

Kramers-Kronig analysis of the optical constants of Zn_3As_2 and Zn_3P_2 *

III. Improvement of Philipp-Taft method

KRZYSZTOF JEZIEWSKI, JAN MISIEWICZ, JANUSZ WNUK, JANUSZ M. PAWLIKOWSKI

Institute of Physics, Technical University of Wrocław, 50-370 Wrocław, Wybrzeże Wyspiańskiego 27, Poland.

The present paper consists of five parts and is devoted mainly to the results of optical constants examination in Zn_3As_2 and Zn_3P_2 which are compounds of II_3-V_2 type with broad energy gap. The analysis of the up to now state-of-affairs so far as the application of the Kramers-Kronig analysis method (KK) for the complex coefficient of reflection is concerned has indicated that it is necessary to develop the research in three additional directions described in the first parts of this series. In the part I a critical analysis of the calculation methods used till now and based on dispersion relations due to Kramers and Kronig has been presented. Also the methods of calculation of optical constants not employing the KK integral have been discussed.

In the parts II and III some suggestions of effective improvements of calculation methods basing on the KK integrals have been discussed. In the present (III) part the method of calculation known as Philipp-Taft method has been described in details and an improved modification of its one variant has been proposed. Additionally, in both these parts the results of calculations for GaAs (as a well known semiconductor) have been presented to be next employed to test the methods of calculation improved in this part.

The analysis of errors and accuracies for these methods is widely presented in the part IV. Finally, the part V contains the results of calculations of optical constants for Zn_3As_2 and Zn_3P_2 , carried out by using the improved methods of calculation and basing on our own measurement of optical properties of these semiconductors.

1. Introduction

The method used most often for calculation of the optical constants in semiconductors from the results of reflection coefficient measurements is the method due to Philipp-Taft (PT method) and more specifically to many of its variants (see [1]). An essential feature of the PT method is that it is based on the assumption that the reflection coefficient decreases exponentially as a function of energy, except for the highest energy, for which this coefficient is known.

The detailed analysis of various variants of this method presented in this paper is based on testing calculations of optical constants for GaAs and on the remarks met in the literature. The energy dependences of the optical constants of GaAs have been taken from the work [2]. In this analysis a number of shortcomings of the PT method have been

* Work sponsored under the Research Project PR-3.08.

described, a detailed algorithm being also given for some variants of the PT method which is rarely met in the literature. Also, an improved modification of one of the most complicated variants of the PT method given by TULVINSKI and TERENTEV in [3] has been suggested.

2. Analysis of the Philipp-Taft method

Optical constants of semiconductors may be determined when knowing the reflection coefficient R , and the change of phase Θ occurring during reflection. The first value may be obtained from the measurements, while the second one may be found by taking account of the dispersion formula

$$\Theta(E) = \frac{E}{\pi} P \int_0^{\infty} \frac{\ln R(x)}{E^2 - x^2} dx. \quad (1)$$

By assuming that the measurement of the reflection coefficient has been made within the interval (E_a, E_b) the integral (1) may be splitted into three components

$$\Theta_1(E) = \frac{E}{\pi} P \int_0^{E_a} \frac{\ln R(x)}{E^2 - x^2} dx, \quad (2a)$$

$$\Theta_2(E) = \frac{E}{\pi} P \int_{E_a}^{E_b} \frac{\ln R(x)}{E^2 - x^2} dx, \quad (2b)$$

$$\Theta_3(E) = \frac{E}{\pi} P \int_{E_b}^{\infty} \frac{\ln R(x)}{E^2 - x^2} dx. \quad (2c)$$

The estimation of $\Theta_2(E)$ is reduced to the respective numerical integration. An accurate description of the applied method will be given in the part IV of this work [4].

The determination of $\Theta_1(E)$ is the next problem. For this purpose independent measurements of the refractive index n within the region of semiconductor transparency, i.e., below the absorption edge [5-7], are often used. The reflection coefficient is then obtained from the relation

$$R(E) = \left[\frac{n(E) - 1}{n(E) + 1} \right]^2 \quad (3)$$

which enables the determination of $\Theta(E)$ from (2a). When such measurements are not available, the extrapolation (for instance, by a straight line $R(E) = R(E_a)$, [8, 9]) is performed.

The problem of determination of $\Theta_3(E)$ remains still open which is connected with the extrapolation of the relation $R(E)$ outside the end point of the measurement. The chosen extrapolation curve should satisfy the following requirements: be smooth, tend to zero at infinity, like E^{-4} (then $\Theta(E) \rightarrow \pi$, when $E \rightarrow \infty$), assure the continuity of $R(E)$

and dR/dE at the point E_b and cause that $\Theta(E) \geq 0$ for each E and $\Theta(E) = 0$ in the region of transparency. The last condition is often replaced by the requirement that the calculated coefficient of absorption be consistent with that measured close to the absorption edge [6, 10, 11]. Thus, assuming some parametrized form of the reflection curve outside the end measurement point, the values of these parameters are determined so that the above assumptions be fulfilled.

The PT method is based on the extrapolation

$$R(E) = B \times E^A, \text{ for } E \geq E_b. \quad (4)$$

The continuity condition $R(E)$ in E_b leads to determination of the constant B and to writing the formula (4) in the following form

$$R(E) = R(E_b) \left(\frac{E}{E_b} \right)^A. \quad (5)$$

Taking account of the above relation, the formula (2c) for $\Theta_3(E)$ may be transformed to the form convenient for calculations

$$\Theta_3(E) = \frac{\ln R(E_b)}{2\pi} \ln \left| \frac{E - E_b}{E + E_b} \right| + \begin{cases} -\frac{A}{\pi} \sum_{m=0}^{\infty} \left(\frac{E_b}{E} \right)^{2m+1} \frac{1}{(2m+1)^2}, & \text{for } E < E_b, \\ \frac{A}{\pi} \sum_{m=0}^{\infty} \left(\frac{E_b}{E} \right)^{2m+1} \frac{1}{(2m+1)^2} - \frac{\pi A}{4}, & \text{for } E > E_b. \end{cases} \quad (6)$$

The further step is to determine the value of the parameter A . It may be assumed that $A = -4$, as it was the case in the works [8, 12–14]. Then, however, some discontinuities of dR/dE may happen at E_b , and also nonphysical behaviour of Θ may be expected in the transparency region. The latter is realized by $\Theta < 0$, which results in negative values of the absorption coefficient. Such a situation may happen very often in the transparency region where the values of Θ should be close to zero.

Another variant of the PT method consists in determination of the parameters A from the continuity condition for dR/dE at E_b [15]

$$A = \frac{E_b}{R(E_b)} \left. \frac{dR}{dE} \right|_{E=E_b}. \quad (7)$$

In this case, however, the fulfilment of the remaining conditions concerning the experimental curve may not be controlled. Moreover, it happens that $R(E)$ increases at the end of the measurement interval, hence the values of A obtained from the formula (7) are positive, and thus nonphysical, since then $R(\infty) = \infty$.

The next, slightly better, variant of the PT method is based on the condition of zeroing of Θ in the transparency region. When $E \ll E_b$, which is usually the case in the trans-

parency region, the dependence (6) becomes linear

$$\Theta_3(E) \cong \left(-\frac{\ln R(E_b)}{\pi} - \frac{A}{\pi} \right) \frac{E}{E_b}, \text{ for } E \ll E_b. \quad (8)$$

In order to make $\Theta(E) = 0$ for $E < E_g$, the sum $\Theta_1(E) + \Theta_2(E)$ should be linear in this region. Due to many different reasons (the measurement error, for instance) the above requirement is not always satisfied. In consequence, the deviation of Θ from zero may be fairly high below the absorption edge. The results may be improved by introducing the extrapolation relations containing more parameters.

If only one parameter is to be determined it may be done (as it was the case in the variant of PT method described above) by demanding that for some $E_0 < E_g$, $\Theta(E_0) = 0$ [16] or that the curve $\Theta(E)$ be the closest to zero in the sense of least squares in the transparency region [17, 18]. The first way suffers from some nonuniqueness of the results caused by nonuniqueness of the choice of E_0 , which is avoided in the second way by using the least square procedure. In the first case the parameter A is defined by the following relation

$$A = \left(\Theta_1(E_0) + \Theta_2(E_0) + \frac{\ln R(E_b)}{2\pi} \ln \frac{E_b - E_0}{E_b + E_0} \right) \left[\frac{1}{\pi} \sum_{m=0}^{\infty} \left(\frac{E_0}{E_b} \right)^{2m+1} \frac{1}{(2m+1)^2} \right]^{-1}, \quad (9)$$

while in the second one the value of the parameter A will result from the fitting procedure. The essential shortcoming of this method is the fact that the received values of the parameter A are positive and thus nonphysical.

In order to eliminate this shortcoming the parameter E_f is introduced [10]. This parameter determines the upper integration limit in the integral (eq. (2c)), in other words

$$\Theta_3(E) = \frac{E}{\pi} P \int_{E_b}^{E_f} \frac{\ln R(x)}{E^2 - x^2} dx. \quad (10)$$

Then the dependences (6), (8) and (9) take the following form

$$\Theta_3(E) = \frac{\ln R(E_b)}{2\pi} \ln \left| \frac{(E - E_b)(E + E_f)}{(E + E_b)(E - E_f)} \right| + \begin{cases} \frac{A}{\pi} \sum_{m=0}^{\infty} \left[\left(\frac{E}{E_f} \right)^{2m+1} - \left(\frac{E}{E_b} \right)^{2m+1} \right] \frac{1}{(2m+1)^2}, & E < E_b < E_f, \\ \frac{A}{\pi} \sum_{m=0}^{\infty} \left[\left(\frac{E_b}{E} \right)^{2m+1} + \left(\frac{E}{E_f} \right)^{2m+1} \right] \frac{1}{(2m+1)^2} - \frac{\pi A}{4}, & E_b < E < E_f, \end{cases} \quad (11)$$

$$\Theta_3(E) \cong \left(\frac{E_b}{E_f} - 1 \right) \left(\frac{A}{\pi} + \frac{\ln R(E_b)}{\pi} \right) \frac{E}{E_b}, \quad E \ll E_b < E_f, \quad (12)$$

$$A = \left[\Theta_1(E_0) + \Theta_2(E) + \frac{\ln R(E_b)}{2\pi} \ln \frac{(E_b - E_0)(E_f + E_0)}{(E_b + E_0)(E_f - E_0)} \right] \times \left\{ \frac{1}{\pi} \sum_{m=0}^{\infty} \left[\left(\frac{E_0}{E_b} \right)^{2m+1} - \left(\frac{E_0}{E_f} \right)^{2m+1} \right] \frac{1}{(2m+1)^2} \right\}^{-1}. \quad (13)$$

As may be seen from the formula (13), in this case $A = A(E_f)$ and by a suitable choice of E_f it is possible to enter the range of nonpositive values of A , thus to avoid the non-physical situation appearing earlier. Here, for given E_0 one obtains an infinite number of pairs of A and E_f such that $\Theta(E_0) = 0$ and A is nonpositive, moreover, for different pairs A and E_f being the solution of eq. (13), $\Theta_3(E)$ behaves almost identically in the region of transparency which results from the relation (12). On the other hand, for high energies the discrepancies between the values of $\Theta_3(E)$ may be significant. Thus, an additional condition must be here introduced to obtain a unique result. The requirement that E_f be maximum seems to be a reasonable condition. From a simple analysis of the formula (12) it follows that A increases with the increase of E_f . Thus, E_f be maximal for the greatest possible A , i.e., $A = 0$. This leads to extraordinarily simple calculations (formula (11) for $A = 0$). However, the method presented is not rigorous from the mathematical viewpoint, neither the condition of decreasing of $R(E)$ with E^{-4} is fulfilled for high energies.

In order to eliminate the shortcomings described above, a more complicated variant of PT method is introduced [6, 16, 19]. This one differs from the variant described above by the fact that the extrapolation described by the following formula

$$R(E) = R(E_f) \left(\frac{E}{E_f} \right)^{-4} \quad (14)$$

is applied additionally in the interval (E_f, ∞) . Now, E_f need not to be maximal, as it was the case earlier. Thus, it is possible to choose such value of E_f that the values of the parameter A assure the continuity of dR/dE in E_b .

3. Extrapolation of $R(E)$ by the power series

The variants of the PT method described in the previous Section are characterized by a very simple shape of extrapolation of $R(E)$ for high energies. The measurements of the coefficient R for high energies (see [20], for instance) show that it diminishes by no means monotonically to zero in this region, while its structure is very complex. This indicates the way of improving the methods of optical constant calculation, which may be achieved by assuming a more complex shape of extrapolation curve. It is realized by the method due to Leveque which, however, is rather troublesome in application [1]. Also, the suggestion of TULVINSKI and TERENCEV [3] may be here mentioned, a modified version of which will be shortly presented.

It is assumed that outside the final measurement point of the reflection coefficient $R(E)$ behaves in the following way

$$R(E) = a_0 f(E) \left[\frac{E}{E_b} \right]^{-4}, \quad (15)$$

where $f(E) \rightarrow 1$, when $E \rightarrow \infty$. The function $f(E)$ plays the role of the function modulating the simple relation $R(E) \sim E^{-4}$. By taking natural logarithm of the equality (15) and expanding $\ln f(E)$ in a power we easily obtain

$$\ln R(E) = \sum_{m=1}^{\infty} a_{2m} \left(\frac{E_b}{E} \right)^{2m} + \ln \left[a_0 \left(\frac{E}{E_b} \right)^{-4} \right], \quad (16)$$

taking account of the fact that $R(E)$ is an even function of E . The $\Theta_3(E)$ calculated from the relation (16) is expressed as follows

$$\begin{aligned} \Theta_3(E) = & \sum_{m=1}^{\infty} \frac{a_{2m}}{2\pi} \left[\sum_{i=1}^m \frac{2}{2i-1} \left(\frac{E_b}{E} \right)^{2m-2i+1} + \left(\frac{E_b}{E} \right)^{2m} \ln \left| \frac{E_b-E}{E_b+E} \right| \right] + \frac{\ln a_0}{2\pi} \ln \left| \frac{E_b-E}{E_b+E} \right| \\ & + \begin{cases} \frac{4}{\pi} \sum_{l=0}^{\infty} \left(\frac{E}{E_b} \right)^{2l+1} \frac{1}{(2l+1)^2}, & \text{for } E < E_b, \\ \pi - \frac{4}{\pi} \sum_{l=0}^{\infty} \left(\frac{E_b}{E} \right)^{2l+1} \frac{1}{(2l+1)^2}, & \text{for } E > E_b. \end{cases} \end{aligned} \quad (17)$$

An arbitrary number of coefficients a_{2m} are determined from the system of linear equations, by assuming that Θ is zero in suitable number of points below the value E_g . On the other hand, the value of the constant a_0 is obtained from the continuity condition of $R(E)$ in E_b

$$\ln a_0 = \ln R(E_b) - \sum_{m=1}^{\infty} a_{2m}. \quad (18)$$

In the method described the condition of continuity of dR/dR in E_b may be also considered in a simple way. Thus, all the conditions imposed on the extrapolating curve are satisfied.

In practice, when using this method it is necessary to restrict the series in (16) to few first terms, since the error caused by the inaccuracy of computer calculations increases quickly with the increase of m . This restriction to the few first terms of the series is justified by the fact that $\Theta_3(E)$ is a slow-varying function of energy in the transparency region.

4. Testing calculations for GaAs

As it has been mentioned in the Introduction a series of calculations of optical constants GaAs have been carried out in order to recognize the practical advantages and shortcomings associated with different methods used up till now and also the possibilities of their improvements.

The values of the reflection coefficient for GaAs ranging within 0–25.4 eV have been taken from the work [2]. The values of $\Theta_2(E)$ have been calculated for several different extrapolations under assumption that the reflection coefficient changes linearly between the points read out of the graph. Next, the optical constants were determined with the help of the well known relations. All the calculations have been made on the Odra 1305 computer.

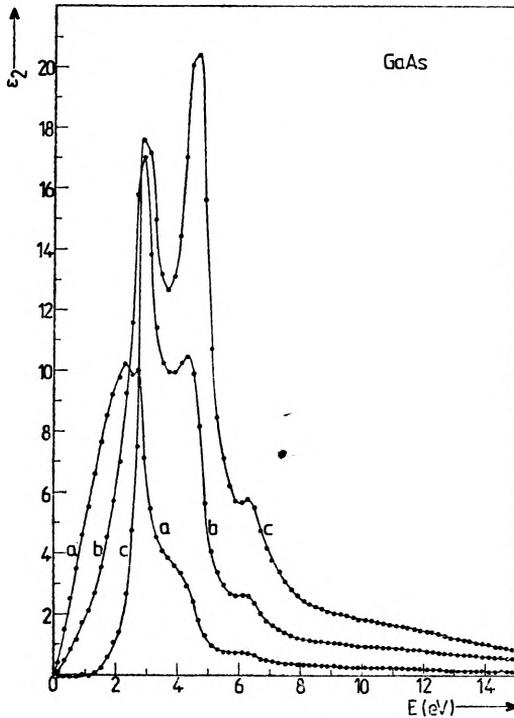


Figure 1 presents the obtained relations $\varepsilon_2(E)$ for three variants of the PT method described in Section 2, namely: for extrapolation according to (5) for parameter $A = -4$ (a), and $A = -12.29$ (b), determined from the continuity condition for dR/dE at the point E_b according to the formula (7). $\varepsilon_2(E)$ has been calculated also for the case described by relation (11), where A has been assumed to be zero and $E_f = 2000$ eV chosen so that the values of Θ be below $E_g = 1.35$ eV the closest to zero in the sense of least-squares procedure

Fig. 1. The energy dependence of the imaginary part of the complex coefficient of the dielectric permittivity obtained for the various variants of the Philipp-Taft method (see the description in text) for GaAs. The reflection spectrum of GaAs reported in [2] has been used for calculations

(c). There are great differences in the results obtained. The positions of the characteristic points suffer from the relatively smallest changes. For certain extrapolations some maxima of the curve may disappear, while the relative heights of maxima change. The value of the dielectric permittivity is determined in the least accurate way.

The calculations for the method described in Section 3 have been also made. In Figs. 2a and 2b the results obtained have been presented in comparison with the results published in the work [2]. The obtained agreement of the results is very high, considering the low read-out accuracy of the „muster” optical constant. For instance, the divergence in the values of $\varepsilon_2(E)$ for two main maxima are 1.6% for 2.8 eV, and 4.1% for 4.6 eV. The zeroing of values of Θ in the points 0.3 and 0.7 eV assumed in the calculations resulted in the following parameter values (when using the extrapolation given in (16)): $a_2 = 17.402$, $a_4 = -19.244$ and $a_0 = -6.925$.

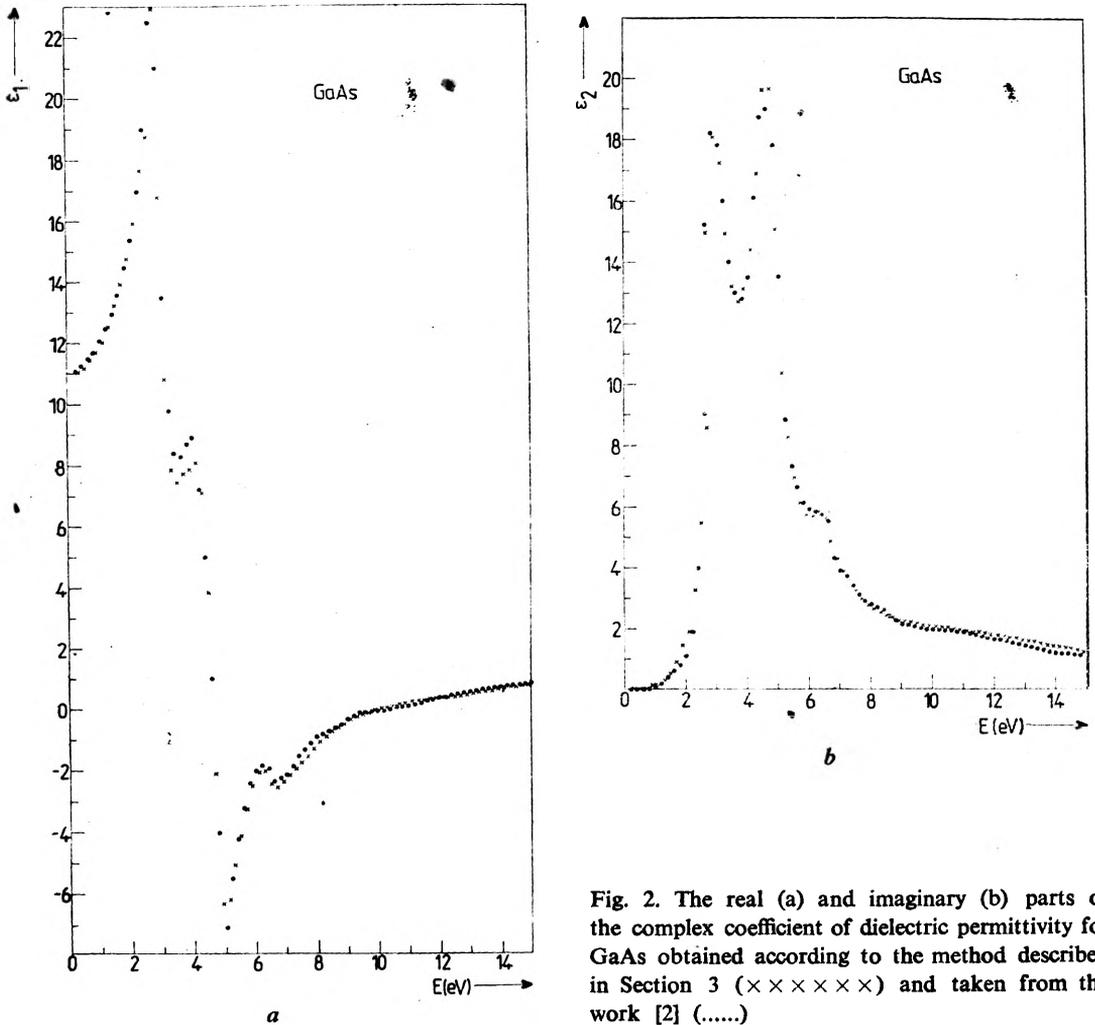


Fig. 2. The real (a) and imaginary (b) parts of the complex coefficient of dielectric permittivity for GaAs obtained according to the method described in Section 3 ($\times \times \times \times \times$) and taken from the work [2] (.....)

5. Conclusions

The results presented in Section 4 illustrate to what degree the right choice of the extrapolation method influences the final results. Hence the tendency, mentioned in Section 2, to find an extrapolation which would fulfil in the best way the basic conditions following from the physical model (i.e. lack of absorption below the fundamental edge $E < E_g$, the continuity of both the $R(E)$ function and its derivative at the border of the measurement and the extrapolation regions and its diminishing with E^{-4} for high energies) is justified. But also justified is an attempt of giving the extrapolation a more complex form which would better represent the expected distribution of reflection in the extrapolation region, although the fulfillment of the basic conditions would not be changed essentially. This is the task to be solved by the method presented in Section 3 of this paper. Unfortunately, it does not solve all the difficulties connected with the calculations of optical

constants which will be considered in the next part of this work devoted to a detailed analysis of the calculational errors [4].

The fact that in the course of last years very primitive methods have been often used to calculate optical constants may be explained by noticing that some of these methods may be just unexpectedly good for a special spectrum and that often only the basic structure of the obtained spectra is of concern. But, as it is indicated in Fig. 1, when applying the simple method of calculations and restricting ourselves only to the analysis of the basic structure of the spectra an error may be made if only one simple variant of the PT method is applied. Under these conditions it may be stated that the application of the method described in Section 3 of this paper may give the solution of this problem.

References

- [1] JEZERSKI K, MISIEWICZ J., WNUK J., PAWLIKOWSKI J. M., *Optica Applicata* **11** (1981),
- [2] PHILIPP H. R., EHRENREICH H., *Phys. Rev.* **129** (1963), 1550.
- [3] TULVINSKI V. B., TERENCEV N. I., *Opt. i Spectr.* **28** (1970), 894.
- [4] JEZERSKI K., MISIEWICZ J., WNUK J., PAWLIKOWSKI J. M., *Optica Applicata* **12** (1982), 87.
- [5] RIMMER M. P., DEXTER D. L., *J. Appl. Phys.* **31** (1960), 775.
- [6] GRILLI E., et al., *Solid State Commun.* **27** (1978), 105.
- [7] SLIVKA V. Yu., et al. *Ukr. Fiz. Zhurnal* **19** (1974), 1870.
- [8] LISITZA M. P., et al., *ibidem* **22** (1977), 1484.
- [9] GOTLIEB M., *J. Opt. Soc. Am.* **50** (1960), 343.
- [10] PHILIPP H. R., TAFT E. A., *Phys. Rev.* **113** (1959), 1002.
- [11] MORRISON R. F., *Phys. Rev.* **124** (1961), 1314.
- [12] CARDONA M., GREENAWAY D. L., *Phys. Rev.* **133A** (1964), 1685.
- [13] BEAL A. R., HUGHES H. P., *J. Phys.* **C12** (1979), 881.
- [14] BÄUERLE D., et al. *Z. Phys.* **B29** (1978), 179.
- [15] HAUSER E., et al. *Phys. Rev.* **B19** (1979), 6331.
- [16] PHILIPP H. R., TAFT E. A., *Phys. Rev.* **136A** (1964), 1445.
- [17] GROSSO V., et al., *J. Phys.* **C5** (1972), 1101.
- [18] WHITE J. J., III, *J. Opt. Soc. Am.* **62** (1972), 212.
- [19] GÄHWILLER Ch., HARBEKE G., *Phys. Rev.* **185** (1969), 1141.
- [20] BEAUMONT J. H., BOURDILLON A. J., *J. Phys.* **C10** (1977), 761.

Received January 25, 1981

Определение оптических констант полупроводников Zn_3As_2 и Zn_3P_2 с помощью соотношения Крамерса-Кронига

Часть III. Усовершенствование метода Филиппа-Тафта

Настоящая работа состоит из пяти частей и посвящена главным образом результатам исследований оптических констант Zn_3As_2 и Zn_3P_2 , двух соединений типа II_3-V_2 с широким энергетическим интервалом. Результатом анализа существующего до настоящего времени состояния в области применения метода анализа Крамерса-Кронига (КК) для комплексного коэффициента отражения оказалась необходимость развития работ в дополнительных направлениях, описанных в трёх первых частях. В I части представлен критический анализ применяемых до настоящего времени методов расчёта, основанных на дисперсионных соотношениях Крамерса-Кронига. Обсуждены также методы расчёта, оптических констант, в которых не используется интеграл КК.

Во II и III частях описаны предложения эффективных усовершенствований методов расчёта, основанных на применении интеграла КК. В настоящей (III) части работы подробно описан метод расчёта, известный как метод Филиппа-Тафта, а также предложена улучшенная модификация одного из вариантов этого метода. Дополнительно, в обеих частях, приведены результаты расчёта для GaS, который послужил (в качестве полупроводника со сравнительно известными оптическими свойствами) для проверки усовершенствованных нами методов расчёта.

Анализ ошибок и точности этих методов очень подробно описан в IV части работы. Пятая же часть содержит результаты для оптических констант Zn_3As_2 и Zn_3P_2 , произведённых с помощью усовершенствованных методов расчёта, а также на основе наших результатов измерения оптических свойств этих полупроводников.