperature, $\Theta$ (4), and the dispersion coefficient, $\Delta$ (5), in combination with the limiting slope, $\alpha$, and the $T \to 0$ energy position, $E(0)$, determines in unambiguous way the $E(T)$ dependence in regions of elevated temperatures (usually near and above room temperature). To this end we consider the argument $x = \varepsilon/kT$ of the exponential function in the denominator of the integral (2) to be throughout smaller than $2\pi$. (Note that this is the case for $T > \Theta_c/2\pi$, where $\Theta_c = \varepsilon_c/k$ is the phonon temperature associated with the cut-off energy, $\varepsilon_c$, of the phonon spectrum). Using the corresponding Taylor expansion, $x(\varepsilon^m - 1) = 1 - x/2 + x^2/12 - ...$ (for $0 < x < 2\pi$), and observing the definitions (4) and (5) for the empirical parameters $\Theta$ and $\Delta$ we reduce thus the integral expression (2) to a truncated (low order) series expansion of the universal form [28]

$$E(T) = E(0) - \frac{\alpha}{k} \left\{ T - \frac{\Theta}{2} + \frac{(1 + \Delta^2) \Theta^2}{12 T} - ... \right\}. \quad (6)$$
In the high temperature limit, \( T \gg \Theta \), this expression (6) tends to the linear asymptote
\[
E(T) \to E_{\text{lim}}(T) = E(0) - \alpha (T - \Theta / 2) = E_{\text{lim}}(0) - \alpha T ,
\]
where we have denoted by \( E_{\text{lim}}(0) = E(0) \pm \alpha \Theta / 2 \) the point of intersection of this linear asymptote (7) with the energy axis. The energy distance between \( E_{\text{lim}}(0) \) and the measured \( T \to 0 \) energy position, \( E(0) \), corresponds to the \( T = 0 \) renormalization energy \([11, 12]\).

\[
E_{\text{lim}}(0) - E(0) = \alpha \Theta / 2
\]
(cf. Fig. 1 in \[28\]). Thus an accurate determination both of the limiting slope, \( \alpha \), and of the average phonon temperature, \( \Theta (4) \), are the prerequisites also for a reliable estimation of the associated \( T = 0 \) renormalization energy \([11, 12]\) (see below).

### 3 Power function model

A series expansion of type (6) is, in view of its general limitation to a certain region of elevated temperatures \( (T > \Theta / 2 \pi) \), generally unsuitable for the purpose of numerical determinations of the whole set of empirical parameters, \( E(0) \), \( \alpha \), \( \Theta \), and \( \Delta \). It is a matter of principle that a simultaneous determination of all four parameters can be performed only on the basis of an analytical expression that is applicable also to lower temperatures including the \( T \to 0 \) limit \((0 \leq T < \Theta / 2 \pi)\). A suitable analytical continuation of Eq. (6) into the cryogenic region can be constructed via a physically reasonable model ansatz for the weighting function \( w(\varepsilon) \) in Eqs. (2) and (3).

We have seen in Ref. \([9]\) that good fittings of measured \( E(T) \) dependences can be provided, as a rule, when we envisage for the low temperature asymptote some power dependence, \((E(0) - E(T)) \propto T^p\), with an exponent \( p \) of order 2 to 3 \([9]\). Such a behavior can be modeled by an ansatz for the weighting function \( w(\varepsilon) \) in the form of a power dependence \([6, 26]\),

\[
w_\nu(\varepsilon) = \frac{\nu}{\varepsilon_c^\nu - \varepsilon^\nu} ; \quad \nu > 0 ,
\]
for \( \varepsilon \) up to a certain cut-off energy, \( 0 \leq \varepsilon \leq \varepsilon_c \), and \( w_\nu(\varepsilon) = 0 \) elsewhere. The corresponding model-specific moments (3) are \( \langle \varepsilon^m \rangle = \nu \varepsilon_c^\nu / (\nu + 1) \), i.e. in particular \( \langle \varepsilon \rangle = \nu \varepsilon_c^1 / (\nu + 1) \) and \( \langle \varepsilon^2 \rangle = \nu \varepsilon_c^2 / (\nu + 1) \), for the first and second moment, respectively. The effective phonon temperature (4) is thus related to the cut-off phonon temperature \( \Theta_c = \varepsilon_c / k \) by

\[
\Theta = \nu \Theta_c / (1 + \nu)
\]
and the dispersion coefficient (5) is given in terms of the curve shape parameter \( \nu > 0 \) by

\[
\Delta = 1 / \sqrt{(2 + \nu)}
\]
(cf. the curve Fig. in 1). The subclasses of monotonically increasing and decreasing weighting functions (9) are generated by variations of the curve shape parameter \( \nu \) within the intervals \( 0 < \nu < 1 \) and \( 0 < \nu < 1 \), respectively. From (11) we see that these subclasses correspond to intervals of \( 0 < \Delta < 3^{1/2} \) and \( 3^{1/2} < \Delta < \infty \), respectively, for the dispersion coefficient (cf. Fig. 1). These dispersion regimes can be described as regions of moderate (small and intermediate) and large phonon dispersion \([9, 28]\). The special case \( \nu = 1 \) of a constant weighting function, which corresponds to a critical value of \( \Delta = 3^{1/2} = 0.577 \), marks thus the characteristic point of transition between the two qualitatively different dispersion regimes.

Inserting (9) into (2) and observing (10) we can represent the model-specific \( E(T) \) shift generally in form of a product,

\[
E(T) = E(0) - \big( \alpha \Theta / 2 \big) \cdot \eta_\nu (2T/\Theta) ,
\]
(12)
of the magnitude of the $T = 0$ renormalization energy, $\alpha \Theta^2$ (8), and a characteristic (dimensionless) shape function $\eta_\nu(\tilde{\xi})$ (A1) (see appendix A) of a single (dimensionless) temperature argument, $\tilde{\xi} = 2T/\Theta$. A series of discrete $\eta_\nu(\tilde{\xi})$ values, which we have calculated for parameter values $\nu = 2/3, 1,$ and $2$ via numerical integration of (A1), are represented by dots in Fig. 2. From these series of $\eta_\nu(\tilde{\xi})$ points we see that the strength (curvature) of the $\tilde{\xi}$-dependence of shape functions $\eta_\nu(\tilde{\xi})$ in the cryogenic region, $\tilde{\xi} = 2T/\Theta \ll 1$, increases with decreasing $\nu$. This involves, according to the model-specific connection (11) of the dispersion coefficient $\Delta$ with the curve shape parameter $\nu$ (Fig. 1), an increased curvature of the $\eta_\nu(\tilde{\xi} \to 0) \propto \tilde{\xi}^{1+\nu}$ asymptote (A3) with increasing dispersion. It is this monotonic relationship between the actual strength of this curvature and the magnitude of phonon dispersion which provides the physical basis for a reasonable detection of the material-specific value of the dispersion coefficient $\Delta$ (Fig. 1) from the characteristic shape $\eta_\nu(\tilde{\xi})$ (A1) (“fingerprint”; Fig. 2) of a measured $E(T)$ dependence.

Fig. 1 Dependence (11) of the dispersion coefficient $\Delta$ on the parameter $\nu$, which controls the shape of the normalized weighting function (9) for the power function model. The material-specific values for $\nu$ and $\Delta$, which have been estimated for a variety of group IV, III–V, and II–VI materials (cf. Table 1) via Eq. (13) are indicated by dots. The value of $\Delta_c = 3^{1/2}$ marks the characteristic point of transition between the regimes of moderate (small to intermediate) and large dispersion.

Fig. 2 Characteristic shape functions $\eta_\nu(\tilde{\xi})$ for three discrete values of the curve shape parameter $\nu (=2/3, 1,$ and $2)$. The discrete series of dots represent the original (exact) $\eta_\nu(\tilde{\xi})$ values obtained by numerical integration of Eq. (A1). The bold curves show the results of the approximate higher order root expression (A5) with coefficients $a_1(\nu)$ (A9), $a_2(\nu)$ (A10), and $a_3(\nu)$ (A8). The limiting shape function of Bose–Einstein type, $\eta_B(\tilde{\xi}) = \coth (1/\tilde{\xi}) - 1$ (for vanishing dispersion, $\Delta = 0$), is the dashed curve, and the shape function of Varshni type, $\eta_V(\tilde{\xi}) = \tilde{\xi}^2/(1 + \tilde{\xi})$ (for excessively large dispersion, $\Delta > 1$; cf. Sec. 5), is the dash-dot curve.
Numerical fittings of measured $E(T)$ dependences on the basis of Eq. (12) are rather cumbersome and time-consuming as long as we use for the shape functions $\eta(\xi)$ the exact integral representation (A1) (cf. [26, 30]). For practical applications it is as a rule necessary and useful to approximate the original $\eta(\xi)$ functions (A1) by adequate analytical expressions. For this purpose we have studied in due detail (in appendix A) the characteristic $\xi$-dependences of the shape functions $\eta(\xi)$ (A1). We have found that, particularly for magnitudes of the parameter $\nu$ within the experimentally important range $0.6 \ll \nu \ll 2.1$ (cf. Fig. 1), a good approximation for the original $\eta(\xi)$ functions (A1) is given by the fractional higher-order root representation (A5) (with dispersion-related expansion coefficients $a_n(\nu)$ given by Eqs. (A8) to (A10)). The obvious adequacy of this analytical approximation is seen from the numerical comparison (in Fig. 2) between exact and approximate $\eta(\xi)$ curves resulting from Eqs. (A1) and (A5), respectively (for $\nu = 2/3, 1,$ and 2). Henceforth, according to Eq. (12) in combination with the approximation (A5), we can represent the theoretical $E(T)$ dependence (2) implied by the power function model, $w(\xi) \propto \xi^{\nu-1}$ (9), to good approximation by a fractional higher order root expression of the form

$$E(T) = E(0) - \frac{\alpha \Theta}{2} \left[ 1 + \sum_{n=1}^{3} a_n(\nu) \cdot \left( \frac{2T}{\Theta} \right)^{n+\nu} + \left( \frac{2T}{\Theta} \right)^{3+\nu} \right] - 1, \quad (13a)$$

where the expansion coefficients $a_n(\nu)$ are given (cf. the appendix) by

$$a_1(\nu) = \frac{5 + \nu}{6} \left( \frac{\pi}{2} \right)^{2+\nu-1/2}, \quad a_2(\nu) = \frac{1-\nu}{2}, \quad a_3(\nu) = \frac{(5+\nu)(1+\nu)^2}{3\nu(2+\nu)}. \quad (13b)$$

### 4 Results and discussion

The advantage of the analytical $E(T)$ formula (13) (in comparison with Eq. (5) in Ref. [26]) is the avoidance of numerical integrations for evaluations of the characteristic shape functions $\eta(2T/\Theta)$ (A1) (in (12)). This feature is rather important for practical applications (especially with respect to experimental studies). To compare present, theoretically consistent results due to Eq. (13) with earlier approximate results [9], which had been obtained via the relatively simple analytical approximation (1) ("$p$-representation" [4–9]) for the power function model, we have fitted the same $E_\xi(T)$ and/or $E_p(T)$ data sets for group IV materials (15R-SiC [34], Si [35, 36], and Ge [37]), binary III-V compounds (AlN [38], AlAs [39], AlSb [40], GaN [41], GaP [42, 43], GaAs [44, 45], GaSb [46, 47], InN [48], InP [49, 50], InAs [51], and InSb [51, 52]) as well as binary II–VI compounds (ZnS [53], ZnSe [53], ZnTe [53], CdS [54–57], CdSe [58], and CdTe [59]). By corresponding least-mean-square four-parameter fitting procedures using Eq. (13) we have determined the empirical parameters $E(0), \alpha, \Theta$, and $\nu$. (Sample fits are shown for Si and GaAs in Figs. 3 and 4, respectively). The individual zero-temperature positions, $E(0)$ or $E_p(0)$, are not reported here because they turned out to be essentially the same (except for tiny numerical differences of $\pm 0.1$ meV) as those obtained using other dispersion-related models [9, 28, 29]. The empirical parameters $\alpha, \Theta$, and $\nu$ are listed together with the corresponding dispersion coefficients, $\Delta (11)$, and further physically important quantities in Table 1.

Let us compare, first of all, the present results for basic parameter values (due to Eq. (13)) with their approximate counterparts estimated in Ref. [9] via the simplified $E(T)$ formula (1). A comparison with Table I of Ref. [9] shows above all considerable deviations between the present (theoretically consistent) values for the average phonon temperature, $\Theta$ (4), and the preceding estimations [9] for $\Theta_e$ (in Eq. (1)). Instructive are in particular the results for the group of zinc chalcogenides, ZnS, ZnSe, and ZnTe (where we dispose of very detailed and comprehensive $E_p(T)$ data sets [53]). We see that the approximate $\Theta_e$ values [9] for these materials are by about 15% lower than the comparable (self-consistent) $\Theta$ values (4).
due to Eq. (13) (Table 1). In various other cases (like Si, Ge, GaAs, InSb, CdS, and CdSe) the underestimation of the average phonon temperature parameter $\Theta$ (4) by $\Theta$ values [9] is seen to be limited to an order of 10%, whereas it achieves even an order of 25% in the case of SiC (15R). Especially the latter case shows that the use of a duly elaborate \( E(T) \) expression of type (13), instead of the simplified \( E(T) \) formula (1), is actually necessary when we wish to obtain, in addition to fine numerical simulations of given \( E_g(T) \) or \( E_{gx}(T) \) data sets, also theoretically consistent parameter sets.

The availability of theoretically consistent values for the effective phonon temperatures \( \Theta \) (4) is crucial in particular for reliable estimations of \( T = 0 \) renormalization energies. Within the present notation, the magnitudes of these renormalization energies are given by the combination $\alpha \Theta 2$ (8) of the basic parameters $\alpha$ and $\Theta$. [Note that the quantities $E_{\text{lim}}(0) - E(0) = \alpha \Theta 2$ (8) are positive for all group IV, III–V, and II–VI materials in question; cf. Table 1. Within a somewhat different notation used in Cardona’s paper [12], the notion of \( T = 0 \) renormalization energy refers to the quantity $\Delta E(T = 0) = E(0) - E_{\text{lim}}(0) = -\alpha \Theta 2$, the sign of which is opposite to that of (8).] A comparison of the quantities $\Delta E(T = 0)$ listed for the fundamental energy gaps of Ge, Si, GaAs, and ZnSe in Table 1 of Ref. [12] with their counterparts $\alpha \Theta 2$ (8) given in Table 1 of the present paper shows the following: (i) The value $\alpha \Theta 2 = 52$ meV, for Ge, is nearly coincident with $-\Delta E(T = 0) = 53$ meV [12]. (ii) The value $\alpha \Theta 2 = 71$ meV, for Si, is by 11% higher than the earlier estimate of $-\Delta E(T = 0) = 64$ meV [12] (which was taken from our earlier fit [9] of the same data sets [35, 36] using the rougher approximation (1) [5, 9]). Somewhat larger is the
difference between the present $\alpha \Theta_2$ value for Si and the relatively small magnitude of 50 meV estimated via the isotopic substitution effect [12]. (iii) The value $\alpha \Theta_2 = 60$ meV, for GaAs, is considerably lower than Cardona’s estimates of $-\Delta E(T = 0) = 90$ meV or 110 meV [12] (which resulted alternatively from a linear extrapolation of the measured $E(T)$ curve and the isotopic substitution effect, respectively). 

(iv) The value $\alpha \Theta_2 = 55$ meV, for ZnSe, is by about 18% lower than Cardona’s estimate of about 65 meV [12].

We believe that in cases where the estimates quoted for renormalization energies in [12] differ by more than 10% from present values (Table 1), the latter are more adequate. This applies in particular to the case of GaAs, where the estimates quoted in [12] are significantly higher (by factors of order 1.5 to 2) than the present $\alpha \Theta_2$ value (of about 60 meV). Possible uncertainties in presently estimated $\alpha \Theta_2$ values (Table 1) should not exceed an order of 10% in cases like Ge, GaAs, GaSb, InP, InAs, InSb, ZnS, ZnSe, ZnTe, CdS, and CdTe, where the experimental data sets are extended significantly beyond the effective Debye temperatures, $\Theta_\alpha(T \rightarrow \infty)$ [1, 29]. At the same time we would like to point out that a reliable determination of the limiting slope, $\alpha = -dE(T)/dT\big|_{\epsilon = 0}$, in combination with the effective phonon temperature, $\Theta_\alpha(4)$, requires in general a relatively close approach of the measured $E(T)$ curve to the linear asymptote (7) (cf. [53]). This requirement means that in practice measurements must be extended

Table 1  Dispersion-related parameters $\alpha$, $\Theta$, and $\nu$ obtained via least-mean-square fittings of experimental $E_g(T)$ (fundamental band gap) and/or $E_x(T)$ (exciton peak) data available for various group-IV, III–V, and II–VI materials using the $E(T)$ formula (13) for the power function model. Further we have listed the corresponding magnitudes for dispersion coefficients, $\Delta = [\nu(\nu+2)]^{1/2}(11)$, the magnitudes of $T = 0$ renormalization energies, $E_g(0) - E(0) = \alpha \Theta_2$ (8), the average phonon energies, $\nu_\Theta = k \Theta$, and the apparent cut-off phonon energies, $\nu_c = k \Theta = (\nu + 1) \nu_\Theta/\nu$ (cf. Eq. (10)). The values included in brackets are still largely uncertain due to the incompleteness (limitation to $T < 300$ K or scarcity in the cryogenic region) of the data sets in question.

<table>
<thead>
<tr>
<th>Material</th>
<th>Ref.</th>
<th>$T_{\max}$ (K)</th>
<th>$T_{\max}$ (eV/K)</th>
<th>$\alpha \times 10^{-4}$</th>
<th>$\Theta$ (K)</th>
<th>$\nu$</th>
<th>$\Delta$</th>
<th>$\alpha \Theta_2$ (meV)</th>
<th>$\nu_\Theta$ (meV)</th>
<th>$\nu_c$ (meV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiC (15R)</td>
<td>[34]</td>
<td>6 to 645</td>
<td>4.64</td>
<td>913</td>
<td>2.04</td>
<td>0.35</td>
<td>212</td>
<td>78.6</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Si</td>
<td>[35, 36]</td>
<td>2 to 415</td>
<td>3.22</td>
<td>441</td>
<td>1.20</td>
<td>0.51</td>
<td>71</td>
<td>38.0</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Ge</td>
<td>[37]</td>
<td>4 to 416</td>
<td>4.13</td>
<td>253</td>
<td>1.62</td>
<td>0.49</td>
<td>52</td>
<td>21.8</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>AlN</td>
<td>[38]</td>
<td>4 to 298</td>
<td>(10.9)</td>
<td>(887)</td>
<td>(1.75)</td>
<td>(0.39)</td>
<td>(483)</td>
<td>(76.5)</td>
<td>(120)</td>
<td></td>
</tr>
<tr>
<td>AlAs</td>
<td>[39]</td>
<td>4 to 287</td>
<td>3.89</td>
<td>255</td>
<td>1.27</td>
<td>0.49</td>
<td>50</td>
<td>22.0</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>AlSb</td>
<td>[40]</td>
<td>4 to 298</td>
<td>3.49</td>
<td>212</td>
<td>0.59</td>
<td>0.81</td>
<td>37</td>
<td>18.3</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>GaN</td>
<td>[41]</td>
<td>2 to 1067</td>
<td>6.14</td>
<td>585</td>
<td>1.56</td>
<td>0.42</td>
<td>179</td>
<td>50.4</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>GaP</td>
<td>[42, 43]</td>
<td>4 to 680</td>
<td>4.78</td>
<td>356</td>
<td>0.94</td>
<td>0.60</td>
<td>85</td>
<td>30.6</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>GaAs</td>
<td>[44, 45]</td>
<td>2 to 673</td>
<td>4.76</td>
<td>252</td>
<td>1.46</td>
<td>0.44</td>
<td>60</td>
<td>21.7</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>GaSb</td>
<td>[46, 47]</td>
<td>9 to 300</td>
<td>3.86</td>
<td>204</td>
<td>1.46</td>
<td>0.44</td>
<td>39</td>
<td>17.6</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>InN</td>
<td>[48]</td>
<td>4 to 300</td>
<td>(2.27)</td>
<td>(583)</td>
<td>(1.87)</td>
<td>(0.37)</td>
<td>(66)</td>
<td>(50.3)</td>
<td>(77)</td>
<td></td>
</tr>
<tr>
<td>InP</td>
<td>[49, 50]</td>
<td>4 to 873</td>
<td>3.96</td>
<td>273</td>
<td>1.33</td>
<td>0.47</td>
<td>54</td>
<td>23.5</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>InAs</td>
<td>[51]</td>
<td>10 to 300</td>
<td>2.82</td>
<td>146</td>
<td>(0.83)</td>
<td>(0.65)</td>
<td>21</td>
<td>12.6</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>InSb</td>
<td>[51, 52]</td>
<td>10 to 550</td>
<td>2.54</td>
<td>155</td>
<td>(1.66)</td>
<td>(0.41)</td>
<td>20</td>
<td>13.4</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>ZnS</td>
<td>[53]</td>
<td>2 to 541</td>
<td>5.49</td>
<td>285</td>
<td>1.63</td>
<td>0.41</td>
<td>78</td>
<td>24.6</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>ZnSe</td>
<td>[53]</td>
<td>4 to 500</td>
<td>5.00</td>
<td>218</td>
<td>1.76</td>
<td>0.39</td>
<td>55</td>
<td>18.8</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>ZnTe</td>
<td>[53]</td>
<td>2 to 291</td>
<td>4.68</td>
<td>171</td>
<td>1.67</td>
<td>0.40</td>
<td>40</td>
<td>14.7</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>CdS</td>
<td>[54–57]</td>
<td>2 to 289</td>
<td>4.09</td>
<td>165</td>
<td>1.34</td>
<td>0.47</td>
<td>34</td>
<td>14.2</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>CdSe</td>
<td>[58]</td>
<td>15 to 550</td>
<td>4.10</td>
<td>188</td>
<td>1.91</td>
<td>(0.37)</td>
<td>39</td>
<td>16.2</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>CdTe</td>
<td>[59]</td>
<td>2 to 300</td>
<td>3.09</td>
<td>104</td>
<td>0.75</td>
<td>0.70</td>
<td>16</td>
<td>9.0</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>
up to the vicinity of the effective Debye temperature, \( \Theta_D(T \to \infty) \) [1]. (Note that \( \Theta_D(T \to \infty) \) is comparable to the estimated cut-off phonon temperature, \( \Theta_c = (\nu + 1) \Theta_D; \) cf. Eq. (10) and see below). For the materials listed in Table 1, this requirement is obviously not fulfilled for SiC (15R), Si, AlN, AlAs, AlSb, GaP, and InN. Largely uncertain are in particular the parameter values for AlN and InN (we have enclosed in brackets). This is due to the circumstance that the effective Debye temperatures for AlN and InN [29] are higher by factors of about 3 or 2, respectively, than the experimental cut-off temperatures (of about 300 K [38, 48]). More comprehensive measurements for these III-nitrides thus might yield \( \alpha \) and \( \Theta \) values that differ significantly from those quoted in Table 1.

Of considerable physical interest is an information on locations of relevant regions of the phonon energy spectra that are making significant contributions [28, 29] to the measured temperature dependences of the energy gaps in question. Of primary importance in this respect is the average phonon energy, \( \varepsilon_{av} = k \Theta_D \) which represents the position of the center of gravity (cf. Eq. (4)). Referring to known phonon energy spectra [1, 29] we see that the estimated \( \varepsilon_{av} \) values (listed in Table 1) are located either below or within the corresponding phonon energy gaps (separating upper from lower phonon energy bands). These findings are indicative for relatively strong contributions of low-energy bands (LA and TA phonon branches) to the observable gap shrinkage effect (in accordance with Ref. [29]). In any case, the estimated \( \varepsilon_{av} \) positions satisfy throughout the fundamental physical requirement of being located significantly below the cut-off energies, \( \varepsilon_c \) [1, 29], of the upper (optical) bands.

The present analyses by the power function model provided still another physically informative by-product. This consists in estimates of apparent cut-off energies, \( \varepsilon_c = k \Theta_D = (\nu + 1) \varepsilon_{av}/\nu \) (listed in the last column of Table 1). The latter give an at least rough idea of the total size (upper boundary) of the phonon energy spectrum in a given material or heterostructure. Their comparison with known cut-off energies \( \varepsilon_c \) [1, 29] for the upper (optical) phonon energy bands shows that the deviations between both quantities are as a rule limited to an order of 20% (except for case of CdS). For many materials pertaining to the regime of moderate dispersion [28], \( \Delta < 3 \) [12] (cf. Fig. 1), the actual deviations of estimated \( \varepsilon_{c} \) values (Table 1) from true \( \varepsilon_{av} \) positions [1, 29] are even limited to an order of only 10%. This applies in particular to SiC (15R), Si, Ge, GaAs, GaSb, InP, InAs, ZnS, ZnSe, ZnTe, and CdSe. The rather close approaches between both quantities in so many cases might be somewhat surprising, at first sight, in view of the apparent simplicity of the power function ansatz (9) for the weighting function (which a priori ignores, e.g., a basic feature like the existence of energy gaps between lower and higher phonon energy bands in these materials). Yet, the obvious success of the present order-of-magnitude estimations of cut-off energy positions via the theoretical (model-specific) quantity \( \varepsilon_c = k \Theta_D = (\nu + 1) \varepsilon_{av}/\nu \) shows that even the latter can be considered as a physically reasonable piece of information on the phonon spectra in question. Such information can be used in practice in two different ways:

(A) For elemental and binary semiconductor materials, approximate values for cut-off energies, \( \varepsilon_{av} \), are usually known [1] either from theoretical calculations or from other kinds of experimental investigations (in particular from Raman measurements). In such cases, a relatively close approach or a large difference between estimated \( \varepsilon_c \) and known \( \varepsilon_{av} \) values can be taken as an indication of internal consistency or a possible inconsistency, respectively, of the theoretical interpretation of the given \( E(T) \) data set via the power function model.

(B) For many cases of ternary and quaternary materials as well as a large variety of heterostructures (superlattices, MQW structures, quantum dots, etc.), we have little information on the actual phonon energy spectrum. In such a case (see below), the estimated \( \varepsilon_c \) value can be taken as a rough (preliminary) information on the upper boundary of the phonon energy spectrum.

A crucial feature of the novel analytical \( E(T) \) expression (13) basing on the power function model (Sec. 3) is its usability for a more detailed study of \( E(T) \) dependences measured in the cryogenic region. The corresponding low-temperature asymptotes (for \( T \ll \Theta \)) are commonly assumed to be given by certain power dependences, \([E(0) – E(T)] \propto T^p\), with exponents \( p \approx 2 \). In consideration of corresponding
analytical results presented already many years ago in several theoretical papers [31–33], the $T \to 0$ asymptote was believed for a long to be a strictly quadratic one ($p = 2$). At variance to this, there has been given in a recent paper by Cardona [16] a global theoretical argument in favor of a cubic asymptote ($p = 3$). The perspective of an eventually direct determination of the low-temperature exponent $p$ via least-mean-square fittings of measured $E(T)$ dependences can thus be considered as an important goal of forthcoming investigations. A reliable determination of the actual $T \to 0$ asymptote via numerical fittings of experimental $E(T)$ data sets, however, is generally rather difficult. This is due to the fact that non-negligible differences between alternative fittings using structurally different dispersion-related models are continually found to be throughout limited to the cryogenic region and to be usually smaller than \( \pm 0.2 \) meV. A reliable discrimination between different models assumes thus comprehensive sets of high-precision data (with error bars \( \approx 0.2 \) meV) to be available. The majority of experimental data sets available hitherto does not satisfy this requirement.

The existing numerical problem can be well illustrated by a comparison between present fittings using Eq. (13) and the results obtained by fittings using an alternative high-precision model developed recently in Ref. [11]. This model is based on an alternative ansatz for the weighting function $w(\epsilon)$ in (2) and (3) of the type $w_\nu(\epsilon) \propto (1 - \epsilon^2)^{\nu-1}$ (with curve shape parameter $\sigma > 0$ and an associated dispersion coefficient of $\Delta = (\sigma(\sigma + 2))^{1/2}$). A detailed analytical and numerical study of this model in [11] yielded a corresponding high-precision formula for the relevant interval $1/3 \leq \Delta \leq 3/4$ (Fig. 1) of the form

\[
E(T) = E(0) - \frac{a\Theta}{2} \left[ 1 + \sum_{n=2}^{\nu} a_n(\Delta) \left( \frac{2T}{\Theta} \right)^n + \left( \frac{2T}{\Theta} \right)^{\nu} - 1 \right],
\]

where the expansion coefficients $a_n(\Delta)$ are given by

\[
a_n(\Delta) = \frac{\pi^n \Delta^n}{1 + \Delta^n}, \quad a_2(\nu) = 1 - 4\Delta^2 + 3\Delta^4, \quad \text{and} \quad a_3(\Delta) = 2\left(1 + \Delta^2\right).
\]

We see readily that the equations (13) and (14) are coincident for the special case of a constant weighting function (which is represented by $\nu = 1$ in Eq. (13) and $\Delta = 1/3^{1/2} = 0.577$ in Eq. (14)). Consequently, both models yield for this case also the same quadratic $T \to 0$ asymptote, $[E(0) - E(T)] \to (\pi^2/12) aT^2/\Theta$. In contrast to this, for all the other values of $\nu \neq 1$ in Eq. (13), the $T \to 0$ asymptote differs more or less strongly from a quadratic one, $p = 1 + \nu \neq 2$, whereas its counterpart due to Eq. (14) remains a strictly quadratic one, $p = 2$ (for any $\Delta$). This is a principal qualitative difference between both four-parameter formulas which might be detectable, in principle, via numerical analyses of high-precision data sets (at least).

However, the experimental data sets analyzed above (Table 1) turned out to be not yet informative enough for a clear discrimination between these alternatives. In order to visualize the typical state of affairs for the data sets in question we have shown in the insets to Figs. 3 and 4 the low-$T$ sections of the theoretical $E(T)$ curves for Si and GaAs, respectively, that are resulting from alternative fittings using Eqs. (13) and (14). The numerical comparisons show that both theoretical model curves are almost coincident. Properly speaking, the differences between the alternative $E(0)$ positions resulting from the alternative models (13) and (14) are smaller than 0.1 meV (both for Si and GaAs). This means that the model-related differences are of the same order as the experimental uncertainties (error bars) in the cryogenic region. For many other data sets, the experimental uncertainties are even considerably larger. From this comparison it is obvious that, on the basis of the experimental $E(T)$ data sets available up till now for bulk samples (Table 1), it is not possible to distinguish clearly between both models. (Note that the residual variances due to alternative fittings by Eq. (14) vs. Eq. (13) differ usually by less than 5% and the relative deviations between the corresponding $a$ and $\Theta$ values tend to be even smaller).

Better chances for a discrimination between both models can be expected to be given by novel $E(T)$ data sets available from more recent studies of various quantum-well (QW) and quantum-dot (QD) structures. Most promising in this respect appears the single dot photoluminescence spectroscopy [13, 14].
The linewidths associated with ground state exciton recombinations in the cryogenic region reduce in favorable cases to values considerably smaller than 1 meV [13, 14]. This gives us in principle the chance of a much more accurate determination of the temperature shift of QD luminescence peak positions than in macroscopic systems. A corresponding high-precision set of $E_{\text{lum}}(T)$ positions due to an In$_{0.6}$Ga$_{0.4}$As/GaAs self-assembled QD system [13] has been obtained recently by Ortner et al. [14] (see Fig. 5). The parameter values resulting from least-mean-square fittings by Eqs. (13) and (14) are given in Table 2. From the latter we see, among other things, that the $\alpha$ and $\Theta$ values estimated alternatively by Eq. (14) vs. Eq. (13) differ by about 10% or 13%, respectively, which involve a difference of about 25% for the associated renormalization energy, $\alpha\Theta/2$ (8). These differences in magnitudes of the parameters determining the high-$T$ asymptote (7) are a typical consequence of the limitation of measurements to a break-off temperature of $T_{\text{max}} = 100$ K, which is significantly lower than the average phonon temperature, $\Theta \cong (200 \pm 15)$ K. (Note that the effective Debye temperature for the QD in question can be estimated to be of order $\Theta_{\text{D}}(T \to \infty) \approx 300$ K, and the cut-off phonon temperature should be even higher, $\Theta_{\text{c}} \cong \epsilon_{\text{c}}/k \approx 350$ K). As we have pointed out above, it is impossible in such cases to determine accurately the parameters $\alpha$ and $\Theta$ (which analytical model would ever be used). Yet, the most interesting aspect of these alternative fittings (see the inset to Fig. 5) consists in a residual variance of the fit by Eq. (13) that is by 20% lower than that of the fit by Eq. (14). This is a first indication for the present power function model to be more adequate than the alternative model model of Ref. [11] (which assumed a strictly quadratic $T \to 0$ asymptote).

Assessing the corresponding parameter values for this QD system (see Table 2) it might still be interesting to observe the following: Starting from the $\alpha$ and $\Theta$ values obtained above for bulk samples of InAs and GaAs (Table 1) and calculating their averages at weights of 0.6 and 0.4, respectively (according to the composition of the QD in question), we come to magnitudes of $\alpha_{\text{av}} \cong 3.6 \times 10^{-4}$ eV/K and $\Theta_{\text{av}} \cong 188$ K. The latter are, interestingly, of the same order of magnitude as the fitted $\alpha$ and $\Theta$ values listed in Table 2. (Note in particular the approximate equality between $\alpha_{\text{av}}$ and the $\alpha$ value due to Eq. (13) as well as that between $\Theta_{\text{av}}$ and the $\Theta$ value due to Eq. (14)). In view of these approximate equalities
Table 2 Parameter sets obtained via least-mean-square fittings of the novel high-precision $E_{\text{in}}(T)$ data set available from recent single dot photoluminescence spectroscopy measurements by Ortner et al. [14] (for an In$_{0.6}$Ga$_{0.4}$As/GaAs self-assembled QD system [13, 14]) by the alternative four-parameter model expressions (13) and (14).

<table>
<thead>
<tr>
<th>Eq.</th>
<th>$E(0)$ (eV)</th>
<th>$\alpha/10^{-4}$ (eV/K)</th>
<th>$\Theta$ (K)</th>
<th>$\nu$</th>
<th>$\Delta$</th>
<th>$\alpha\Theta/2$ (meV)</th>
<th>$\varepsilon_c$ (meV)</th>
<th>$\varepsilon_{av}$ (meV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13)</td>
<td>1.26744</td>
<td>3.58</td>
<td>209</td>
<td>1.52</td>
<td>0.43</td>
<td>37</td>
<td>18.0</td>
<td>30</td>
</tr>
<tr>
<td>(14)</td>
<td>1.26748</td>
<td>3.23</td>
<td>185</td>
<td>–</td>
<td>0.36</td>
<td>30</td>
<td>16.0</td>
<td>–</td>
</tr>
</tbody>
</table>

of comparable parameters one can suggest that the gap shrinkage mechanism in the In$_{0.6}$Ga$_{0.4}$As/GaAs QD system in question [13, 14] should be rather similar to its counterpart in an In$_{0.6}$Ga$_{0.4}$As bulk sample.

Further we would still like to mention that the dispersion-related four-parameter expressions of type (13) or (14) are, of course, applicable not only to numerical analyses of temperature dependences of fundamental energy gaps, but in principle also to those of various critical point (CP) energies, e.g. the $E_1(T)$ data sets available from literature for the $E_1$ critical points in Si [60], Ge [17], and various III–V materials (like GaP [22], GaAs [19], InP [20], InSb [18], etc.). Unfortunately, due to the relative scarcity of the available experimental data points particular in the cryogenic region ($T < 100$ K), these $E_1(T)$ data sets are as a rule insufficient for a proper determination of the dispersion coefficient $\Delta$. (Note that this common deficiency is especially striking for Ge [17] and InSb [18], where experimental $E_1(T)$ data are not at all available for this region.) In such cases it appears sufficient when we limit ourselves, preliminarily, to somewhat rougher numerical analyses using lower-order approximations like the structurally relatively simple “$p$-representation” [4–9] (i.e. Eq. (1), with fixed $p \rightarrow 5/2$, if necessary) or the Bose–Einstein model [15–24] (see also [5–12]).

5 What is actually wrong with Varshni’s model

Considering the question of applicability of one or the other analytical $E(T)$ model function from physical points of view, there are two basic requirements which should throughout be fulfilled by any model in consideration, namely:

– firstly, the model should be capable of simulating in good approximation the measured $E(T)$ dependence and,

– secondly, the model-specific parameters should be related in reasonable way to basic physical properties of the system like the magnitude of the slope, $\alpha \approx \left| \frac{dE(T \rightarrow \Theta_c)}{dT} \right|$, that is observable in the high temperature region ($T \approx \Theta_c$), the average phonon temperature, $\Theta \approx (0.4 \div 0.7) \Theta_c$ (cf. Table 1 in [29]), and the magnitude of the associated renormalization energy, $E_{\text{in}}(0) - E(0) = \alpha\Theta/2$ (8).

Especially the latter requirement, however, is as a rule not fulfilled (even not in rough approximation) by Varshni’s ad hoc model [2, 3],

$$E_v(T) = E(0) - \frac{\alpha_v T^2}{\beta + T}. \tag{15}$$

For a comparison of the latter with basic equations of the dispersion-related semi-empirical theory (Sec. 2), let us first consider the ratio $x = \beta T$ to be small compared to unity (in analogy to the ratio $\Theta T$ in Eqs. (6) and (7)) and use the Taylor expansion $(1 + x)^{-1} = 1 - x + x^2 - ...$ for the denominator in Varshni’s formula. In this way we obtain, for the region of intermediate to high temperatures, an expansion of the form

$$E_v(T) \rightarrow E(0) - \alpha_v \left( T - \beta + \beta^2/T - ... \right) \tag{16}$$
(which, up to the $1/T$–term, has the same structure as Eq. (6)). In the high temperature limit, $T \gg \beta$, this expression tends to the linear asymptote \[4\], $E_r(T) \to E(0) - \alpha_r(T - \beta) = E_{\text{ren}}(0) - \alpha_r T$ (in analogy to (7)), where $E_{\text{ren}}(0) = E(0) + \alpha \beta$ is the point of intersection of this asymptote with the energy axis. Accordingly, within the frame of Varshni’s model, we obtain for the $T = 0$ renormalization energy the model-specific parameter constellation $E_{\text{ren}}(0) - E(0) = \alpha \beta$. In order to exemplify the notorious inconsistencies of Varshni’s parameter sets, let us make sample comparisons between results of conventional three-parameter fittings using Varshni’s formula (15) and the results of the above four-parameter fittings via the self-consistent $E(T)$ expression (13).

The low temperature sections of $E_r(T)$ curves (15) and their counterparts due to Eq. (13) are shown in the insets of Fig. 3 (for Si), of Fig. 4 (for GaAs), and of Fig. 5 (for the In$_{0.4}$Ga$_{0.6}$As/GaAs QD system). From these insets we see, first of all, that the $E_r(T)$ curves (15) for all three cases show significantly stronger curvatures in the cryogenic region than observed experimentally. The analytical and numerical aspects of such systematic deviations, which are rather pronounced for practically all cases of moderate (small to intermediate) dispersion, $\Delta < 3^{1/2}$ (cf. Fig. 1), have already been discussed in detail in Refs. [5–9] (see also below). Especially unpleasant, however, are the excessive numerical instabilities [8–10] of Varshni’s parameter sets ($\alpha$ and $\beta$ values), which are closely connected with different choices of experimental cut-off temperatures in different studies. For the experimental data constellations presently considered for Si (Fig. 3) and GaAs (Fig. 4), e.g., the parameters in Varshni’s formula (15) took the values $\alpha = 5.36 \times 10^{-4}$ eV/K (with $\beta = 745$ K) and $\alpha_r = 5.73 \times 10^{-4}$ eV/K (with $\beta = 218$ K), respectively. This means an overestimation of limiting slopes by about 66% and 11%, respectively, in comparison with the self-consistent $\alpha$ values (listed in Table 1). The large degree of arbitrariness of Varshni’s parameters has often dramatic numerical consequences especially for the products of both parameters, $\alpha \beta$. We have seen above that these products are representing, within Varshni’s model (15), the counterparts to the dispersion-related, theoretically consistent parameter constellation $\alpha \Theta^2$ (8) for the $T = 0$ renormalization energies. From the $\alpha$ and $\beta$ values quoted above one obtains for the products $\alpha \beta$ magnitudes of about 400 meV for Si and 115 meV for GaAs. This means an overestimation by factors of about 5.6 and 1.9, respectively, in comparison with the $\alpha \Theta^2$ values for Si and GaAs (cf. Table I). [Note that we would be concerned, in the case of Si, in the actual magnitude, $\alpha \beta = 778$ meV is even by a factor 11 higher than the theoretical counterpart of $\alpha \Theta^2 = 71$ meV, of the $T = 0$ renormalization energy of Si.]

The misinterpretations by Varshni’s formula (15) use to be still more dramatic when it is used for fittings of $E(T)$ data sets that are limited by relatively low break-off temperatures ($T_{\text{max}} \ll \Theta = \varepsilon/k$). The $E_{\text{ren}}(T \leq 100$ K) data set (Fig. 5) for the In$_{0.4}$Ga$_{0.6}$As/GaAs QD system in consideration provides a good example. Firstly, we have found that an alternative least-means-square-fitting procedure using Eq. (15) ends with a residual variance that is by almost an order of magnitude (a factor 7.2) higher than the actual fit, in the proper sense of this word, by Eq. (13). This qualitative difference is due to the notorious incapability of Eq. (14) to adjust to the relative weak temperature dependence measured at very low temperatures ($T < 20$ K). Secondly, we have obtained after a sufficiently long least-mean-square fitting process for the parameters of Varshni’s type the (physically obviously quite unreasonable) values $\alpha_r = 2.84$ eV/K and $\beta = 2.31 \times 10^6$ K. [Note that the $\alpha_r$ value is by a factor of about $8 \times 10^5$ higher than the limiting slope $\alpha$ estimated via Eq. (13); cf. Table 2.] This parameter constellation corresponds to an estimate of $\alpha \beta = 6.56 \times 10^6$ eV for the renormalization energy, which is by a factor of about $1.8 \times 10^6$ higher than the theoretically consistent counterpart of $\alpha \Theta^2 = 37$ meV (due to Eq. (13)). This striking order-of-magnitude difference shows once more that a use of Varshni’s formula can lead in certain cases (when $T_{\text{max}}$ is relatively low) to physically quite absurd parameter constellations (cf. Sec. 5 of Ref. [9]).

From physical points of view, the continually found inadequacy of Varshni’s model function can be understood as being due to the fact that this model refers, implicitly, to a physically unrealistic regime of
extremely large dispersion, which is never found in experiment. This can already roughly be seen from the comparison (in Fig. 2) of the characteristic shape function of Varshni’s type, \( \eta_\nu(\xi) = \xi^{\frac{2}{7}}/(1 + \xi) \), with their dispersion-related counterparts \( \eta_\nu(\xi) \) for \( \nu = 2, 1, \) and \( 2/3 \). These sample values for the curve shape parameter \( \nu \) correspond to dispersion coefficients, \( \Delta \) (11), of about 0.35, 0.58, and 0.75. By comparison with Table 1 we see that the lowest and highest magnitudes of these sample values for \( \Delta \) are marking approximately the physically relevant interval for the dispersion coefficient, \( 1/3 \approx \Delta \approx 3/4 \). At the same time we see from Fig. 2 that Varshni’s curve, \( \eta_\nu(\xi) = \xi^{\frac{2}{7}}/(1 + \xi) \), lies clearly outside the physically realistic corridor of the \( \eta - \xi \) plane that is spanned by the sample functions \( \eta_{\nu=0}(\xi) \) and \( \eta_{\nu=2/3}(\xi) \). This suggests that the dispersion coefficient, \( \Delta_\nu \), which is factually implied by Varshni’s model, must be significantly higher than \( 3/4 \) (at least). This is in qualitative accordance with an earlier order of magnitude estimation [10] of \( \Delta_\nu \approx 1 \).

We have performed recently a more detailed study [11] of the effective degree of dispersion associated with Varshni’s model. This has shown, above all, that the dispersion coefficient associated with Varshni’s model is even significantly higher than unity. Yet, the basic inequality \( \Delta_\nu > 1 \) can be obtained also in relatively simple way from a comparison of the universal series expansion (6), for intermediate to high temperatures, with the counterpart (16) due to Varshni’s \( E(T) \) formula (15). We see immediately that the linear and constant terms in the expansions (6) and (16) are coincident just when

\[
\alpha_\nu = \alpha \quad \text{in combination with} \quad \beta = \Theta/2. 
\]

(17)

By means of these equalities (17) we can represent the subsequent (nonlinear) term in Varshni’s expansion (60) in the form \( \beta^2/4T = \Theta^2/12T \). Comparing the latter with the corresponding \( (1 + \Delta^2) \Theta^2/12T \) term in Eq. (6) we see that both nonlinear \( (\propto T^{-1}) \) terms are coincident just for \( (1 + \Delta^2) \to 3 \). This means that the dispersion coefficient associated with Varshni’s model (15) should be on the order of \( \Delta_\nu \approx 2^{1/2} \approx 1.4 \). Of course, the latter value represents only a rough estimate. The comprehensive analytical and numerical study [11] yielded \( \Delta_\nu = (\pi^2/6 - 1)^{1/2} \approx 1.25 (\pm 0.01) \). Yet, apart from such finer numerical details, the physically most important conclusion which can be drawn unanimously both from the order-of-magnitude estimation sketched above and the detailed study in [11], can be summarized by the global statement that the magnitude of dispersion coefficient, that is effectively involved by Varshni’s hypothetical model function (15), is significantly higher than unity, \( \Delta_\nu > 1 \). In contrast to this, as seen from Table 1, the empirical values for dispersion coefficients are throughout significantly lower than unity. This qualitative difference (numerical gap) between a hypothetical regime with \( \Delta_\nu > 1 \) and the realm of physically realistic degrees of phonon dispersion, \( 1/3 \approx \Delta_{\exp} \approx 3/4 \) (Fig. 1), is the objective cause of the general inadequacy of Varshni’s formula (15).

6 Summary and conclusions

The main achievement of this study is the development of a successful higher-order root expression for monotonic temperature dependences of band gaps of semiconductors and insulators, based on the power function model for the weighting function. The derivation of the final analytical formula has been given here in detail. In contrast to the earlier, simpler formula of the power function model ("p-representation"), the new analytical expression provides – in addition to excellent numerical fits – also theoretically consistent values for dispersion-related parameters. As main targets of application of the new \( E(T) \) expression we consider above all high-precision data sets, in particular those obtainable from current single QD luminescence peak position measurements (in analogy to [13, 14]). Detailed numerical analyses of such data sets by the present model (Eq. (13)) give us, among other things, more information on \( E(T \to 0) \) asymptotes. Comparisons with results of fittings using the alternative model developed recently in [11] (Eq. (14)) give us a chance to distinguish between strictly quadratic and various higher-order power dependences for \( E(T \to 0) \) asymptotes.
At the same time we would like to point out that the exponent \( \nu + 1 \) of the theoretical asymptote, \([E(0) - E(T)]_{\text{low}} \propto T^{\nu+1}\), which follows from a fit using Eq. (13), can not yet be considered as a quite accurate representative of the exponent \( p \) of the asymptote, \([E(0) - E(T)]_{\text{exp}} \propto T^p\), that is inherent to the experimental data points. It is a basic feature of the four-parameter model in question that the parameter \( \nu \) in Eq. (13) (similarly as \( \Delta \) in Eq. (14)) plays a double role, namely: (1) it controls the asymptotic behavior at low temperatures, \( T \ll \Theta \), and (2) it co-determines, owing to its a priori connection with the dispersion coefficient, \( \Delta = (\nu(\nu + 1))^{-1/2} \) (11), also the temperature dependence (6) at intermediate temperatures, \( T \approx \Theta \). A separate determination of the low-\( T \) exponent, \( p \), in conjunction with other adjustable parameters (such as \( E(0), \alpha, \Theta, \) and \( \Delta \)) requires thus the use of a more sophisticated analytical model, comprising a total set of at least five adjustable parameters. In order to get first sample results for the real low-\( T \) exponents in question we have still performed corresponding five-parameter fittings of the high-precision data sets available for Si [35], GaAs [45], ZnSe [53], and ZnTe [53]. As an interesting and agreeable result we have found that the maximum differences between the real low-precision data sets available for Si [35], GaAs [45], ZnSe [53], and ZnTe [53]. As an interesting and agreeable result we have found that the maximum differences between the real low-precision data sets available for Si [35], GaAs [45], ZnSe [53], and ZnTe [53].

**Appendix: Characteristic shape function**

Inserting the normalized weighting function (9) into (2) and observing (10) we see that the characteristic (dimensionless) shape function \( \eta(\xi) \) of the single (dimensionless) temperature argument \( \xi = 2T/\Theta \) in Eq. (12) is given by integrals of the (equivalent) form

\[
\eta(\xi) = \frac{(\nu\xi)^{1+\nu}}{(2(1+\nu))^{1+\nu}} \int_0^{z(\xi)} \frac{z^\nu dz}{\exp(z) - 1} = 2(1+\nu) \int_0^1 x^\nu \left[ \exp\left(\frac{2(1+\nu)}{\nu\xi} x\right) - 1 \right]^{-1} dx
\]

(A1)

(where the upper limit of integration in the first representation is given by \( z(\xi) = 2(1+\nu)/\nu\xi \). Consider now in detail the asymptotic low- and high-temperature dependences of these dispersion-related shape functions (A1). For the cryogenic region, \( T \ll \Theta \), the upper limit of integration \( z(\xi) \) (in the first representation) is much higher than unity, \( z(\xi) = 2(1+\nu)/\nu\xi = (1+1/\Theta\nu)T \gg 1 \). Consequently, in view of the fast exponential decrease of the corresponding integrand (for finite \( \nu \)), we can formally extend the integration region to infinity, \( z(\xi) \to \infty \). Numerical values for integrals of this type are well known to be given by products of the familiar \( \Gamma \)-function and Riemann’s \( \zeta \)-function (with arguments \( 1+\nu \)), i. e.

\[
\int_0^\infty \frac{dz}{\exp(z) - 1} = \Gamma(1+\nu) \zeta(1+\nu).
\]

(A2)

Using the relation \( \Gamma(2+\nu) = (1+\nu) \Gamma(1+\nu) \) for \( \Gamma \)-functions with arguments differing by unity, we obtain thus for the low temperature asymptote of the shape function \( \eta(\xi) \) (A1) an expression of the form

\[
\eta(\xi) \to 2\Gamma(2+\nu) \zeta(1+\nu) \left( \frac{\nu^2}{2(1+\nu)} \right)^{1+\nu}.
\]

(A3)

This means, among other things, that the \( E(T) \) curve (12) in the cryogenic region reduces to a certain power dependence of temperature, \( (E(0) - E(T)) \propto \xi^{2+\nu} \propto T^{2+\nu} \), whose exponent \( p = \nu + 1 \) [6, 9] is just higher by 2 than the exponent controlling the \( \nu \)-dependence of the weighting function considered, \( w(\nu) \propto \nu^{\nu+1} \) (9).
For the high-temperature region, $T \gg \Theta$, the argument of the exponential in the second representation of (A1) is small compared to unity, $2(1 + \nu) x / \nu \xi \equiv (1 + \nu) \Theta / \sqrt{vT} \ll 1$. Thus we can use a Taylor expansion of the form $z/(e^z - 1) = 1 - z/2 + z^2/12 - z^4/720 + ...$ (in analogy to Sec. 2) in order to transform the integrand in (A1) into a corresponding series expansion (in terms of powers of the argument $z = 2(1 + \nu) x / \nu \xi$, with fractional exponents). Performing hence the $x$-integration through the range $0 \leq x \leq 1$ we obtain an asymptotic expression of the form

$$\eta_\nu(\xi) \rightarrow \xi^{-1+\frac{(1+\nu)^2}{3v(2+\nu)\xi}}. \quad (A4)$$

Let us now attack the hitherto unsolved problem of constructing a generally good analytical approximation for shape functions $\eta_\nu(\xi)$ (A1) for arbitrary magnitudes of $\xi$. As a necessary condition we must, obviously, require that the low- and high-temperature asymptotes be coincident with the corresponding asymptotes (A3) and (A4), respectively, of the original $\eta_\nu(\xi)$ function. This condition is satisfied by a higher-order root ansatz of the form

$$\eta_\nu(\xi) \approx \left[ 1 + a_1(\nu) \xi^{1+\nu} + a_2(\nu) \xi^{2+\nu} + a_3(\nu) \xi^{3+\nu} + \xi^{5+\nu} \right]^{1/\nu} - 1. \quad (A5)$$

Actually, when we consider the limiting ranges of low and high temperatures (i.e. $\xi \ll 1$ and $\xi \gg 1$), we see that the root expression (A5) tends to asymptotes of the desired type,

$$\eta_\nu(\xi) \rightarrow \frac{a_1(\nu)}{5+\nu} \xi^{1+\nu} \quad \text{and} \quad \eta_\nu(\xi) \rightarrow \xi^{-1+\frac{a_1(\nu)}{(5+\nu)\xi}}, \quad (A6)$$

respectively. Comparing the latter with their counterparts (A3) and (A4) from the original $\eta_\nu(\xi)$ function (A1), we see that the required coincidence of the respective asymptotes takes place when the lower and higher order expansion coefficients, $a_1(\nu)$ and $a_3(\nu)$, are given by

$$a_1(\nu) = \frac{(5+\nu) \Gamma(2+\nu) \xi(1+\nu)}{2^\nu} \left( \frac{\nu}{1+\nu} \right)^{1+\nu} \quad (A7)$$
and

\[ a_1(\nu) = \frac{(5 + \nu)(1 + \nu)^2}{3\nu(2 + \nu)}, \quad (A8) \]

respectively (cf. Fig. 6). At the same time it should be noted that no unambiguous analytical expression can be derived for the intermediate order coefficient \( a_2(\nu) \) (see below).

Unfortunately, especially for fractional values of \( \nu \), one does not dispose hitherto of usable analytical representations for \( I^- \) and \( \zeta \)-functions which co-determine the low-temperature coefficient \( a_1(\nu) \) (A7). Since knowledge of this coefficient is crucial for the description of the low temperature dependence at different magnitudes of dispersion, we focussed on finding a reasonable analytical approximation for the coefficient \( a_2(\nu) \) (A7). In particular with respect to the experimentally relevant \( \nu \)-range (of about \( 2/3 < \nu < 10/3 \); cf. Table 1 and Fig. 1) we have found a very good approximation to be given by

\[ a_2(\nu) \approx \frac{5 + \nu}{6} \left( \frac{\pi}{2} \right)^{2+\nu-1/2}. \quad (A9) \]

A numerical comparison (Fig. 6) between the original \( a_1(\nu) \) expression (Eq. (A7)) and (A9) shows that the difference between them is limited to less than 5% (within a \( \nu \)-region of \( 2/3 < \nu < 3 \), at least). This is sufficient for practical purposes. (Note that Eq. (A9) is exact for \( \nu = 1 \), where the magnitude of \( a_1(\nu) \rightarrow (\pi/2)^2 \) (A9) is coincident with (A7).) Approximate values for the intermediate order coefficient \( a_2(\nu) \) can be estimated by numerical comparisons (especially for intermediate temperatures, \( \zeta \approx 1 \); cf. Fig. 2) of original \( \eta(\xi) \) values (A1) obtained via numerical integration with approximate values following from the root expression (A5) (Fig. 2). In this way we found a good approximation for this coefficient to be

\[ a_2(\nu) \approx (1 - \nu)^{1/2}. \quad (A10) \]

References