Acta Physica Hungarica

VOLUME 57, NUMBERS 1–2, 1985

EDITOR–IN–CHIEF

I. KOVÁCS

EDITORIAL BOARD

Z. BAY, R. GÁSPÁR, N. KÜRTI, K. NAGY,
L. PÁL, A. SZALAY, I. TARJÁN, B. TELEGDI,
L. TISZA, E. WIGNER

Akadémiai Kiadó, Budapest

ACTA PHYS. HUNG. APAHAQ 57 (1–2) 1–154 (1985) HU ISSN 0231–4428

ACTA PHYSICA HUNGARICA

A JOURNAL OF THE HUNGARIAN ACADEMY OF SCIENCES

EDITED BY
I. KOVÁCS

Acta Physica publishes original papers on subjects in physics. Papers are accepted in English, French, German and Russian.

Acta Physica is published in two yearly volumes (4 issues each) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1054 Budapest, Alkotmány u. 21

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
1389 Budapest P.O. Box 149

or to its representatives abroad.

Acta Physica is indexed in Current Contents, in Physics Abstracts and in Current Papers in Physics.

ACTA PHYSICA HUNGARICA

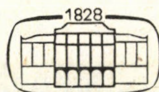
EDITORIAL BOARD

Z. BAY, R. GÁSPÁR, N. KÜRTI, K. NAGY, L. PÁL, A. SZALAY, I. TARJÁN, B. TELEGDI, L. TISZA, E. WIGNER

EDITOR-IN-CHIEF

I. KOVÁCS

VOLUME 57



AKADÉMIAI KIADÓ, BUDAPEST

1985

CONTENTS

Volume 57

GENERAL PHYSICS

Radiative self-interaction in classical field theory. *P. Kálmán* 119

ELEMENTARY PARTICLES AND FIELDS

Space-time foliations. *G. Sardanashvily* 31

The foundations of non-associative classical field theory. *D. F. Kurdgelaidze* 79

ATOMIC AND MOLECULAR PHYSICS

The expectation value of the inelastic loss in heavy atom collisions for keV energies. *I. Nagy, J. László and J. Giber* 113

Mössbauer effect study of Libyan desert silica glass. *H. A. Sallam, N. A. Eissa, I. Dézsi and D. L Nagy* 125

Electronic stopping power for different elements and inelastic energy loss in collision cascades. *I. Nagy, J. László and J. Giber* 149

CONDENSED MATTER

A note on the ferromagnetic lattice gas: study of the metastable state. *C. Palmieri* 3

Математическая модель флуктуационной равновесной термодинамики, *В. Б. Роганков* ... 13

Interface state density and capture cross section measurements on MOS structures by TSC. *D. Marton and A. Hübler* 41

Low energy electron diffraction by thin films. Part I. *K. Stachulec* 55

Thermodynamic properties of anharmonic noble metals in terms of interatomic potentials from experimental phonon spectra. *C. Malinowska—Adamska* 69

Study of defect morphology in the metal and dielectric layers of MIS structure by SIMS and SEM. *P. I. Didenko, V. G. Litovchenko, R. I. Marchenko, G. F. Romanova, M. Németh—Sallay, I. C. Szép and I. Cseh* 73

Correlation properties and the structural instability of the twodimensional anharmonic crystal. The virial approach. *V. K. Pershin and I. S. Gersht* 99

Determination of the mean free path of electrons in solids from the elastic peak. Part I. *A. Jablonski, P. Mrozek, G. Gergely, M. Menyhárd and A. Sulyok* 131

Determination of the mean free path of electrons in solids from the elastic peak. Part II. *G. Gergely, M. Menyhárd, A. Sulyok, A. Jablonski and P. Mrozek* 139

Preface to HCCG-3. *I. Tarján* 159

The history of crystal growth. *J. Bohm* 161

Single crystal research — crystal quality control — industrial production. *R. Voszka* 179

Phantom controlled automatic Czochralski puller sensing the crystal weight. *B. Mező, L. Bálint* and *R. Voszka* 185

Finite element analysis of process control and operating limits in edge-defined filmfed growth (EFG) of silicon and sapphire ribbons: a review. *J. P. Kalejs, H. M. Ettouney* and *R. A. Brown* . 189

Preparation of aluminium–sapphire bicrystals. *H. G. Lindenkreuz, M. Jurisch* and *S. Sahling* 205

Preparation of ZnO single crystals by chemical transport reaction and determination of ZnO homogeneity range. *H. Oppermann, G. Stöver, A. Heinrich, K. Teske* and *E. Ziegler* 213

Preparation of GaAs structures for different devices by chloride transport. *T. Görög, I. Gyuró* and *K. Somogyi* 223

Growth of PbSe crystals from lead melt. *V. Rakovits* and *M. Hársy* 233

Selective epitaxial growth of $Al_xGa_yIn_{1-x-y}Sb$. *E. Lendvay, L. Petrás* and *V. A. Gevorkian* 237

Heat conduction vs crystal defects in GaSb ingots grown by the vertical Bridgman method. *M. Hársy, F. Koltai* and *T. Görög* 245

X-ray and SEM investigation of crystal defects in N-type, P-type and semi-insulating GaAs. *P. Franzosi* and *G. Salviati* 251

LEC-growth and the major electrical and structural characteristics of semi-insulating gallium arsenide. *R. Fornari, C. Paorici, L. Zanotti* and *G. Zuccalli* 263

Electrochemical etching. *Chr. Nanev* and *K. Dicheva* 271

Investigation of morphology and surface structure of small NaCl particles. *M. Rösler* 285

The growth of TeO_2 crystals from the boiling solution. *L. Ötsi, E. Hartmann* and *B. Vajna* 295

Growth and luminescence properties of GaSb single crystals. *W. Lewandowski, M. Gajewska, B. Pastuszka* and *M. Bugajski* 303

Evolution of impurity centres in NaCl:Ca single crystals under laser radiation. *M. V. Galustashvili* and *D. G. Driyaev* 309

KDP Q-switches and second harmonic generator for high-power solid-state lasers. *J. Bakos, T. Juhász, Cs. Kuti* and *L. Vannay* 313

On the variations in surface diffusion coefficient of LFM crystal. *M. Rak* 317

Preliminary studies on crystal growth of $Pb_{1-x}Cd_xS$ solid solutions. *P. Şteţiu* 323

Bunching of monomolecular steps during growth and evaporation of crystals. *L. Malicskó* 335

Macro- and microanalytical investigations of impurities during production of TeO_2 single crystals. *L. Malicskó, I. Cravero* and *L. Ötsi* 347

TITLES OF OTHER PAPERS PRESENTED AT HCCG 3 355

CONTENTS

GENERAL PHYSICS

Radiative self-interaction in classical field theory. *P. Kálmán* ... 119

ELEMENTARY PARTICLES AND FIELDS

Space-time foliations. *G. Sardanashvily* ... 31

The foundations of non-associative classical field theory. *D. F. Kurdgelaidze* ... 79

ATOMIC AND MOLECULAR PHYSICS

The expectation value of the inelastic loss in heavy atom collisions for keV energies. *I. Nagy, J. László and J. Giber* ... 113

Mössbauer effect study of Libyan desert silica glass. *H. A. Sallam, N. A. Eissa, I. Dézsi and D. L. Nagy* ... 125

Electronic stopping power for different elements and inelastic energy loss in collision cascades. *I. Nagy, J. László and J. Giber* ... 149

CONDENSED MATTER

A note on the ferromagnetic lattice gas: study of the metastable state. *C. Palmieri* ... 3

Математическая модель флуктуационной равновесной термодинамики, *В. Б. Роганков* ... 13

Interface state density and capture cross section measurements on MOS structures by TSC. *D. Marton and A. Hübler* ... 41

Low energy electron diffraction by thin films. Part I. *K. Stachulec* ... 55

Thermodynamic properties of anharmonic noble metals in terms of interatomic potentials from experimental phonon spectra. *C. Malinowska—Adamska* ... 69

Study of defect morphology in the metal and dielectric layers of MIS structure by SIMS and SEM. *P. I. Didenko, V. G. Litovchenko, R. I. Marchenko, G. F. Romanova, M. Németh—Sallay, I. C. Szép and I. Cseh* ... 73

Correlation properties and the structural instability of the twodimensional anharmonic crystal. The virial approach. *V. K. Pershin and I. S. Gersht* ... 99

Determination of the mean free path of electrons in solids from the elastic peak. Part I. *A. Jablonski, P. Mrozek, G. Gergely, M. Menyhárd and A. Sulyok* ... 131

Determination of the mean free path of electrons in solids from the elastic peak. Part II. *G. Gergely, M. Menyhárd, A. Sulyok, A. Jablonski and P. Mrozek* ... 139

Manuscript received by Akadémiai Kiadó: 14 February 1984
Manuscript received by the Printers: 16 February 1984
Date of publication: 4 July 1985

PRINTED IN HUNGARY
Akadémiai Kiadó és Nyomda, Budapest

Acta Physica Hungarica 57 (1—2), pp. 3—11 (1985)

A NOTE ON THE FERROMAGNETIC LATTICE GAS: STUDY OF THE METASTABLE STATE

C. PALMIERI*

Institute of Theoretical Physics, 80125 Napoli, Italy

(Received 19 May 1983)

The Bethe—Peierls approximation is applied to a ferromagnetic lattice gas to study the behaviour of the system in the metastable region. The values of the critical indices γ above and below T_c are calculated

The Bethe—Peierls approximation in the simple form given by Domb [1] gives straightforward results for the calculations of the thermodynamics quantities such as the spontaneous magnetization, the internal energy etc. and the critical behaviour in the ferromagnetic Ising model or lattice gas.

Here we study the ferromagnetic lattice gas with the Bethe—Peierls approximation, assuming translational invariance, and discuss the metastable behaviour for different values of the magnetic field H.

Since, as it is well known [2], the Bethe—Peierls approximation gives exact results on the Bethe lattice, some of our results on the nature of the transition along the ferromagnetic critical line can be compared with Eggarter [3], Müller—Hartmann and Zittarz [4], paying attention to the fact that these authors do not neglect the surface effects. Finally since we have an explicit form for the susceptibility we can study its critical behaviour for $T > T_c$ and $T < T_c$ around the critical point. We find for $T > T_c$ ($H=0$) the results given in [1] for the critical index γ; that is in this region $\gamma = 1$. Furthermore we find γ also below T_c, that is $\gamma' = 1 (T < T_c,\ H=0)$. This result is in agreement with that of Katzura and Takizawa [5] and is expected in terms of Fischer's conjecture.

Let our system be described by the lattice gas Hamiltonian

$$-\beta\kappa = K \sum_{\langle ij \rangle} n_i n_j + \Delta \Sigma_i n_i, \tag{1}$$

where $n_i = 1$ if site i is occupied, 0 otherwise, $\beta = \frac{1}{K_B T}$, K is the nearest neighbour coupling constant related to the Ising coupling constant K_I by $K_I = \frac{K}{4}$, Δ is the

* Istituto Nazionale di Fisica Nucleare-Sezione di Napoli.

chemical potential related to the Ising magnetic field H and the coordination number $c = \sigma + 1$ by $-2H = \Delta + \frac{c}{2}K$. Positive and negative values of K correspond respectively to ferromagnetic and antiferromagnetic coupling. The sum $\sum_{\langle ij \rangle}$ is over nearest neighbours.

As it is well known in the Bethe—Peierls approximation Hamiltonian (1) is studied on a reduced lattice. For the ferromagnetic case $K > 0$ it is sufficient to consider an elementary cluster of a central site and its nearest neighbours on which acts the

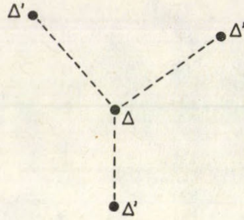


Fig. 1. Reduced lattice for $\sigma = 2$

auxiliary field Δ' (Fig. 1). The derivation of the susceptibility and of the critical line for the susceptibility is straightforward. For the cluster of Fig. 1 the partition function is

$$Z = [1 + e^{\Delta'}]^{\sigma+1} + e^{\Delta}[1 + e^{\Delta'+K}]^{\sigma+1}. \tag{2}$$

If ρ and ρ' denote respectively the densities (or magnetization per site) on the central site and on its nearest neighbours, the translational invariance $\rho = \rho'$, i.e.

$$\rho \equiv \frac{d}{d\Delta} \ln Z = \frac{1}{\sigma+1} \frac{d}{d\Delta'} \ln Z \equiv \rho'$$

implies

$$e^{\Delta} = e^{\Delta'} \frac{[1 + e^{\Delta'}]^{\sigma}}{[1 + e^{\Delta'+K}]^{\sigma}}, \tag{3}$$

therefore we can write

$$\rho = \frac{1}{Z} e^{\Delta}[1 + e^{\Delta'+K}]^{\sigma+1} = \frac{e^{\Delta'}(1 + e^{\Delta'+K})}{(1 + e^{\Delta'}) + e^{\Delta'}(1 + e^{\Delta'+K})}. \tag{4}$$

From (4) the auxiliary field Δ' can be given as a function of ρ, i.e.

$$e^{-\Delta'} = \frac{(1 - 2\rho) \pm \sqrt{(1 - 2\rho)^2 + 4e^{K}(1 - \rho)\rho}}{2\rho}. \tag{5}$$

The susceptibility can be easily evaluated by means of Eqs (3) and (5)

$$\chi = \frac{d\rho}{d\Delta} = \frac{d\rho}{d\Delta'} \frac{d\Delta'}{d\Delta}. \tag{6}$$

We obtain

$$\chi=\rho(1-\rho)\frac{1+\dfrac{e^{\Delta'+K}}{1+e^{\Delta'+K}}-\dfrac{e^{\Delta'}}{1+e^{\Delta'}}}{1-\sigma\left(\dfrac{e^{\Delta'+K}}{1+e^{\Delta'+K}}-\dfrac{e^{\Delta'}}{1+e^{\Delta'}}\right)}, \tag{7}$$

which diverges along the line

$$\left(\frac{d\Delta'}{d\Delta}\right)^{-1}=0, \tag{8}$$

therefore the critical line for the susceptibility is:

$$\rho_{\text{crit}}=\frac{1}{2}\left(1\pm\sqrt{1+\frac{4\sigma}{(\sigma-1)^2(1-e^K)}}\right). \tag{9}$$

Fig. 2 is a plot of two of these lines in the $(\rho, T/T_c)$ plane. The value $\rho=\frac{1}{2}$ is reached at the Curie point K_c such that

$$e^{K_c/2}=\frac{\sigma+1}{\sigma-1}. \tag{10}$$

Eq. (8) represents the spinoidal and describes the critical behaviour of metastable states.

Eq. (9) can be studied in the (H, T) plane giving:

$$e^{2H}=\left[\frac{(\sigma+1)e^{-K}+(\sigma-1)\pm g}{2\sigma e^{-K/2}}\right]^\sigma\left[\frac{(\sigma+1)e^{-K}-(\sigma-1)\pm g}{2e^{-K/2}}\right], \tag{11}$$

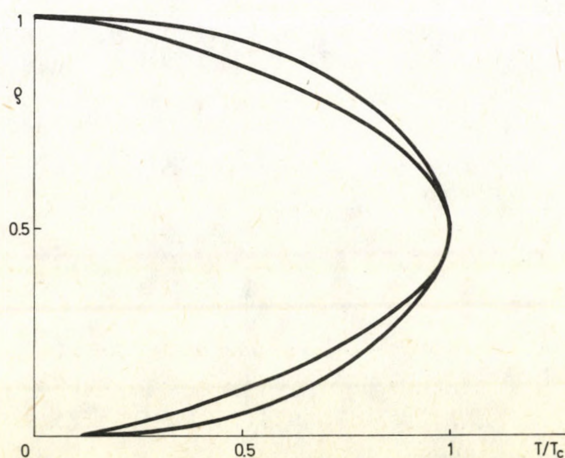


Fig. 2. Two spinoidal lines for $\sigma=3$ and $\sigma=11$ in the $(\rho, T/T_c)$

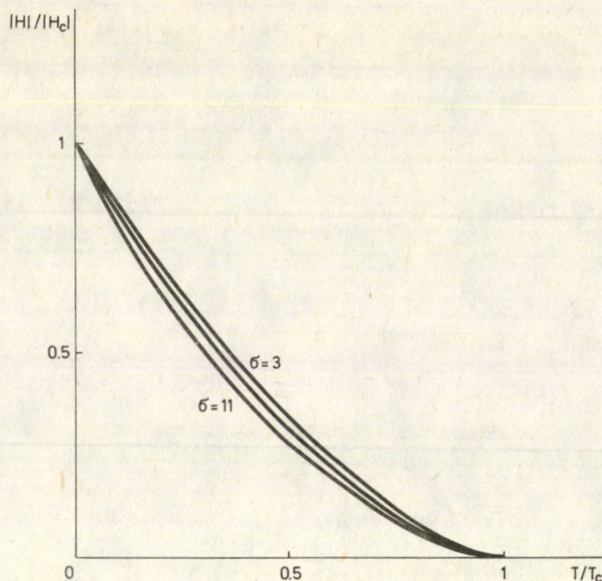


Fig. 3. The spinoidal lines for $\sigma=3$, $\sigma=5$, $\sigma=11$ in the plane of the reduced variables $(H/H_c, T/T_c)$

where

$$g^2=(1-e^{-K})\,[(\sigma-1)^2-(\sigma+1)^2e^{+K}]\,.$$

In Fig. 3 we have plotted the critical lines in the $(H/H_c, T/T_c)$ plane for $\sigma=3, 5, 11$.

Let us now investigate on the character of the singularity along the spinoidal critical line for different values of the magnetic field H. To do this we have done a computer calculation of Eq. (3) written as

$$\varphi=e^{-2H-\sigma}\varphi, \tag{12}$$

where

$$\varphi=e^{\Delta'+K/2} \tag{13}$$

and

$$\bar{\varphi}=\frac{e^{-K/2}+\varphi}{1+e^{-K/2}\varphi} \tag{14}$$

come from considering the density written in the following useful form, i.e.

$$\rho=\frac{\varphi\bar{\varphi}}{1+\varphi\bar{\varphi}}\,. \tag{15}$$

One can see that in the $H=0$ case Eq. (12) has a unique solution for $T>T_c$ at $\varphi=1$, and three solutions for $T<T_c$. The susceptibility given by Eq. (7) can be evaluated for each solution.

Figure 4 shows the values of the magnetization $m=2\rho-1$ in the $H=0$ case. We find that χ is positive on the solid line and negative on the dashed line, which therefore

represents an unphysical solution: m_+ and m_- denote the spontaneous magnetizations. At T_c, χ is obviously divergent. For $H<0$ and $T<T_c(H)$ there are always three solutions. In Fig. 5 a) and b) the magnetizations are plotted for the following values of the magnetic field $H/K=-0.1$ and $H/K=-0.05$. In both cases the dashed line represents the unphysical magnetization; m_- is stable, m_+ is unstable and represents a residual

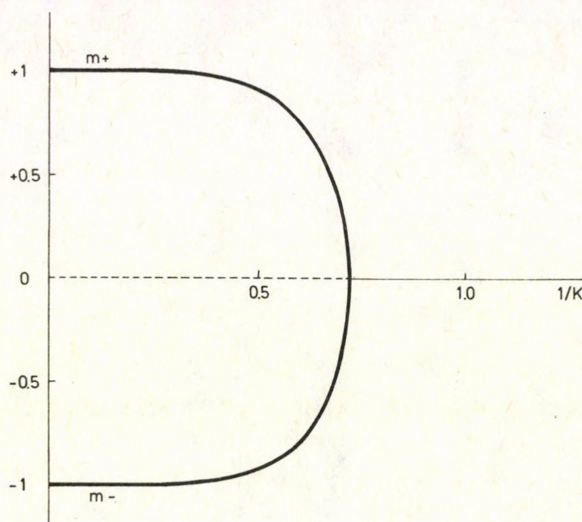


Fig. 4. The magnetization at $H=0$. On the dashed line $\chi<0$, while on the solid line $\chi>0$; at $T_c\chi=\infty$. m_+ and m_- are the spontaneous magnetizations

magnetization which survives up to $T=T_c(H)$. For $T>T_c(H)$ we find only the right magnetization m_-.

We can see that m_- is the stable solution for $H<0$ not only on a physical group but also by examining for each φ the behaviour of the free energy which can be written

$$F=-2H\rho-\frac{\sigma+1}{4}K\left[\rho\frac{e^{K/2}\varphi}{1+e^{K/2}\varphi}+(1-\rho)\frac{1}{1+e^{K/2}\varphi}\right]+$$

$$+\sigma[\rho\ln\rho+(1-\rho)\ln(1-\rho)]+\frac{\sigma+1}{2}\left\{\rho\frac{e^{K/2}\varphi}{1+e^{K/2}\varphi}\cdot\ln\rho\frac{e^{K/2}\varphi}{1+e^{K/2}\varphi}+\right.$$

$$\left.+\frac{(1-\rho)}{1+e^{-K/2}\varphi}\ln\frac{1-\rho}{1+e^{-K/2}\varphi}+\frac{2\rho}{1+e^{K/2}\varphi}\ln\frac{\rho}{1+e^{K/2}\varphi}\right\}.$$

From the previous analysis it is clear that all along the spinoidal line we have a first order transition, except for $H=0$ and $T=T_c$ where the transition is of the second order.

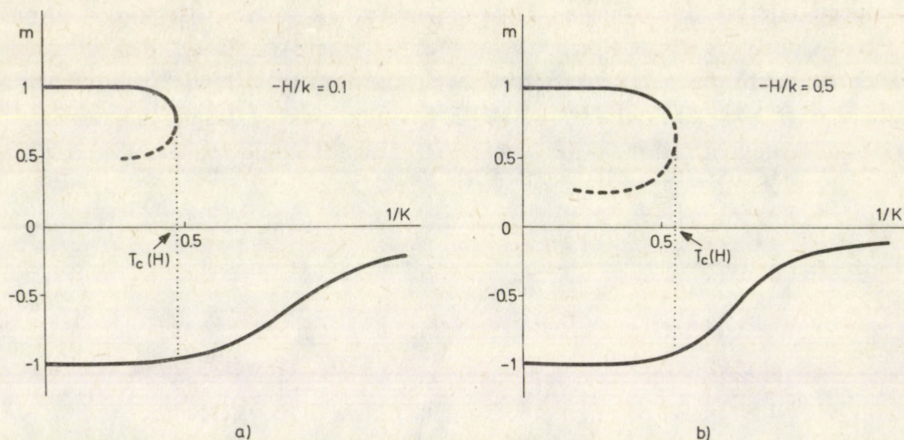


Fig. 5. a) – the magnetizations for $H/K = -0.1$ b) — the magnetizations for $H/K = -0.05$. Observe that $T_c(H)$ is greater in b). The dashed and the solid lines have the same meaning as in Fig. 4

Let us analyze the behaviour of the susceptibility above and below T_c and give the values for the critical indices γ, γ'. In terms of the magnetic variables Eq. (6) is

$$\chi = \frac{dM}{dH} = 4\frac{d\rho}{d\Delta},$$

moreover Eq. (7) can be written as

$$\chi = 4\rho(1-\rho)\left(1 + \frac{\varphi(e^K - 1)(\sigma+1)}{(e^{K/2}+\varphi)(1+e^{K/2}\varphi) - \sigma(e^K-1)\varphi}\right). \tag{16}$$

For $H=0$ and $T > T_c \left(\text{i.e. } \rho = \frac{1}{2}\right)$ we recover the classical result quoted in Domb [1].

Indeed for $T > T_c$ $\varphi = 1$ is the unique solution and since near K_c

$$Z \simeq Z_c + C_3|K - K_c|, \tag{17}$$

where

$$Z = e^{K/2} \quad \text{and} \quad Z_c = \frac{\sigma+1}{\sigma-1} \quad C_3 = \pm\frac{Z_c}{2}$$

we have

$$\chi_0 \simeq \frac{1}{(\sigma-1)|K-K_c|} \simeq \frac{1}{|T-T_c|}, \tag{18}$$

therefore the critical index $\gamma = 1$.

For $H=0$ and $T < T_c$, i.e. $\rho \simeq \frac{1}{2}$ we can solve analytically Eq. (12) for $\sigma = 2$ and find near K_c that

$$\varphi \simeq 1 + c_1|K-K_c|^{\frac{1}{2}} + c_2|K-K_c|. \tag{19}$$

At the end of this note we will show that this behaviour of φ is true for every σ. Let us now use (19) to estimate the behaviour of χ. By means of (17) and (19) we have

$$e^{K/2}+\varphi\simeq Z_c+1+c_1|K-K_c|^{\frac{1}{2}}+(c_2+c_3)|K-K_c|$$

$$1+e^{K/2}\varphi\simeq 1+Z_c+Z_c c_1|K-K_c|^{\frac{1}{2}}+(Z_c c_2+c_3)|K-K_c|$$

$$e^K\simeq Z_c^2+2Z_c c_3|K-K_c|$$

which replaced in (16) give

$$\chi_0=1+\frac{(\sigma+1)(Z_c^2-1)+R_1|K-K_c|^{\frac{1}{2}}+R_2|K-K_c|}{|K-K_c|(c_1^2 Z_c-2\sigma c_3)}\simeq\frac{1}{T-T_c},$$

where

$$R_1=c_1(\sigma+1)(Z_c^2-1)\qquad\text{and}\qquad R_2=c_2(\sigma+1)(Z_c^2-1)+2Z_c c_3,$$

therefore the value of the critical index is $\gamma'=1$.

The fact that the critical indices are symmetric above and below T_c has been first shown by Fisher [6] with classical Landau theory and the exact 2-dimensional Ising model. The result we find here agrees with the results obtained by Domb and Guttman [7] by series expansions. These authors found that the complexities of the behaviour of the low temperature series for different types of lattices could reasonably be explained taking account of the Caley tree type configurations.

On the other hand it is known [8], [9] that the high temperature $(T>T_c)$ behaviour is dominated by Caley tree type ramified configurations.

Our result is therefore as expected since the Bethe approximation coincides with the Caley tree solution.

We remark that the tree configurations are the ones that more naturally give rise to the metastable states since their linear character makes them grow quickly. On the other hand they are not suitable to describe nucleation because of the absense of intramolecular bonds.

This argument could explain why on the treatment of the site-bond correlated percolation in the Bethe approximation the spinoidal line is recovered almost exactly [10]. On the other hand with the Migdal—Kadanoff Renormalization Group the Coniglio—Klein site-bond correlated percolation does not give the metastable behaviour [11], [12], therefore a modification of the definitions is necessary in order to select the ramified configurations. This problem is under investigation.

Let us now prove that for every σ, φ behaves for $T<T_c$ as in (19). Eq. (12) at $H=0$ magnetic field defines the function

$$F(\varphi, Z)\equiv\varphi(Z+\varphi)^\sigma-(1+Z\varphi)^\sigma, \tag{20}$$

which is obviously equal to zero for every φ solution of Eq. (12) in $H=0$.

Let us expand F near $\varphi=1$; since $F(1,Z)=0$, we have

$$F(\varphi,Z)=(\varphi-1)F'(1,Z)+(\varphi-1)^2\frac{F''}{2!}(1,Z+)(\varphi-1)^3\frac{F'''}{3!}(1,Z). \tag{21}$$

Moreover, since $F(\varphi,Z)\simeq 0$ for $\varphi=1$ and $Z\simeq Z_c$, retaining the terms up to order three in (21) we have near Z_c

$$(\varphi-1)\left[F'(1,Z)+(\varphi-1)\frac{F''}{2!}(1,Z)+(\varphi-1)^2\frac{F'''}{3!}(1,Z)\right]=0. \tag{22}$$

Near $Z_c (Z\geq Z_c)$ we have

$$F'(1,Z)\simeq F'(1,Z_c)+\left.\frac{\partial F'}{\partial Z}\right|_{Z_c}(Z-Z_c)=A_1+B_1(Z-Z_c),$$

$$F''(1,Z)\simeq F''(1,Z_c)+\left.\frac{\partial F''}{\partial Z}\right|_{Z_c}(Z-Z_c)=A+B(Z-Z_c),$$

$$F'''(1,Z)\simeq F'''(1,Z_c)+\left.\frac{\partial F'''}{\partial Z}\right|_{Z_c}(Z-Z_c)=A_3+B_3(Z-Z_c).$$

One can see that $A_1=A_2=0$

$$B_1=(Z_c+1)^{\sigma-1}(1-\sigma),\qquad B_2=2\sigma(1+Z_c)^{\sigma-2},(-\sigma)$$

$$A_3=\sigma(\sigma-1)(1+Z_c)^{\sigma-3}\{3(Z_c+1)+(\sigma-2)(1-Z_c^2)\}.$$

Therefore (22) becomes

$$B_1(Z-Z_c)+(\varphi-1)\frac{B_2}{2}(Z-Z_c)+\frac{(\varphi-1)^2}{3!}|A_3+B_3(Z-Z_c)|=0,$$

hence

$$(\varphi-1)\simeq\frac{-\frac{B_2}{2}(Z-Z_c)\pm(Z-Z_c)^{\frac{1}{2}}\left(\frac{B_2}{2}\right)^2(Z-Z_c)-\frac{4}{3!}B_1|A_3+B_3(Z-Z_c)|}{2\left|A_3+\frac{B_3}{3!}(Z-Z_c)\right|}$$

retaining the lower order terms we have

$$(\varphi-1)\simeq-\frac{B_2 3!}{4A_3}(Z-Z_c)\pm(Z-Z_c)^{\frac{1}{2}}\sqrt{-\frac{B_1 3!}{A_3}},$$

i.e.

$$(\varphi-1)\simeq-\frac{B_2 3!}{4A_3}\frac{Z_c}{2}|K-K_c|\pm|K-K_c|^{\frac{1}{2}}\sqrt{-\frac{Z_c B_1 3!}{2A_3}}.$$

References

1. C. Domb, Adv. in Phys., *9*, 1949, 1960.
2. D. N. Burley, "Closed form approximation for lattice system" in Domb—Green: Phase transitions and critical phenomena, Vol. II, Academic-Landau, 1973.
3. T. P. Eggarter, Phys. Rev., *B9*, 2989, 1973.
4. E. Muller—Hartmann and J. Zittarz, Phys. Rev. Letters, *33*, 893, 1974.
5. S. Katsura and M. Takizawa, Progr. of Theor. Phys., Vol. 51, n. 1. p. 82, 1974.
6. M. E. Fisher, Rep. Progr. Phys., *30*, 615, 1967.
7. C. Domb and J. Guttmann, J. Phys. C: Solid State Phys., *3*, 1652, 1970.
8. C. Domb, J. Phys. C: Solid State Phys., *6*, 39, 1973.
9. C. Domb, J. Phys. A: Math. Gen., *9*, 283, 1976.
10. F. Di Liberto, G. Monroy and C. Palmieri, to appear in J. Phys. A. 1982.
11. A. Coniglio and W. Klein, J. Phys. A.: Math. Gen., *13*, 2775, 1980.
12. A. Coniglio, F. Di Liberto and G. Monroy, J. Phys. A: Math. Gen., *14*, 3017, 1981.

Acta Physica Hungarica 57 (1—2), pp. 13—30 (1985)

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ФЛУКТУАЦИОННОЙ РАВНОВЕСНОЙ ТЕРМОДИНАМИКИ

В. Б. Роганков

Одесский технологический институт холодильной промышленности, Одесса, СССР

(Поступило 12 июня 1983)

Предлагается математическая модель флуктуационной равновесной термодинамики (ФРТ), позволяющая установить существенно динамическую структуру данной теории в случае учёта флуктуаций интенсивных переменных. Последние рассматриваются в качестве локальных операторов термодинамического пространства состояний, что наряду с введением формального параметра — термодинамического «времени» τ, связанного с масштабом системы, приводит к изоморфности модели ФРТ с квантово-механической теорией измерений.

Методом изучения локальных флуктуаций в предлагаемой работе является операция непрерывного изменения (уменьшения) масштаба измерения некоторого заданного объёма системы. Данный приём близок по смыслу процедуре масштабного преобразования, использованной Вильсоном при построении микроскопической теории критической области, но важное отличие заключается, как в противоположном выборе направления изменения масштаба, так и в существенно макроскопическом подходе к изучению равновесного состояния флюида в математической модели ФРТ. Установлено, что в развиваемой теории локальное давление подсистем играет роль термодинамического локального потенциала, на основе которого может быть выведено уравнение, определяющее термодинамическое «движение» состояния.

Показано, что непрерывное преобразование масштаба объёма вызывает спонтанное нарушение симметрии состояния и возникновение дуальных характеристик данного процесса — параметров порядка и беспорядка, имеющих в модели ФРТ смысл квазисредних по Боголюбову. Их отношение определяется в работе, как мера симметрии состояния (совпадающая с энтропией, отнесённой к частице), которая сохраняется при изменении τ.

1. Введение

Равновесная термодинамика (РТ) занимает особое положение в ряду теорий, рассматривающих поведение флюидных систем, состоящих из большого числа частиц, поскольку объектом её изучения являются состояния равновесия, инвариантные относительно физического времени t. Отсюда принято считать [1] т.1, что в РТ динамическая теория (ДТ) играет тривиальную роль, так как равновесные состояния соответствуют, независящим от t, решениям уравнения движения, описывающего эволюцию системы. В данной работе предлагается введение в РТ таких основных понятий ДТ, как время τ (будем

называть его «термодинамическим», Т-временем, чтобы отличать от физического времени t, определяющего изменение состояния, например, в неравновесной термодинамике) и уравнение движения (также «термодинамического», Т-движения), что позволит нам установить существенно динамическую структуру РТ. Интуитивно понятно, что целью указанного подхода является расширение традиционных представлений РТ для конструктивного рассмотрения таких важных и до сих пор нерешённых полностью проблем флюидного состояния, как построение теоретически обоснованного уравнения состояния (УС), а также теории, описывающей с единых позиций, как область однофазных равновесных состояний (1), так и области двухфазных равновесных состояний (2) и критических явлений (3). Отсутствие физической теории данной общности особенно явно сказывается при попытке интерпретировать с помощью УС (ван-дер-Ваальсова типа) классической РТ поведение вещества в окрестности критической точки [2] т.5.

С физической точки зрения основанием работы является последовательный учёт флуктуаций интенсивных параметров РТ, определяющих равновесное состояние термодинамической системы. Математическая модель флуктуационной РТ (ФРТ), формулируемая ниже, позволяет установить важность роли, которую играют, при выявлении динамической структуры РТ, флуктуации интенсивных переменных не только в области (3), где их учёт привёл к созданию теории [3], удовлетворительно описывающей экспериментальные данные, но и в областях (1), (2), где общим выводом флуктуационной теории [2] т.5 является указание на пренебрежимую малость флуктуаций экстенсивных параметров макроскопической равновесной системы. Последнее обстоятельство не исключает возможности значительных локальных флуктуаций интенсивных переменных РТ в малых объёмах данной системы. Средством изучения указанных локальных флуктуаций, определяющих, по нашему мнению, физическую сущность всех трёх (1—3) возможных типов равновесных состояний однокомпонентного флюида, является в предлагаемой модели ФРТ операция непрерывного изменения (уменьшения) масштаба измерения некоторого постоянного объёма системы V_0. Данный приём напоминает процедуру масштабного преобразования, использованную при построении микроскопической теории критической области (3) вещества [3, 4]. Следует отметить, однако, что помимо противоположного выбора направления изменения масштаба (в методе ренормгруппы [4] рассматриваются всё более крупные блоки модельной системы) в математической модели Т-движения используется существенно макроскопическое представление о равновесном состоянии флюида. Здесь имеется в виду, что состояние вещества в любом (и, в том числе, микроскопическом с точки зрения классической РТ) произвольном объёме системы предполагается далее равновесным, если при изменении масштаба V некоторая величина, характеризующая состояние системы в целом, остаётся инвариантной. Отсюда следует, что в основе предлагаемого подхода лежит то же представление о

локальном равновесии, использованное как в микроскопической теории критической области [4], так и в макроскопическом подходе обобщённой термодинамики [5, 6], целью которого явилась экстраполяция РТ далеко за пределы устойчивости термодинамической ветви. Как будет показано ниже, по сравнению с упомянутыми методами, предположение о локальном равновесии в модели Т-движения в большей степени отвечает физическому содержанию задачи, которая в качестве начального рассматривает произвольное равновесное состояние макроскопической системы в объёме V_0. Это обстоятельство позволяет надеяться на адекватное описание с помощью формулируемой ниже модели ФРТ поведения системы для всей шкалы явлений — от макроскопических масштабов до микроскопических — на основе единого математического формализма.

2. Моделирование равновесного состояния

Начальным этапом процесса моделирования является определение вида рассматриваемых математических переменных. Учёт флуктуаций заставляет нас отказаться от принятого в классической РТ чисто детерминированного определения интенсивных параметров. В дальнейшем используется смешанный характер описания, где наряду с величинами, которые могут быть, в принципе, точно измерены (V, τ, $\bar{x}$, t) будут введены стохастические переменные, вероятностный смысл которых, как следует подчеркнуть, не связан в данной работе с их определением в статистической механике [2] т.5.

Введём в терминах гамильтоновой ДТ [2] т.1, 5 понятие фазового Т-пространства, каждая точка которого характеризуется набором (q, p) сопряжённых обобщённых координат q и обобщённых импульсов p и изображает собой одно из допустимых равновесных макросостояний системы. Элементы множества интенсивных, локально определённых переменных $u \in \mathcal{U}$ образуют набор (q, p) гладко меняющихся в фазовом Т-пространстве величин и точка данного пространства имеет, согласно классической РТ, две ($n=2$) термодинамических степени свободы, соответствующие значениям указанных функций в этой точке.

Перейдём к формулировке модели ФРТ и установим смысл интенсивных переменных u в данной модели. Исследуем равновесную систему тождественных точечных частиц, рассматриваемую, формальным образом, как реализацию понятия Т-предела [1] т.1 с некоторым важным дополнением ($\star$), суть которого будет обсуждена ниже. Назовём указанную предельную систему L-системой и определим пространственную структуру модели Т-движения наличием в каждый момент Т-времени такого количества конгруэнтных кубических V_τ в произвольной подсистеме V_0 L-системы, что выполняется соотношение:

$$V_0 = \tau V_\tau . \quad (2.1)$$

Отсюда выбор фиксированного значения V_0 соответствует заданию начала отсчёта ($\tau=1$ при $V_\tau=V_0$) на шкале Т-времени. Допускается, что положение подсистемы V_0 может быть свободно выбрано в физическом пространстве $\bar{x}$ L-системы и объём, дополняющий V_0 до полного объёма V_L, плотно заполнен в каждый момент τ конгруэнтными кубическими подсистемами V_τ. В задаваемой, таким образом, решёточной структуре L-системы начальный момент Т-времени характеризуется значением объёма V_0, допустимыми состояниями которого считаются состояния, совместимые с упомянутым выше условием (★), для каждой из бесконечного числа подсистем ансамбля, плотно заполняющего конгруэнтными кубами V_0 объём V_L предельной L-системы. Введение понятий указанного ансамбля (назовём его ансамблем V_τ) и L-системы является чисто формальным приёмом изучения равновесного состояния реальной открытой системы объёмом $V=V_0$, которая (как и любая из подсистем ансамбля V_τ) обменивается с окружением веществом и энергией. Важно отметить, что мыслимая процедура деления объёма V_0 на плотно заполняющие его V_τ не нарушает, естественно, равновесного состояния исходной системы. Будем считать далее, что термодинамические интенсивные переменные $u \in U$ определяют каждое из допустимых состояний однокомпонентной системы тождественных частиц для конечного интервала значений τ, естественные границы которого задаются величинами V_0 (начального объёма) и V_G (объёма порядка размера объёма отдельной молекулы). Для каждого τ из этого интервала выполняется соотношение (2.1).

Допустим теперь существование бесконечного числа шкал измерения объёма V, отличающихся друг от друга величиной единицы измерения V_τ, меняющейся непрерывно от V_0 до V_G. Тогда моменты Т-времени τ_i и τ_j можно упорядочить таким образом, что равновесное состояние подсистемы V_{τ_i} предшествует состоянию V_{τ_j}, если выполняется неравенство: $\tau_i < \tau_j$ Последовательная трансформация масштаба измерения объёма, приводящая к непрерывному делению начального значения V_0 на всё более малые подсистемы V_τ и составляет суть, предлагаемой в настоящей работе, модели Т-движения.

Покажем, что данный подход позволяет естественным образом ввести в рассмотрение флуктуации интенсивных параметров РТ. Для равновесного начального состояния подсистемы V_0, однозначно определяемого в рамках РТ средними значениями экстенсивных переменных N, S и E (N—число частиц; E—внутренняя энергия; S—полная энтропия), зафиксируем момент Т-времени τ и исследуем изменение состояния, принадлежащей V_0 произвольной подсистемы объёмом V_τ, в течение некоторого отрезка физического времени t. Эмпирический принцип макроскопической эквивалентности [1] т.1, утверждающий, что при изучении локального поведения макроскопической системы объём V является несущественным параметром, позволяет дать общее определение интенсивных переменных РТ в виде:

$$u_\tau = u_L + \Delta u_\tau(V_\tau)\,, \tag{2.2}$$

где u_L—объёмный предел интенсивной величины, характеризующий L-систему, а Δu_τ—зависящий от V_τ член с асимптотикой:

$$\lim_{V_\tau \to \infty} \Delta u_\tau(V_\tau) = 0 . \tag{2.3}$$

В общем случае, величина u_τ должна зависеть также от пространственных координат $\bar{x}$ и физического времени t. Это означает, что проводя «измерение» состояния отдельной подсистемы V_τ (при некотором заданном $\bar{x}$) в разные моменты физического времени t, мы будем получать дискретные значения u_τ, изменяющиеся по вероятностным законам. Понятно, что физической причиной, вызывающей указанное изменение, является взаимодействие выделенной подсистемы V_τ с её окружением.

3. Изоморфность модели ФРТ с квантовомеханической теорией измерения

Процедура «измерения» состояния в модели Т-движения допускает интерпретацию, полностью согласующуюся с квантовомеханической теорией измерения [2] т.3, [7]. Рассматриваемая подсистема V_τ, состояние которой в данный момент физического времени t мы «измеряем», взаимодействует с окружающей её большой системой — «прибором», находящейся в определённом, известном состоянии. Практически с абсолютной достоверностью малая подсистема V_τ имеет при этом вполне определённое состояние равновесия, «измеряемое» значением u_τ системы — «прибора», поведение которой предполагается с достаточной точностью подчиняющимся в модели Т-движения законам классической механики. Процесс «измерения» приводит, таким образом, к изменению состояния системы — «прибора», которое служит количественной характеристикой состояния подсистемы V_τ. С другой стороны, воздействие на объект «измерения» окружающей подсистему V_τ большой системы принципиально не может быть сколь угодно слабым при данной точности «измерения». Чем выше эта точность, тем сильнее воздействие, оказываемое на подсистему V_τ.

Из предыдущего видно, что использованная нами терминология квантовой механики (КМ) [2] т.3 позволяет, самым общим образом, описать флуктуации интенсивных переменных РТ. Развивая указанный подход, введём вектор состояния $\bar{r}$ в фазовом Т-пространстве системы, который может изменяться двумя способами: а) непрерывным образом при изменении τ и б) по вероятностным законам при измерении координат этого вектора. Наличие флуктуаций интенсивных переменных позволяет ввести в модель ФРТ малую, но всегда конечную величину неопределённости, с которой известно равновесное состояние каждой подсистемы, принадлежащей L-системе (аналог кванто-

вомеханической постоянной Планка [1] т.1):

$$\delta p \delta q \geqslant h_T . \quad (3.1)$$

Здесь и далее знак δ указывает, что речь идёт о флуктуации интенсивной переменной ФРТ. Отсутствие, в данном смысле, полной информации о системе приводит при её описании к наличию статистики и равновесное состояние должно задаваться, в принципе, не числом, а некоторым распределением вероятностей. В том случае, когда указанное распределение известно для заданного τ, то дальнейшее Т-движение состояния будет подчинено в модели ФРТ законам классической механики.

Известно, что описание состояния квантовой системы осуществляется меньшим числом величин, чем в классической механике, где для этого должны быть заданы в некоторый момент времени $2n$ сопряжённых обобщённых координат q и обобщённых импульсов p. На основании рассмотренной выше аналогии модели ФРТ и КМ введём комплексную, в общем случае, волновую Т-функцию координат $\Psi(q)$, квадрат модуля которой $|\Psi(q)|^2$ даёт плотность вероятности того, что система находится в точке (q) конфигурационного Т-пространства. Укажем здесь, что предполагается задание с помощью $\Psi(q)$ максимально возможной информации о состоянии системы. Точные значения волновой Т-функции определяют, так называемые, чистые состояния [2] т.3, которые рассматриваются без учёта флуктуаций в классической РТ.

Покажем, что последовательный учёт флуктуаций интенсивных переменных РТ, в предлагаемой нами модели Т-движения, приводит к выводу об отсутствии полной системы измерений, дающей однозначно предсказуемый результат для состояния V_0. Рассмотрим в каждый момент Т-времени τ L-систему, как замкнутую систему, состоящую из произвольной подсистемы V_τ и соответствующей системы — «прибора», «измеряющей» состояние данной подсистемы. То обстоя тельство, что L-система, а также система — «прибор» с достаточной точностью могут считаться классическими объектами и обладать, таким образом, определёнными волновыми Т-функциями не гарантирует, однако, наличия воличия волновой Т-функции и для подсистемы V_τ в любой момент τ. Процесс «измерения» состояния V_τ, описанный выше, предполагает взаимодействие объекта «измерения» и системы — «прибора» и данное «измерение» соответствует установлению определённого статистического соотношения между «измеряемой» величиной u_τ и состоянием системы — «прибора». Отсюда интервал физического времени t, в течение которого производится «измерение» в модели Т-движения при заданном τ, должен быть настолько велик, чтобы исключить возможность рассмотрения подсистем V_τ в качестве квазизамкнутых и обладающих, таким образом, волновой Т-функцией. Отметим, что в этом состоит принципиальное отличие предлагаемой модели ФРТ от статистикомеханического подхода [2] т.5, в котором представление о квазизамкнутости макроскопических подсистем, приводящее к статистической

независимости, прямо следует из рассмотрения этих подсистем в течение такого промежутка физического времени t, когда взаимодействием их можно пренебречь.

Введём теперь, в полном соответствии с КМ [2] т.3, Т-матрицу плотности, описывающую состояние подсистемы V_τ:

$$s(q_\tau, q'_\tau) = \int \Psi^*(q; q'_\tau)\Psi(q; q_\tau)dq\,, \tag{3.2}$$

где q_τ и q обозначены, соответственно, совокупность координат подсистемы V_τ и системы — «прибора», а $\Psi(q; q_\tau)$ — волновая Т-функция L-системы. Диагональные элементы $s(q_\tau, q_\tau)$ определяющие распределение вероятности для координат подсистемы V_τ выражаются как:

$$s(q_\tau, q_\tau) = \int |\Psi(q; q_\tau)|^2\, dq\,. \tag{3.3}$$

Исследуем понятие Т-матрицы плотности, введение которого связано с отсутствием в модели Т-движения полной системы «измерений», однозначно предсказывающей результат для состояния V_0. В произвольный момент Т-времени τ у нас имеется бесконечное количество тождественных копий подсистемы V_τ, макроскопические состояния которых изображаются точками фазового Т-пространства и совместимы с некоторой информацией (★) о начальном состоянии подсистемы V_0, являясь, в этом смысле, допустимыми состояниями. Вполне очевидно, что принимаемая пространственная структура модели Т-движения, отвечает заданию, для каждого фиксированного значения τ, ансамбля бесконечного числа подсистем V_τ и именно статистические свойства данного ансамбля описываются Т-матрицей плотности (3.2). Следует подчеркнуть, что вводя указанное определение ансамбля подсистем V_τ, мы отнюдь не предполагаем (как то принято в статистической механике [1] т. 1) замены двух типов усреднения: а) по физическому времени t и б) по термодинамическому времени τ на статистическое усреднение по ансамблю. Напротив, усреднение по ансамблю рассматривается далее, как первичное определение интенсивных переменных ФРТ, которые могут непрерывно изменяться с Т-временем τ и статистические флуктуации которых с течением физического времени t, приводят к наличию в модели Т-движения термодинамического соотношения неопределённостей (3.1). Отсюда понятно, что при формулировке модели ФРТ мы отказываемся от постулирования аналога эргодической теоремы [1] т. 2 в самом начале рассмотрения, как это делается в статистической механике. Такой подход связан с предполагаемым неустойчивым поведением состояния V_0 при значениях τ, превышающих некоторое граничное τ_c, что будет обсуждаться ниже. Выбор ансамбля, усредняющего интенсивные переменные ФРТ по Т-времени τ, в этом случае оказывается вполне произвольным и его целесообразно провести на основе предварительного исследования траекторий, как в области устойчивого, так и неустойчивого Т-движения. Вопросы существования и определения временных средних для неустойчивых динамических систем (ДС),

как известно, далеко не тривиальны [8, 9] и связаны со структурой естественных распределений вероятностей в данных системах.

Если ограничить множество возможных начальных состояний подсистемы V_0 их наиболее общим видом, а именно однофазными равновесными состояниями (1), то фазовое Т-пространство допустимых состояний однокомпонентной системы в данном случае представляет собой в классической, не учитывающей флуктуаций интенсивных переменных РТ [10] гладкое подмногообразие размерности $2n=4$ в пятимерном евклидовом пространстве $\mathbb{R}^5$. Введём локальную координатную систему (u^1, u^2, u^3, u^4) в произвольной точке $\bar{r}$, принадлежащей указанному подмногообразию, являющуюся локальным диффеоморфизмом некоторого открытого подмножества $\mathbb{R}^4$. Отсюда, в каждой точке $\bar{r}$ имеется, непрерывно зависящая от координат фазового Т-пространства, единственная касательная гиперплоскость, проекция подмногообразия на которую взаимно однозначна вблизи точки касания. Таким образом, отдельное допустимое состояние РТ рассматривается в данном подходе с точки зрения классической гамильтоновой механики, что является сутью развиваемых в последнее время, так называемых геометрических формулировок РТ [10]. Последние обеспечивают инвариантность законов РТ относительно преобразования независимых переменных.

Как ясно из предыдущего, основная концепция данной работы не предполагает возможности получения однозначно предсказуемого результата (точки в фазовом Т-пространстве) для состояния V_0 и этим принципиально отличается от геометрического подхода [10—13]. В результате сохранения конечной неопределённости, в смысле уравнения (3.1), для любого момента τ, далее рассматривается распределение вероятностей только для конфигурационного Т-пространства системы $\mathscr{P}$, размерность которого равна $n=2$. Вычисление диагональных матричных элементов оператора плотности $s(q_r, q_\tau)$ позволяет тогда определить вероятность нахождения системы в точке $\bar{r}$, принадлежащей гладкому подмногообразию $\mathscr{P}$. Поскольку любое из трёх возможных типов допустимых равновесных состояний РТ должно соответствовать точке $\bar{r} \in \mathscr{P}$ в качестве исходной системы декартовых координат, характеризующих непрерывную поверхность состояний, из множества интенсивных переменных $\mathscr{U}$ следует выбрать подмножество, так называемых, полевых переменных, одновначно определяющих не только однофазные состояния (1), но и состояния равновесия сосуществующих фаз (2) и (3) [14, 15]. Отсюда, для рассматриваемой однокомпонентной системы вводится взаимно однозначное отображение множества допустимых состояний РТ во множество точек пространства полевых переменных $(\mu, T, P)=\bar{r}$, декартовы координаты которых удовлетворяют системе уравнений:

$$\mu=\mu(\rho, \sigma) \qquad T=T(\rho, \sigma) \qquad P=P(\rho, \sigma) \tag{3.4}$$

таким образом, в модели ФРТ допустимые состояния однозначно определены собственными значениями локальных операторов множества интенсивных переменных $\mathscr{U}$. Действительные параметры ρ и σ, входящие в уравнение (3.4), являются объёмными плотностями экстенсивных величин N и S, соответственно. Данные параметры неоднозначно определены в точках двухфазного равновесия (2) и принадлежат подмножеству плотностных интенсивных переменных РТ [14, 15].

Математическая модель Т-движения рассматривает равновесное состояние системы, используя бесконечное число масштабов объёма V_τ, изменяющихся непрерывно от макроскопических значений V_0 до величин, сопоставимых с объёмом отдельных молекул V_G. Разумеется, в модели ФРТ нет физических причин, препятствующих изучению бесконечного Т-времени τ, то есть континуального предела $\tau \to \infty$ дискретного множества подсистем V_τ, образующих в физическом пространстве $\bar{x}$ кубическую решётку. Выбирая конечное значение верхней границы Т-времени τ_G, мы хотим подчеркнуть тем самым то обстоятельство, что по достижению τ_G, модель ФРТ приобретает все наиболее характерные черты, так называемой модели Глаубера, трактующей микроскопическую модель решёточного газа в качестве ДС [16]. Вопрос о связи модели Глаубера с асимптотическим поведением при $\tau \to \tau_G$ модели ФРТ требует специального обсуждения и будет исследован отдельно.

4. Основные определения модели ФРТ

В статистической механике произвольному равновесному состоянию системы N частиц можно поставить в соответствие определённое минимальное значение объёма V_c, где данное состояние все ещё известно (если учитывать флуктуации интенсивных переменных с погрешностью не превышающей некоторых величин Δu_τ (2.2), характеризующих начальное состояние. Линейный размер $\sim V_c^{1/3}$ называется корреляционным радиусом, инвариантным относительно изменения масштаба длины системы. Последнее оправдывается эмпирическим принципом макроскопической эквивалентности, приводящим к произвольности выбора макроскопического образца системы при изучении её свойств и, таким образом, к понятно Т-предела, играющем важную роль в концепции статистической механики [1] т. 1:

$$\lim_{\substack{N \to \infty \\ V \to \infty}} \frac{N}{V} = \rho = \text{const} . \tag{4.1}$$

Применительно к модели Т-движения это означает, что в интервале $[1, \tau_c]$ изменение геометрической формы конгруэнтных подсистем V_τ и произвольный выбор покрытия объёма V_0 единицей измерения V_τ не будут приводить к

изменению абсолютной величины $\Delta u_\tau(V_\tau)$ и, таким образом, «измеряемые» значения интенсивных переменных не зависят, в этом смысле, от пространственных координат $\bar{x}$.

Определение Т-предела с помощью (4.1), вообще говоря, предполагает однородность интенсивных переменных по объёму системы. Отсюда следует, что равновесные состояния бесконечного набора макроскопически эквивалентных систем с различными значениями V и N должны изображаться одной и той же точкой $P=P(\mu, T)$ поверхности $\mathscr{P}$ при выполнении равенства $N/V=\text{const}$. Распространение принципа макроскопической эквивалентности на системы неоднородные в физическом пространстве — времени обеспечивается в статистической механике тем условием, что увеличение размеров систем указанного набора производится с полным сохранением симметрии начальной системы последовательности [1] т. 1. Граничные условия, требующиеся для выполнения этого правила, таким образом, доопределяют понятие Т-предела (4.1).

Вводимая нами модель ФРТ качественно отличается от описанной интерпретации принципа макроскопической эквивалентности в статистической механике тем, что не предполагает сохранения пространственно — временной симметрии системы при изменении τ. В связи с этим, определение понятия Т-предела (4.1) должно быть дополнено в модели Т-движения некоторым условием (★), упоминавшимся ранее. Для выяснения физического смысла (★) уточним, вначале, то значение, которое имеет корреляционный объём V_c в модели ФРТ:

Определение 1. Минимальный объём подсистемы $V_\tau = V_c$, в котором флуктуации плотностных интенсивных переменных $\delta\rho, \delta\sigma$ по ансамблю V_τ не превышают значений этих величин для подсистемы V_0, называется корреляционным объёмом V_c.

Значение V_c однозначно определено полевыми интенсивными переменными μ_L, T_L, характеризующими состояние L-системы. Строго говоря, вследствие того, что в модели Т-движения состояние подсистемы V_0 известно в каждый момент τ с точностью до $\delta\mu$ и δT, V_c является непрерывной функцией, изменяющейся в некотором ограниченном интервале значений, но для дальнейшего изложения данное уточнение несущественно и предполагается $V_c = V_c(\mu_L.\ T_L)$.

Состояние полного статистического равновесия определяется для замкнутой макроскопической системы [2] т. 5, как состояние в котором для любой макроскопической подсистемы интенсивные физические величины $u \in \mathscr{U}$ с большой относительной точностью (до флуктуаций δu) равны своим средним значениям. Введение ансамблей статистической механики позволяет при этом отказаться от изучения u с физическим временем t, заменяя усреднние по t на статистическое усреднение. Важно отметить, что с точки зрения модели ФРТ,

развиваемой в данной работе, такой приём означает переход к чисто детерминированному (не учитывающему флуктуационного неравенства (3.1) описанию интенсивных переменных u. Последнее справедливо в модели Т-движения только для области τ, соответствующей макроскопическим масштабам измерения объёма: $V_\tau \geqslant V_c$. Поскольку нами предполагается исследование с помощью общего формализма также микроскопических, сопоставимых с объёмом отдельных молекул, значений $V_\tau \in (V_c, V_G]$, ниже приводится, используемое далее, определение локального равновесия подсистемы V_τ в модели ФРТ:

> *Определение 2.* Каждому заданному значению $\tau \in [\tau_0, \tau_G]$ отвечают допустимые состояния локального равновесия подсистем ансамбля V_τ, в случае взаимно однозначного соответствия этих состояний точкам, принадлежащим некоторой, не увеличивающейся с ростом τ, ограниченной области Δ_j непрерывной поверхности полевых интенсивных переменных $P(\mu, T)$.

Отметим важное отличие определения (2) от традиционного определения в классической РТ равновесного состояния двух открытых, обменивающихся между собой веществом и энергией макроскопических подсистем, как состояния с общими для этих подсистем значениями μ и T. В модели ФРТ не предполагается инвариантность полевых переменных μ и T (при заданном τ) относительно изменения физического времени t и пространственных координат $\bar{x}$ подсистемы V_τ. Определение 2 требует выполнения более слабого (и с физической точки зрения более реального) условия, заключающегося в том, что изменение μ и T с Т-временем τ, то есть собственно Т-движение, должно осуществляться при сохранении конечной неопределённости значений $\delta\mu$ и δT и инвариантности функциональной зависимости:

$$P = P(\mu, T). \tag{4.2}$$

Укажем здесь, что масштабирование однокомпонентной системы с помощью V и физический смысл интенсивных полевых переменных μ и T, характеризующих обмен открытой подсистемы V_τ с её окружением, соответственно, веществом и энергией, приводят к выводу об исключительном положении давления P в подмножестве полевых переменных ФРТ. Данное положение полностью аналогично введению в рамках КМ (изоморфность которой модели ФРТ предполагается на основании предыдущего изложения) гамильтониана системы [2] т. 3. Смысл функции (4.2), как уравнения определяющего гиперповерхность гамильтоновского Т-оператора (давления P), на которой осуществляется Т-движение равновесного состояния V_0, позволяет сделать ряд качественных выводов о характере траекторий и общей топологии двухмерной поверхности $\mathscr{P}$, не решая уравнения Т-движения:

$$P_\tau = \tau[P_0], \tag{4.3}$$

где τ — рассматриваемое нами, преобразование масштаба измерения V.

Как следует из определения 2, вектор состояния $\bar{r}$ принадлежит в каждый момент $\tau \in [\tau_0, \tau_G]$ некоторой ограниченной области Δ_j непрерывной поверхности (4.2), соответствующей одному из возможных типов равновесных состояний системы j. Если к трём, перечисленным ранее областям $\mathscr{P}$ добавить область $P=0$ (4), то есть область идеального газа, то, тем самым, указанные возможности флюидного состояния однокомпонентной системы будут исчерпаны и, следовательно, $j=1, 2, 3, 4$. Отсюда, в принципе, начальные условия для V_0 могут быть выбраны соответствующими любому дискретному значению j. Основным физическим различием состояний с разными j является то, что корреляционный объём V_c имеет конечное значение для $j=1, 2$; равен бесконечности для $j=3$ и равен нулю для $j=4$. Задавая начальные условия для реального (конечного по величине) объёма подсистемы V_0, можно получить следующие соотношения между V_0 и V_c: $V_0 \gtrless V_c$, $j=1, 2$; $V_0 < V_c$, $j=3$; $V_0 > V_c$, $j=4$. Поскольку непрерывное увеличение τ приводит к уменьшению значения V_τ до некоторого малого V_G, следует отметить, что при Т-движении смысл данных неравенств для $j=3$, 4 будет сохраняться, тогда как для $j=1$, 2, по достижению $V_\tau = V_c$ он будет изменяться на противоположный, если вначале было $V_0 > V_c$.

Из определения 1 для корреляционного объёма V_c прямо следует, что переход из области $\tau \in [1, \tau_c]$ в область $\tau \in (\tau_c, \tau_G]$ приведёт к увеличению флуктуаций $\delta\rho$, $\delta\sigma$ плотностных интенсивных переменных. Физически это означает, что рассматриваемые в масштабе V_τ, $\tau \in (\tau_c, \tau_G]$, отдельные подсистемы существенно коррелируют в указанном интервале τ, обмениваясь между собой веществом и энергией. Состояния подсистем V_τ с $\delta\rho$ и $\delta\sigma$, значительно превышающими данные величины по сравнению с $V_\tau \geqslant V_c$, мгновенно возникают и мгновенно исчезают на шкале физического времени t, являясь неустойчивыми относительно t. Взаимодействие частиц флюида приводит, естественно, к полной стохастичности описанного процесса и отсутствию динамических составляющих, вызывающих движение частиц в каком-либо преимущественном пространственном направлении. Данный факт отсутствия макроскопического переноса вещества и энергии в объёме всей системы V_0 оправдывает распространение определения 2 локального равновесия в модели ФРТ также на область $V_\tau < V_c$, неустойчивых относительно физического времени t, состояний.

Принципиальными для дальнейшего построения модели Т-движения, являются вопросы о характере изменения $\delta\rho$ и $\delta\sigma$ относительно Т-времени τ в области $V_\tau < V_c$ и, в частности, вопрос о существовании инвариантов Т-движения в указанной области. Последнее прямо связано с изучением свойств симметрии предлагаемой модели ФРТ. Определим основное состояние предельной L-системы P_L, как состояние инвариантное относительно вращений и трансляций в четырёхмерном физическом пространстве — времени. Данное основное состояние будет инвариантным относительно Т-движения, или, другими словами, действия на P_L однопараметрической непрерывной группы преобразований τ в том случае, если под действием указанной группы оно переводится в

себя, являясь неподвижной точкой рассматриваемого преобразования:

$$\tau[P_L]=P_L. \quad (4.4)$$

Вполне очевидно, однако, что принятая нами пространственная структура модели ФРТ, в данный момент τ образующая бесконечную в каждом из трёх пространственных измерений $\bar{x}$ решётку кубов V_τ, нарушает трансляционную и вращательную симметрию L-системы, понижая её до кубической симметрии. Операторы интенсивных переменных u_τ задаются, таким образом, на дискретном множестве, имеющем в координатном пространстве $\bar{x}$ структуру решётки и являются, по существу, операторами квантовой теории поля на решётке [4], для которых инвариантность относительно преобразований Лоренца восстанавливается только в континуальном пределе $\tau\to\infty$. Важным свойством гамильтонианов, используемых в теории поля на решётке и имеющих вид, соответствующий связанным гармоническим или ангармоническим осцилляторам [4], является условие их локальности. Здесь имеется в виду, что Т-гамильтониан отдельной подсистемы V_τ зависит только от операторов u_τ, относящихся к подсистемам, непосредственно окружающим V_τ. Указанное свойство прямо использовано нами в определении 2, где состоянию локального равновесия в любой момент Т-времени $\tau\in[\tau_0, \tau_G]$ отвечает, не возрастающая при увеличении τ, ограниченная область Δ_j непрерывной поверхности полевых переменных (4.2).

Для $\tau\in[\tau_0, \tau_c]$ указанной области Δ_j соответствуют, как то следует из определения 1 корреляционного объёма, значения флуктуаций $\delta\rho$, $\delta\sigma$ не превышающие некоторых начальных $\delta\rho_0$ и $\delta\sigma_0$. При переходе в область $\tau\in[\tau_c, \tau_G]$ данные флуктуации существенно увеличиваются по абсолютной величине. При этом, как было отмечено выше, сильно увеличивается корреляция отдельных подсистем V_τ, что приводит к необходимости исследования кооперативного поведения всей системы, рассматриваемой, как целое. Аддитивный Т-гамильтониан системы P (величина механическая по определению) не содержит в этом случае, полной информации о поведении системы. Данная информация должна заключаться, естественным образом, в изменяющихся значениях $\delta\rho$ и $\delta\sigma$, причём это изменение должно осуществляться так, чтобы некоторая характеристика (★) ансамбля V_τ была в любой момент τ совместима с характеристикой (★) L-системы. Такой величиной, инвариантной относительно изменения τ, во всём рассматриваемом отрезке $\tau\in[\tau_0, \tau_G]$ является в модели ФРТ, в силу предполагаемой нами замкнутости L-системы, частичная плотность энтропии $s_L=S_L/N_L$ (★).

Рассмотрим физическое содержание инвариантности s_L при Т-движении системы. На основании предполагаемой нами изоморфности модели ФРТ с КМ интенсивные переменные $u\in\mathscr{U}$ имеют смысл абстрактных линейных операторов, действующих в гильбертовом пространстве [1] т. 1. Бесконечный объём предельной L-системы приводит, вообще говоря, к несепарабельному гильбертову пространству [4], но данное различие, по видимому, несущественно в

рамках модели ФРТ. Уравнение Т-движения в наиболее общем виде (4.3) обладает при этом симметрией, как по отношению к изменению знака τ, так и по отношению к трансляциям и вращениям в рассматриваемом гильбертовом пространстве. Отражением данной симметрии для каждого момента τ является инвариантность s_L при унитарных преобразованиях среди физически допустимых по определению 2 сотояний V_τ. Указанная величина соответствует состоянию полного статистического равновесия [2] т. 5 L-системы, не зависящему от Т-движения состояний локального равновесия её подсистем V_0.

Процесс измерения в КМ (и, следовательно, в развиваемой нами модели ФРТ) приводит, как это подчёркнуто в [2] т. 3, 5, к физической неэквивалентности обоих направлений времени. Последнее связано с тем очевидным фактом, что изменяя Т-время τ в противоположном направлении от некоторого его момента, мы будем получать допустимые состояния, неэквивалентные с микроскопической точки зрения состояниям, пройденным в положительном направлении τ. Указанная необратимость Т-движения является прямым следствием нарушения симметрии путём введения, как отмечено выше, конкретной пространственной структуры модели ФРТ. Нарушение симметрии характеризуется неинвариантностью основного состояния L-системы, которая проявляется в существовании неинвариантных средних этого состояния для ряда локальных операторов u, то есть в наличии флуктуаций интенсивных переменных δu. Неинвариантность основного состояния при инвариантности гамильтониана системы в теории многих тел носит название спонтанного нарушения симметрии [17, 18] и приводит к возникновению определённой упорядоченности системы, которая связана с возникновением макроскопического параметра порядка. В флюидной системе, модель которой рассматривается, данная величина, понижающая симметрию состояний, имеет смысл $\delta\rho$. С другой стороны, мерой необратимости в ФРТ естественно выбрать $\delta\sigma$, и изучать, таким образом, процесс Т-движения, как конкуренцию тенденций к упорядочиванию ($\delta\rho$) и беспорядку ($\delta\sigma$) при сохранении величины $s_L = \delta\sigma/\delta\rho$ *), интерпретируемой нами в качестве меры симметрии системы. В данном подходе $\delta\sigma$ приобретает смысл энтропии ДС, характеризующей в каждый момент Т-времени число различных, при заданной точности наблюдения (соотношение (3.1)), траекторий Т-движения. Отсюда следует принципиально важное для модели ФРТ.

Определение 3. Преобразование τ, нарушающее симметрию основного состояния системы, приводит в модели ФРТ как к появлению параметра порядка $\delta\rho$, связанного с упорядочиванием вследствие спонтанно нарушенной симметрии, так и к возникновению параметра беспорядка $\delta\sigma$, характеризующего стремление системы восстановить начальную симметрию основного состояния $s_L(^\star)$.

Нетрудно видеть, что вводимый, таким образом, параметр беспорядка $\delta\sigma$ близок по содержанию к используемому в ДТ понятию топологической

энтропии [9]. Последняя равна нулю в области периодического или квазипериодического поведения траекторий системы, где они устойчивы по Ляпунову (близкие вначале точки не расходятся при увеличении τ на значительное расстояние) и обращение топологической энтропии в нуль соответствует определению этой величины в пределе бесконечного физического времени. В области стохастического, непредсказуемого поведения траекторий системы указанное определение приводит к тому, что топологическая энтропия становится положительной [9]. Важно отметить, что наличие в модели ФРТ на шкале Т-времени момента τ_c, приводящего к существенной разнице абсолютной величины флуктуаций интенсивных переменных $\delta\rho$, $\delta\sigma$ для $\tau \leqslant \tau_c$ и $\tau > \tau_c$, соответственно, заставляет нас, вследствие конечных размеров отрезков $[\tau_0, \tau_c]$ и $(\tau_c\ \tau_G])$, исследовать Т-движение данных переменных именно в этих пределах и отказаться, таким образом, от их определения при $\tau \to \infty$. Последнее связано просто с выбором верхней границы τ_G, рассматриваемых τ, и не имеет принципиального влияния на, утверждаемую нами близость понятий параметра беспорядка в модели ФРТ и топологической энтропии в ДТ.

Определение 1 однозначно предполагает существование Т-временных средних $\delta\rho$ и $\delta\sigma$ для области $[\tau_0, \tau_c]$. Статистическое распределение данных параметров вблизи значений ρ_L и σ_L, соответствующих полному статистическому равновесию [2] т. 5, можно изобразить при этом в виде графика с чрезвычайно острым максимумом в районе указанных средних значений. Переход в область $[\tau_c, \tau_G]$ означает постепенное расплывание такого графика в силу увеличения интервала возможеных значений $\delta\rho$ и $\delta\sigma$. Если предположить существование пределов Т-временных средних данных параметров также для области неустойчивых по отношению к физическому времени t состояний $\tau > \tau_c$, то разность указанных пределов и значений Т-временных средних для текущего момента τ будет стремиться с увеличением τ к нулю, как $1/\tau$. В ДТ это означает, что рассматриваемая нами ДС относится к типу неустойчивых ДС, статистическое распределение которых в континуальном пределе $\tau \to \infty$ является классическим распределением Гаусса [8].

На практике, при изучении термодинамического поведения реальных систем, масштаб измерения V всегда фиксирован и нам требуется дать определение средних по ансамблю V_τ для параметров порядка $\delta\rho$ и беспорядка $\delta\sigma$ в произвольный момент Т-времени τ. Следует ещё раз подчеркнуть, что в модели ФРТ такое определение отнюдь не означает отказ от изучения Т-временного поведения указанных средних, как то принято в статистической механике [1] т. 1, где исследуется объёмный предел интенсивных величин u_L. Как уже отмечалось, необходимую для определения u_L асимптотику уравнения (2.3) обеспечивает понятие Т-предела (4.1), основанное на эмпирическом принципе макроскопической эквивалентности. Выше установлено, что глубокое отличие модели ФРТ от концепций классической РТ и статистической механики заключается в исследовании равновесного состояния с учётом спонтанно

нарушенной симметрии, тогда как замена одной реальной системы другой, макроскопически эквивалентной первой, прямо предполагает сохранение первоначальной симметрии задачи [1] т. 1 В соответствии с указанным различием, ансамбль V_τ системы с нарушенной симметрией определяется переходом к предельной L-системе, для которой, взамен условий (4.1), выполняется:

$$\lim_{\substack{N\to\infty\\ V\to\infty\\ S\to\infty}} \frac{\delta\sigma}{\delta\rho} = \frac{S_L}{N_L} = s_L = \text{const.} \qquad (\star)$$

Данное требование вполне эквивалентно определению Т-предела (4.1) в области $\tau \leqslant \tau_c$, но позволяет также ввести понятие ансамбля состояний V_τ для области сильно развитых флуктуаций интенсивных переменных $\tau > \tau_c$.

С учётом различия условий (4.1) и ($\star$), параметры порядка $\delta\rho$ и беспорядка $\delta\sigma$ могут быть определены в модели ФРТ, как квазисредние в смысле Боголюбова [17] для систем с нарушенной симметрией. Действительно, введём на пространстве $\mathscr{P}$ (пространство Т-гамильтониана флюидного состояния) инфинитезимальное преобразование, переводящее начальное состояние системы P_τ в состояние $P_{\tau+\delta\tau}$:

$$P_{\tau+\delta\tau} = P_\tau + \delta\tau \cdot \tau[P_\tau] \qquad (4.5)$$

При этом действие указанного преобразования рассматривается нами, как возникновение инфинитезимального внешнего поля, нарушающего симметрию Т-гамильтониана P_τ. Отсюда, в соответствии с работой [17], введём определение $\delta\rho$ и $\delta\sigma$ в модели ФРТ, как квазисредних по Боголюбову:

$$\delta\rho = \{\delta\rho(\bar r)\} = \lim_{\delta\tau\to 0} \lim_{S_L} \langle \delta\rho(\bar r)\rangle_{\delta\tau} \qquad (4.6)$$

$$\delta\sigma = \{\delta\sigma(\bar r)\} = \lim_{\delta\tau\to 0} \lim_{S_L} \langle \delta\sigma(\bar r)\rangle_{\delta\tau} \qquad (4.7)$$

где $\langle\ \rangle$ означает среднее по ансамблю V_τ, с Т-гамильтонианом P_τ. Отметим, что определение $\delta\rho$ и $\delta\sigma$ с помощью (4.6) и (4.7) обеспечивает физически корректное определение интенсивных переменных ФРТ, поскольку на первом месте в указанных формулах стоит предел $\delta\tau\to 0$, что сохраняет конечную неопределенность значений данных величин для любых τ [9]. Отражением этой неоднозначности результата «измерения» состояния V_τ является, введённое нами выше, понятие Т-матрицы плотности (3.2). Отсюда следует, что определение в уравнениях (4.6) и (4.7) с помощью операции $\{\ \}$ флуктуаций интенсивных переменных $\delta\rho$ и $\delta\sigma$ означает, по сути дела, использование диагональных элементов (3.3) Т-матрицы плотности $s(q_\tau, q_\tau)$. При этом значение s_L, сохраняющееся при Т-движении, является собственным значением указанной Т-матрицы. Введя, таким образом определённые, параметры порядка $\delta\rho$ и беспорядка $\delta\sigma$,

мы в состоянии рассматривать далее эти величины, как объекты подчиняющиеся законам классической механики. Именно в указанном смысле будет идти далее речь о Т-движении состояния, определённого с помощью (4.6) и (4.7) по гиперповерхности $\mathscr{P}$.

Формальным образом, понижение симметрии под действием на Т-гамильтониан системы преобразования τ можно связать с тем, что решение уравнения Т-движения (4.3) задаёт только одно, из симметричного в целом, набора возможных направлений Т-движения. Таким образом, симметрия упорядоченного состояния отвечающего этому решению и характеризующегося параметрами $\delta\rho$ и $\delta\sigma$, оказывается ниже симметрии Т-гамильтониана системы, то есть симметрии уравнения Т-движения (4.3), связанной с инвариантностью s_L. В модели ФРТ экстенсивной переменной V, нарушающей симметрию состояния, соответствует дополнительное условие (*) сохранения при Т-движении интенсивной переменной s_L. Как и выше, при определении V_c с помощью μ_L и T_L, мы не учитываем здесь, несущественные для данного этапа рассмотрения модели ФРТ, флуктуации этой величины в исходной подсистеме V_0. Наличие условия (*) согласуется с общими выводами работы [19], в которой РТ интерпретируется на основе симметрии.

Как установлено выше, переход из области $\tau \leqslant \tau_c$ в область $\tau > \tau_c$ сопровождается быстрым увеличением абсолютной величины значений $\delta\rho$ и $\delta\sigma$ и расплыванием графика, напоминающего δ-функцию до формы типа распределения Гаусса для ансамбля состояний V_τ, неустойчивых по отношению к физическому времени t. Поскольку при этом в соответствии с определением 2, должна, во всяком случае, не возрастать величина элемента площади Δ_j, изображающего локальное равновесное состояние, принадлежащее $\mathscr{P}$, следует предположить возможность смещения указанного элемента Δ_j для значений $\tau > \tau_c$, относительно распределений $\delta\rho$ и $\delta\sigma$ с центром в ρ_L и σ_L. Для увеличения абсолютных величин $\delta\rho$ и $\delta\sigma$ данное смещение должно осуществляться так, чтобы центр их распределния характеризовался непрерывно уменьшающимися с ростом τ параметрами ρ, σ при сохраняющемся отношении $s_L = \sigma/\rho$(*). Действительно, с физической точки зрения, непрерывное уменьшение масштаба измерения V_0 будет приводить к допустимым состояниям отдельных подсистем V_τ, средние характеристики которых ρ и σ должны отклоняться в сторону меньших значений от начальных ρ_L и σ_L. Последнее отвечает исчезающе малой вероятности скопления в любой из подсистем V_τ в течение достаточно длительного промежутка физического времени t некоторого числа частиц и, соответствующей этим частицам энтропии, превышающих средние значения данных величин по L-системе. Отсюда следует важный вывод:

Определение 4. В любой произвольный момент Т-времени из интервала $[\tau_c, \tau_G]$ Т-гамильтониан системы P является не возрастающей функцией τ.

В заключение отметим, что введённый определениями 1—4 и сопровождающим их обсуждением формализм модели Т-движения устанавливает вполне равноправное место ФРТ в ряду теорий, описывающих эволюцию системы и, таким образом, обладающих существенно динамической структурой.

Литература

1. Р. Балеску, Равновесная и неравновесная статистическая механика, Мир, Москва, 1978, т. 1, 408 с.; 1978, т. 2, 400 с.
2. Л. Д. Ландау и Е. М. Лифшиц, Теоретическая физика, Наука, Москва, 1973, т. 1, 208 с.; Физматгиз, Москва, 1963, т. 3, 704 с.; Наука, Москва, 1976, т. 5, 584 с.
3. А. З. Паташинский и В. Л. Покровский, Флуктуационная теория фазовых переходов. Наука, Москва, 1975, 256 с.
4. К. Вильсон и Д. Когут, Ренормализационная группа и ε-разложение. Мир, Москва, 1975, 256 с.
5. П. Гленсдорф и И. Пригожин, Термодинамическая теория стуктуры, устойчивости и флуктуаций. Мир, Москва, 1973, 280 с.
6. I. Gyarmati, Nonequilibrium Thermodynamics, Field Theory and Variational Principles, Springer Verlag, Berlin, New York, 1970; Dissertation, Budapest, 1957 (in English: 1958).
7. Е. Вигнер, Этюды о симметрии, Мир, Москва, 1971, 320 с.
8. Я. Г. Синай, Природа. Сер. Математика. 1981, 3, с. 72.
9. М. И. Рабинович, Усп. Физ. Наук, *125*, 123, 1978.
10. R. Mrugala, Rep. Math. Phys., *14*, 419, 1978.
11. L. Mistura, Nuovo Cim., *51B*, 125, 1979.
12. M. A. Peterson, Amer. J. Phys. *47*, 488, 1979.
13. Е. В. Преснов и С. Н. Малыгин, Термодинамика и кинетика биологических процессов. Химия, Москва, 1980, 197.
14. R. B. Griffiths and J. C. Wheeler, Phys. Rev., *2A*, 1047, 1970.
15. М. Фишер, Устойчивость и фазовые переходы. Мир, Москва, 1973, 376 с.
16. Г. Стенли, Фазовые переходы и критические явления, Мир, Москва, 1973, 424 с.
17. Н. Н. Боголюбов, Квазисредние в задачах статистической механики, ОИЯИ, Дубна, Д. 781, 1961.
18. А. А. Гриб, Е. В. Дамаскинский и В. М. Максимов, Усп. физ. наук, *102*, 587, 1970.
19. H. A. Callen, Symmetry Interpretation of Thermodynamics, Found. Continuum Thermodyn. L. B. 1974, V. 4. p. 61—79.

Acta Physica Hungarica 57 (1—2), pp. 31—40 (1985)

SPACE-TIME FOLIATIONS

G. SARDANASHVILY

Physics Faculty of Moscow University
117234, Moscow, USSR

(Received in revised form 23 June 1983)

The correspondence between gravitational fields and foliations describing a space-time structure of the world is established, and the criterion of gravitational singularities as foliation ones is suggested.

In General Relativity a space-time X can be characterized by four interdependent structures. They are the structure of an underlying manifold, the Riemannian, pseudo-Riemannian and space-time structures on X. A space-time is usually postulated to be a Haussdorf, connected, locally compact, paracompact, orientable, smooth four-manifold without boundary [1]. Such a manifold is metrizable, and it seems natural to provide X with the Riemannian metric g^R induced by a gravitational field g on X.

A space-time structure on X is represented as some $(3+1)$ decomposition of world vectors in time-like and space-like components relative to a given gravitational field g [2]. The $(3+1)$ decomposition is defined by a certain tetrad form $h^{\perp}$ of the field g, and if this form is integrable, a space-time is foliated in spatial hypersurfaces.

At present the method of $(3+1)$ decomposition of space-time-covariant quantities is widely applied to problems of the gravitation theory [3]. We are interested in space-time foliations because they represent a certain topological construction on a space-time [2], and their behaviour around a singularity may display the topological structure of the singularity. We aim at examining gravitational singularities as foliation singularities identified with zero points of tetrad forms $h^{\perp}$ as differential 1-forms on a space-time [4, 5]. The theory of such points is well elaborated [6].

1. Gravitational fields

A gravitational field on a manifold X is defined as a global section g of the fiber bundle B of pseudo-Euclidean bilinear forms in tangent spaces over X. The bundle B is associated with the tangent bundle $T(X)$ possessing the structure group $\mathrm{GL}^+(4, \mathbb{R})$, and it is isomorphic with the bundle in quotient spaces $\mathrm{GL}^+(4, \mathbb{R})/\mathrm{SO}(3, 1)$, whose global section h describes a gravitational field in the tetrad form. Thereby for a gravitational

field to exist on X the contraction of the structure group $GL^+(4, \mathbb{R})$ of $T(X)$ to the Lorentz group SO(3, 1), and consequently to SO(3) is necessary and sufficient [7]. It means that there exists such an atlas $\Psi^g = \{U_i, \psi_i^g\}$ of $T(X)$, whose transition functions gluing charts (U_i, Ψ_i^g) of trivializations of $T(X)$ reduce to elements of the gauge group SO(3, 1) (X), and the gravitational field g becomes the Minkowski metric η in all charts of Ψ^g. There is a one-to-one correspondence between various contractions of the structure group of $T(X)$ to the Lorentz group and gravitational fields on X.

The tetrad gravitational field h taken relative to Ψ^g assumes values in the center of the quotient space $GL^+(4, \mathbb{R})/SO(3, 1)$. The field h can be represented relative to an arbitrary atlas Ψ by the family of matrix functions $\{h_i(x), x \in U_i\}$ acting in the typical fiber of the tangent bundle and realizing the gauge transformation from Ψ^g to Ψ [7]. For instance, the well-known relation $g_i = h_i \eta$ (or in the index form $g_{ab} = h_a^A h_b^B \eta_{AB}$) holds. Note that the tetrad functions h_i are determined up to multiplying them on the right by gauge Lorentz operators, and this freedom reflects the nonuniqueness of an atlas Ψ^g.

2. The Riemannian structure and the space-time structure

There is the diagram

$$
\begin{array}{ccccc} & & SO(4) & & \\ & \nearrow & & \searrow & \\ GL^+(4, \mathbb{R}) & & & & SO(3) \\ & \searrow & & \nearrow & \\ & & SO(3,1) & & \end{array} \tag{1}
$$

of contractions of the structure groups of a tangent bundle over a manifold X admitting gravitational fields. Its subdiagram $GL^+(4, \mathbb{R}) \to SO(3, 1)$ defines a certain gravitational field g on X, as stated above. The contraction $SO(3, 1) \to SO(3)$ supplies X with a nonvanishing covector field (differential 1-form) ω as a global section of the bundle in quotient spaces $SO(3, 1)/SO(3)$ over X. The contraction $GL^+(4, \mathbb{R}) \to SO(4)$ defines a certain Riemannian metric g^R on X. Being combined in diagram (1), these contractions result in the following theorem.

Theorem 1.

(i) Let g be a gravitational field on a manifold X. There exists a nonvanishing differential 1-form ω and a Riemannian metric g^R on X such that

$$
g = g^R - 2\omega \otimes \omega / |\omega|^2, \tag{2}
$$

where $|\omega|^2 = g^R(\omega, \omega) = -g(\omega, \omega)$. Inversely, let ω be a nonvanishing 1-form on a manifold X. For any Riemannian metric g^R on X there exists a pseudo-Riemannian metric g on X such that the collection (g, g^R, ω) satisfies relation (2) [8].

(ii) For any collection (g, g^R, ω) obeying relation (2) there is an atlas Ψ^g such that $g, g^R, \omega/|\omega|$ form respectively the Minkowski metric η, the Euclidean metric η^E, and the form (1, 0, 0, 0) relative to Ψ^g.

(iii) The form $\omega/|\omega|$ coincides with the tetrad form $h^{\perp}=h^{0}_{\mu}dx^{\mu}$ of the gravitational field g, and forms ω and ω', connected with the same gravitational field g, differ in some Lorentz and dilatation gauge transformations. Thereby they fail to be antipodal at any point $x \in X$, and consequently they are homotopic [8]. The corresponding Riemannian metrics g^R and g may be nonequivalent (if there are points where the gauge Lorentz rotation between $h^{\perp}$ and h goes to infinity), but define the same metric topology on X.

The g-compatible pair $(g^R,\ \omega)$ from theorem 1 defines a certain $(3+1)$ decomposition of the tangent bundle $T(X)$ in the Whitney sum of the three-dimensional subbundle $T^{||}(X)$ evaluated from the equation $\omega=0$, and its orthocomplement $T^{\perp}(X)$ relative to the Riemannian metric g^R. The form $h^{\perp}$ has the dual (relative to g^R) space-like (relative to g) nonvanishing vector field τ on X, which represents the field of directions of the bundle $T^{\perp}(X)$. The metrics g and g^R coincide with each other on the subbundle $T^{||}(X)$ and induce a spatial Riemannian metric γ on it.

A gravitational field g on a space-time X admits many g-compatible $(3+1)$ decompositions describing various space-time structures on X. To choose a certain $(3+1)$ decomposition means to choose a certain reference frame (atlas Ψ^g) and to fix the Riemannian metric g^R on X such that observers connected with different reference frames perceive the same space-time as different Riemannian spaces. The well-known relativistic change of sizes of moving bodies exemplifies this phenomenon.

Thus we have defined the collection of quantities $(X, g, g^R, h^{\perp})$ characterizing the regular space-time.

3. Space-time foliations

The $(3+1)$ decomposition is named integrable if there is a hypersurface through each point of X such that its tangent space at each point is a fiber of $T^{||}(X)$. These spatial hypersurfaces form a g-compatible space-time foliation on X.

Let X be an n-dimensional connected smooth manifold without boundary. One says that there is a smooth foliation F of codimension $p<n$ on X, if X is represented as the union of disjoint sets F_α possessing the following property. For each point $x \in X$ there is a coordinate chart $(U \ni \varphi)$ such that φ maps connected components of intersections $F_\alpha \cap U$ onto $(n-p)$-dimensional parallel planes in $\mathbb{R}^n$. The sets F_α are named slices of the foliation F, and X is called the total space of F [9, 10]. Foliation slices are provided with the topology whose basis consists of connected components of intersections of slices with open subsets of X. Slices are $(n-p)$-dimensional manifolds relative to this topology, and they are submanifolds of X if this topology coincides with the induced topology.

The tangent bundle $T(X)$ over the total space X of a foliation F has the subbundle $T(F)$ of all tangent vectors to foliation slices. Foliations F and F' on a manifold X are called transversal if $T(X)=T(F)\oplus T(F')$. A foliation F is called orientable if the bundle $T(F')$, where F' is transversal to F, is orientable.

We shall deal with one-codimensional foliations.

Theorem 2

Any one-codimensional orientable foliation F on a manifold X can be defined by the equation $\omega=0$, where ω is a certain nonvanishing differential 1-form on X which satisfies the integrability condition

$$\omega \wedge d\omega=0 \qquad (d\omega=\sigma \wedge \omega). \tag{3}$$

Inversely, any such form generates a foliation on X. Generating forms ω and $\omega'=N\omega$, where N is an arbitrary nonvanishing real function on X, define the same foliation [9, 10].

We shall say that an orientable one-codimensional foliation F on a space-time X is a space-time foliation relative to a given gravitational field g if the generating form of F is the tetrad form $h^{\perp}$ of g.

The following theorem being the corollary of theorems 1 and 2 establishes the correspondence between space-time foliations and gravitational fields on a space-time X.

Theorem 3

Any one-codimensional foliation F on a space-time is a space-time foliation relative to a certain gravitational field g on X. Inversely, a gravitational field g whose tetrad form $h^{\perp}$ obeys the integrability condition (3) defines a space-time foliation generated by this form on X.

The integrability of the tetrad form $h^{\perp}$ can be expressed in terms of the gravitational connection Γ and reads $\Gamma^{\perp}_{[ab]}=0$, where $a, b=1, 2, 3$ index orthonormal basis of tangent and cotangent spaces of foliation slices.

4. The causality condition

The causality condition that the velocity of one's propagation cannot be greater than the velocity of light is the corner-stone of Special Relativity. However, in General Relativity this causality condition fails to be sufficient because closed time-like curves, curves trapped by a compact set, and other singular situations can exist. To describe these situations, various causality criteria are applied [1, 11]. The stable causality is the strongest causality condition which states that there is no causality violation of any type [1]. We formulate this causality condition in terms of foliations.

We shall say that a space-time foliation F is causal if no curve, transversal to F, intersects any slice of F more than once. It means that slices of a causal foliation are linearly ordered, and for verifying the causality any one curve intersecting a given slice is sufficient to be considered. Thereby slices of a causal space-time foliation represent the partial Cauchy hypersurfaces.

The following analog of a theorem by Hawking [12, 11] describes causal space-time foliations.

Theorem 4

A space-time foliation is causal only if it represents a foliation of level surfaces of a certain smooth real function f with nonvanishing differential df on X.

Note that the integrability of $(3+1)$ decomposition may be interpreted as the condition of sui generis local causality which claims the existence of a neighbourhood of any point $x \in X$ where the tetrad form $h^{\perp}$ of a given $(3+1)$ decomposition can be brought to the form $h^{\perp} = N df$, where N denotes a nonvanishing real function on X.

There exist gravitational fields not admitting integrable $(3+1)$ decompositions or causal space-time foliations. Such fields being regular themselves define singular space-time structures. At the same time by theorem 3 any one-codimensional foliation defines a regular gravitational field on a space-time. Thereby any gravitational singularity must be accompanied by a singularity of space-time foliations, that motivates our attempts to describe gravitational singularities via their foliation images.

5. Gravitational singularities

One faces gravitational singularities in the majority of physically significant solutions of Einstein's equations, but even the criteria of these singularities are unclear [13]. Singularities remain one of the principal problems of gravitation theory.

It seems natural to identify gravitational singularities with singular values of the metric g, or with the curvature R, or else with polynomial invariants of R and its derivatives. However, the regularity of all these quantities fails to guarantee against facing such singular situations as incomplete geodesics, causality breaking etc. At the same time, if nonscalar gravitational quantities are singular relative to a certain reference frame while scalar polynomials are regular, such singularities are usually considered as fictitious singularities being real for some observers but absent for others. However, there may be no holonomic reference frame where all components of metric and the Riemann curvature would be regular. Only the singular value of some scalar curvature polynomial indicates a gravitational singularity unmistakeably. However, one can exclude the point of this singularity from a space-time, and although the remainder is singular too, the criterion under discussion fails to indicate this singularity.

At present the criterion of gravitational singularities, which is based on the notion of so called "bundle-completeness" generalizing the familiar geodesic completeness, is most favoured [11]. By this criterion a gravitational singularity is present if any smooth curve in a space-time X cannot be prolonged up to any finite value of its generalized affine parameter. To describe this singularity, the b-boundary $\partial_b X$ is constructed and is attached to X, and then the behaviour of gravitational quantities measured in a tetrad propagated in parallel is analysed as one approaches $\partial_b X$. One divides singularities into the following main classes: (i) regular ("removable")

singularities; (ii) scalar and nonscalar curvature singularities; (iii) quasi-regular ("locally extendible") singularities [14, 15].

The b-criterion looks excellent from the position of a lone observer moving along a given curve in a space-time, but it is not devoid of defects either [16, 17].

Firstly, it is impossible to examine a continuum of curves. In reality, having found a singularity of gravitational quantities, one tests it by the b-criterion. Moreover, the definition of the b-boundary is not constructive, and the b-boundary can be built only in one or two cases.

Secondly, by the b-criterion singularity points must be excluded from a space-time X, and the remainder $\bar{X}$ is treated as the true space-time. However, such exclusion obscures the description of singularities because singular gravitational fields on the same manifold possess some common properties connected with the global structure of the manifold, e.g., Euler numbers of compact manifolds. But the unambiguous reconstruction of X from $\bar{X}$ is impossible in the general case. For example, sometimes $X \neq \bar{X} \cup \partial_b X$ if $\bar{X}$ equals X minus even regular points. The reconstruction is possible only when $\bar{X}$ inherits the Riemannian structure of X.

Thirdly, all constructions of the b-criterion, e.g., the generalized affine parameter, or the b-boundary depend only on the affine connection Γ and do not use anywhere the pseudo-Riemannian metric g or that Γ is the metric connection. Only that Γ is the SO(3, 1)-connection is used in [11] for constructing the b-complement $\tilde{\bar{X}}$ of $\bar{X}$, but this construction can be generalized to SL(4, $\mathbb{R}$)-connections too. The b-criterion ignores the pseudo-Riemannian structure of a space-time, and this causes that the b-criterion fails to indicate singularities when only the space-time causality is broken.

Thus the main defect of this criterion is that the behaviour of one or two curves tells only a little about the gravitational singularity structure, whose description needs an analysis of the behaviour of some family of curves or hypersurfaces near the singularity point.

We believe that space-time foliations are applicable to this problem.

6. Space-time singularities

Any breakdown of the regular structure of a space-time, which is characterized by the collection $(X, g^R, g, h^\perp)$ defined above, can be treated as a singularity.

We shall say that a gravitational field g on a space-time X possesses a singularity, if there is no g-compatible causal space-time foliation on X.

We distinguish three types of gravitational singularities.

The first type includes gravitational fields admitting regular $(3+1)$ decompositions (nonvanishing smooth forms $h^\perp$), but no causal space-time foliations. Such singularities destroying only the causality of a space-time need not possess singular values of gravitational quantities.

The violation of causality (i.e. the deviation of a space-time foliation from a level surface foliation) can be characterized by the nonzero values of the cohomology class of ω (if $d\omega=0$), the Reinhart and Godbillon—Vey classes A, B, GV of a space-time foliation [2].

For instance, let X be a cylinder $S^1\times\mathbb{R}^1$ possessing the pseudo-Riemannian metric $ds^2=-d\psi^2+dr^r$, where ψ is the S^1-coordinate. This metric admits a space-time foliation generated by the closed, but not exact form $h^\perp=d\psi$. Other classes A, B, GV are zero. The causality is broken because of closed time-like curves $r=\text{const}$. Note that the b-criterion fails to indicate such singularities.

The Taub—Misner metric $ds^2=2d\psi dr+rd\psi^2$ on $X=S^1\times\mathbb{R}$ exemplifies the case when a space-time foliation possesses only the non-zero Reinhart class A (i.e. $\sigma\neq0$) measuring the rotation of the vector field τ. For example, time lines of the space-time foliation generated by the form $h^\perp=dr+1/2(r-1)d\psi$ have the limiting circle $r=-1$. These lines are complete relative to the geodesic parameter $(r-\psi)$. But the Taub—Misner world is known to possess incomplete null geodesics.

In general singularities of the first type represent quasi-regular singularities and nonscalar curvature singularities (e.g., in the Taub-NUT model) by the b-criterion.

The second type of gravitational singularities includes gravitational fields admitting singular $(3+1)$ decompositions and foliations.

Singular foliations are defined as closing the class of foliations under the operation of the foliation induction.

A smooth map $f\colon Y\to X$ of a manifold Y into the foliation total space X is called transversal to F, if $T_x(X)=T_x(F)\oplus\mathrm{Im}\,(df)_x$ at each point $x\in X$. If the map f is transversal to F, the preimages of slices of F compose the induced foliation f^*F on a manifold Y, and codim $f^*F=$ codim F. When the smooth map $f\colon Y\to X$ fails to be transversal to the foliation F on X, the induced construction f^*F makes a certain geometric sense too, this may be interpreted as a singular foliation [9].

The following theorem describes singularities in causal space-time foliations.

Theorem 5

Singular foliations closing the class of causal space-time foliations represent foliations of level surfaces of real smooth functions f on a space-time. Singularities in such foliations are identified with the critical points of functions f, i.e. where $df=0$.

However, in the general case a 1-form ω vanishing at some points of X fails to define any singular foliation. But such a form may be considered as defining some singular $(3+1)$ decomposition on X. Thus, the second type of gravitational singularities includes gravitational fields admitting tetrad forms which being multiplied by smooth functions reduce to smooth 1-forms possessing zero points on X. The following theorem holds.

Theorem 6

If a gravitational field g admits a singular $(3+1)$ decomposition generated by the form ω, any other g-compatible $(3+1)$ decomposition generated by the form ω' possesses the singularities at the same points where $\omega=0$. If such a singularity is isolated, it can be characterized by the indices of the vector fields τ and τ' at the singular point, and these indices coincide with each other.

However, there are cases when a form ω being equal to zero at a point $x \in X$ can define a regular $(3+1)$ decomposition or a foliation on X, e.g., when $\omega = N\omega'$ where ω' is a nonvanishing form, and N equals zero at x. To discern such a fictitious singularity one can use the additional criterion that the divergence of the vector field τ tends to infinity about a true singularity. For instance, it takes place when an isolated zero of a field τ is characterized by the non-trivial index.

Let a form ω define a singular foliation F on a space-time X. Then by well-known theorems [18] div τ is related to the exterior curvature of slices of F as follows:

$$K = -\gamma^{ab}\Gamma^{\perp}_{ab} = -\operatorname{div}\tau - 1/2\partial_{\perp}\ln|g|. \tag{4}$$

This shows that the scalar exterior curvature of foliation slices tends to infinity near a foliation singularity. Then K becoming infinite can indicate the existence of space-time singularities. Moreover, this singularity criterion is rather handy because K calculated by formula (4) is independent of the choice of a reference frame.

Singularities in homogeneous cosmological models exemplify singularities of the second type. A space-time X of such a model is foliated into three-dimensional space-like orbits of the group of motion of this model, and these orbits represent the level surfaces of the morphism f of X onto the time axis T. Then the singularities of this model can be described and classified as the critical points of the function f [19]. For instance, the singularity in the closed Friedman model can be described as the nondegenerate critical point (whose index equals zero) of a real function on S^4.

The second type singularities represent the scalar and nonscalar curvature singularities, and the quasi-regular singularities (e.g., the cone singularity) by the b-criterion.

Gravitational fields of the types considered above are assumed to admit regular Riemannian metrics, compatible with them, and therefore singularities of these types can be described as caused by singularities only of the space-time structure of world, i.e. as foliation singularities.

The third type of gravitational singularities includes gravitational fields g not admitting regular g-compatible Riemannian structure on a space-time X.

The topology and the standard manifold structure of X fails to be compatible with g^h and consequently g at points where gravitational singularities are accompanied by singularities of a Riemannian metric. Therefore it seems reasonable to exclude such points from a space-time. Then the following cases are possible.

(i) The remainder represents a regular space-time.

(ii) Singularities reduce to singularities of the first and second types.

(iii) The remainder fails to possess the conventional attributes of the standard manifold structure of a space-time, which hinders the foliation analysis of the singularities. For instance, the tetrad form $h^{\perp}$ generating a space-time foliation can be regular and exact whereas a singularity is present. At the same time the $K=\infty$ criterion can indicate such singularities.

The $r=0$ singularity in the Schwarzschild solution exemplifies the third type singularity combined with the cone singularity.

7. The singularity theorem

Let us consider gravitational fields obeying the Einstein equations.

The theorems by Hawking and Penrose establish that solutions of the Einstein equations possess singularities in the sense of the b-criterion if the matter energy-momentum tensor T satisfies the strong energy condition (i.e. $T_{ab}v^a v^b \leq 1/2T$ for any unit time-like vector v), a certain causality condition and some other requirements hold [11].

The proof of these theorems is based on the solution of the Raychaudhuri equation describing the geodesic deviation. This solution shows the existence of the conjugate points on a finite segment of the geodesic curve γ if there holds the condition $R_{ab}v^a v^b \geq 0$, where v is a tangent time-like vector to γ. Just this condition being substituted in the Einstein equations results in the strong energy condition. The causality condition is necessary to obtain a solution of the deviation equation on a finite interval of the geodesic curve. And the theorems themselves relate the existence of the conjugate point to the existence of a gravitational singularity in the sense of the b-criterion under some supplementary requirements.

This translation is unnecessary if one uses the singularity criterion formulated above. Indeed, let γ be integral curves of the vector field τ, dual to the generating form $h^{\perp}$ of some causal space-time foliation F. Then γ are the geodesic curves, the tensor of the volume deviation Θ_{ab} coincides with the exterior curvature tensor K_{ab} of foliation slices, and the Raychaudhuri equation represents the equation for the scalar exterior curvature:

$$\frac{d}{ds}K = -R_{\perp\perp} - 2\sigma^2 + 1/3K^2, \tag{5}$$

where $\sigma_{ab} = K_{ab} - 1/3\gamma_{ab}K$. This equation can be obtained directly by pairing the spatial indices d, a in the first from the following expressions for the F-compatible $(3+1)$ decomposition of the Riemann tensor [3]:

$$\begin{aligned} R^d_{\perp a \perp} &= \partial_\perp K^d_a - K^d_b K^b_d, \\ R^d_{cab} &= {}^3R^d_{cab} + K^d_a K_{bc} - K^d_b K_{ac}, \\ R^\perp_{cab} &= \nabla_b K_{ac} - \nabla_a K_{bc}. \end{aligned} \tag{6}$$

Here, because F is the causal foliation, we choose the holonomic spatial indices and the affine parameter s as the coordinate along integral curves of the field τ.

Note that the second and third equations (6) are respectively the well-known Gauss and Codazzi equations.

Solving equation (5) under the strong energy condition, we obtain that the scalar exterior curvature K becomes infinite at the finite value of the affine parameter in future from the point $x \in \gamma$, if $K(x) < 0$, or in past, if $K(x) > 0$.

By theorem 6 this singularity is the attribute of all $(3+1)$ decompositions, compatible with the given gravitational field, and consequently this is the attribute of the gravitational field itself. Thus the following theorem holds.

Theorem 7

The solution of the Einstein equations with the matter energy-momentum tensor obeying the strong energy condition does not admit any regular causal space-time foliation, i.e. possesses the gravitational singularity by our criterion.

This theorem is analogous to the singularity theorems by Hawking and Penrose.

We have discussed above the general aspects of the description of gravitational singularities as foliation ones. One of the promising applications of such a description is the classification of gravitational singularities in the types of critical points of the foliation generating 1-forms [19]. These are the topological characteristics of gravitational singularities.

References

1. R. Geroch, GRG, *2*, 61, 1971.
2. G. Whiston, GRG, *5*, 517, 525, 1974.
3. J. Isenberg and J. Nester, In "General Relativity and Gravitation", Vol. 1, Plenum Press, N.-Y., London, 1980, p. 23.
4. G. Sardanashvily and V. Yanchevsky, Izv. vuzov SSSR, Physika, N9, 20, 1982 (in Russian).
5. D. Ivanenko and G. Sardanashvily, Phys. Lett., *91A*, 341, 1982.
6. V. Arnold, A. Varchenko and S. Husein-Zade, Singularities of differential mappings, Nauka, Moscow, 1982 (in Russian).
7. D. Ivanenko and G. Sardanashvily, Phys. Rep., *94*, 1, 1983.
8. R. Von Sulanke, P. Wintgen, Differentialgeometric und Faserbundle, Deutscher Verlag d. Wissenschaften, Berlin, 1972.
9. D. Fuks, In "Modern problems of mathematics", Vol. 10, VINITY, 1978, p. 179 (in Russian).
10. H. Lawson, Bull. Am. Math. Soc., *80*, 369, 1974.
11. S. Hawking and G. Ellis, The large scale structure of a space-time, Cambridge Univ. Press., Cambridge, 1973.
12. S. Hawking, Proc. Roy. Soc., *308A*, 433, 1969.
13. R. Geroch, In "Relativity; Proc. Relativity Conf., Cincinnati, June 2—6. 1969", N.-Y., London, 1970.
14. G. Ellis and B. Shmidt, GRG, *10*, 989, 1979.
15. C. J. S. Clarke, Comm. Math. Phys., *41*, 65, 1975.
16. B. Bossard, GRG, *10*, 963, 1979.
17. R. Johnson, GRG, *10*, 967, 1979.
18. S. Kobayashi and K. Nomizu, Foundation of differential geometry, Int. Publ. J. Wiley and Sons, N.-Y., London, 1963, 1969.
19. G. Sardanashvily and V. Yanchevsky, Vestn. Mosk. Univ., Physika, *23*, 71, 1982 (in Russian).

Acta Physica Hungarica 57 (1—2), pp. 41—54 (1985)

INTERFACE STATE DENSITY AND CAPTURE CROSS SECTION MEASUREMENTS ON MOS STRUCTURES BY TSC

D. MARTON and A. HÜBLER*

Physics Institute of the Technical University of Budapest
1521 Budapest, Hungary

(Received in revised form 9 July 1983)

The Thermal Dielectric Relaxation Current method, a version of TSC, has been applied to MOS capacitor probes. Capture cross sections and densities of electronic states have been determined as functions of energy for samples obtained by different technological procedures. The results are in good qualitative agreement with the few available literature data. The decrease of the capture cross sections near the band edges is interpreted as a consequence of the increasing density of states. This relationship may give rise to the well-known systematic differences between the interface state densities obtained from CV (capacitance measurements) and thermal methods.

Introduction

In the past decade several methods have become commonly used for the measurement of interface state densities in MIS structures. First of all, there are the AC capacitance and conductivity measurements and that of quasistatic capacitance by means of which the interface state distributions in the middle of the gap can be determined. On the basis of the conductivity measurements some capture cross section data have been published, too [1—4]. Similar data have also been published as obtained from the photoionization of interface states and from the DLTS method as well [5].

In the early seventies Simmons and co-workers [6—8] suggested the use of non-equilibrium methods to investigate the interface states by measuring thermally generated and isothermal transient currents (TSC methods). The theory of the TSC method was elaborated for MOS and its use was demonstrated with experiments on MNOS structures [9]. The method seems to be very useful to investigate MOS structures, since in principle

* Present address: College for Electrical Engineering K. Kandó, Institute for Computer Technique, Székesfehérvár, Vörös Hadsereg út 45. 8002 Hungary

1. it is relatively simple by comparison with the methods listed above,
2. it provides more information: one pair of measurements carried out on a single sample is enough to obtain the $D(E)$ energy spectrum of interface states in both halves of the band gap, and the capture cross section σ_n (or σ_p on p-type samples) can also be obtained,
3. the $D(E)$ spectrum can be obtained as close to the band edges as desired which is not so when the CV and the conductance methods [10] are used.

Despite all these advantages the TSC method has been applied only to a very limited extent in MOS investigations, and even when it was used, it was generally not in its original form [11]. This situation is due to sensitivity problems and some severe difficulties in the evaluation of the results.

In the present paper the results of interface state density and capture cross section measurements obtained on MOS structures by the TSC method are reported together with some considerations on the accuracy and evaluation.

Principles of the TSC method

The electronic states of an n-type semiconductor–insulator interface in a positively biased MOS capacitor (the metal is $+$) will be filled with electrons. After cooling the capacitor in this state and changing the polarity of the voltage the surface of the semiconductor will be in deep depletion for a long time. Meanwhile, the filled surface states lying close to the conduction band start to emit electrons [7]. The Shockley–Read theory can be applied to describe this process. The velocity of electron *emission* is

$$\frac{d\delta Q_t}{dt} = -e_n \delta Q_t \,, \tag{1}$$

where $\delta Q_t = Q_t \delta E$ is the charge in surface states in the energy interval δE, and e_n is the emission probability. The time dependence of the current and energy of the emptying levels is obtained by integration of this differential equation. The result of this integration will be different for different possible conditions: isothermal or changing temperatures. The methods are called correspondingly IDRC and TDRC, the latter is carried out with a temperature ramp,

$$T = T_0 + \beta t \,, \tag{2}$$

where T_0 is the initial value of the temperature, t is the time and β is constant.

The integration leads in the IDRC case to

$$I = \frac{qA_g kT}{t} N_D(E) \tag{3}$$

and

$$E=E_c-kt\ln\nu t, \tag{4}$$

where q is the elemental charge, A_g is the surface area of the electrode, k is the Boltzmann constant, $N_D(E)$ is the energy distribution of electron states, E_c is the conduction band edge and

$$\nu=v_{\mathrm{th}}\,\sigma_n N_c, \tag{5}$$

where v_{th} is the thermal velocity of charge carriers (electrons) in the conduction band, σ_n is the capture cross section for electrons and N_c is the effective density of states in the conduction band.

The deeper energy levels will not be emptied by the emission process described above. The electrons captured in these levels can take part in a *generation* process which can again be described according to the Shockley—Read theory [8]. The velocity of this process is

$$\frac{d\delta Q_t}{dt}=\frac{e_n e_p}{\bar{p}}\,\delta Q_t, \tag{6}$$

where e_p is the hole emission probability and $\bar{p}$ is the effective concentration of holes. The process will be effective when the ΔE interval is between the quasi Fermi levels of holes and electrons. Expressions similar to those cited above can be obtained for the currents and energies in this case, too. The TDRC process can be described by equations deduced on a similar basis.

Evaluation problems and accuracy

The obvious advantages of the TSC methods are accompanied by two limitations. First, the accuracy in determining the energy scale of the interface state distribution can be improved by decreasing the slope of the temperature ramp β, but at the same time the sensitivity (which is relatively poor anyway) decreases. This means in practical cases of $\beta=(0.1-0.5)\,\mathrm{Ks}^{-1}$ that $(0.2-1)\cdot 10^{-13}$ A transient current is to be detected to achieve the usual sensitivity of CV methods $D=10^{10}\,\mathrm{eV}^{-1}\,\mathrm{cm}^{-2}$. The authors of the MOS—TSC theory avoided this problem by measuring only on MNOS structures having 2—3 orders of magnitude higher interface state densities [9]. On the other hand, by the TSC method the capture cross section can be determined by measuring two TDRC curves with two different ramps, viz, β_1 and β_2.

Then the expression

$$\lg\nu=\frac{T_2\lg\beta_2-T_1\lg\beta_1}{T_2-T_1}+\mathrm{const}. \tag{7}$$

is used where the constant is determined theoretically, temperatures T_1 and T_2 are corresponding ones in the two measurements (e.g. the temperatures of the current peaks obtained).

It can be seen that a relatively small error in the measurement of T can lead to a considerable error in the determination of ν and σ (compare Eq. (6)) since T_1 and T_2 are close in value.

Among the few attempts that have been made to use the TDRC method for MOS structures a characteristic example is Ning's work [12], whose results show that σ_n could have been determined only with an uncertainty of ± 4 orders of magnitude, and the error of the energy scale of the interface state distribution was ± 0.2 eV.

Saunders and Wright [13] suggested the use of this method for MOS studies as early as 1970, and the main problem they met was to determine the attempt-to-escape-frequency. (In the evaluation of the measured results Saunders and Wright used $\nu = 10^{10}\,\mathrm{s}^{-1}$ as a reasonable assumption.) Wei and Simmons published the value $\nu = 2.5 \cdot 10^{12}\,\mathrm{s}^{-1}$ obtained from a typical pair of measurements, but no data were reported regarding the accuracy of this value [9].

Our preliminary experiments showed that the temperature of the generation current peak shifted for an n-type sample by about 5 K by changing the ramp in the range $\beta = (0.1 - 0.5)\,\mathrm{Ks}^{-1}$ mentioned above. Thus the accuracy of the temperature measurement at the value where the TDRC peak appears has to be below ± 0.5 K, if ν is to be determined within ± 2 orders of magnitude. This accuracy can be achieved only by careful and precise experiments.

First of all we have to point out that if this accuracy, i.e.

$$\delta(\lg \nu) = \pm 2 \tag{8}$$

is achieved, the accuracy of the energy determination will be satisfactory.

Simmons and Taylor [6] have shown that for the emission process in the TDRC measurement the equation

$$\exp\left(\frac{\Delta E}{kT}\right) = \nu k T \frac{1}{\beta \Delta E} \tag{9}$$

is true, where $\Delta E = E_c - E$. Linearization of Eq. (9) in the form

$$\Delta E = T\left(A \lg \frac{\nu}{\beta} + B\right) + C \tag{10}$$

led to the numerical equation

$$E_c - E = 10^{-4} T\left(1.98 \lg \frac{\nu}{\beta} + 3.2\right) - 0.0155\,, \tag{11}$$

where ν, β and T are given in their usual units (s^{-1} and K, respectively) and the energy is obtained in eV-s. These numerical calculations have been repeated in the present study and the $\lg \frac{\nu}{\beta}$ versus $\frac{\Delta E}{T}$ functions were numerically obtained with the parameter T. These results are shown in Fig. 1. It can be seen that the linearity is obtained in spite of neglecting the temperature dependence of B in Eq. (10). This dependence has been

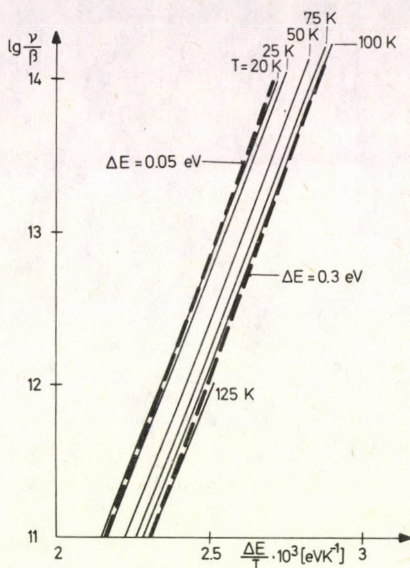


Fig. 1. The $\lg \frac{\nu}{\beta}$ versus $\frac{\Delta E}{T}$ relation for the emission process in the TDRC method. The parameter is the temperature (K). Energy (ΔE) is given in eV-s. The diagram covers the full realistic range of ν, when $0.1\ \mathrm{Ks}^{-1} < \beta < 1\ \mathrm{Ks}^{-1}$. The dotted lines correspond to constant energies

calculated and is shown in Fig. 2 compared with the constants calculated by Simmons et al. The difference is too small to be detected experimentally except perhaps in the range of very low temperatures.

The dotted lines in Fig. 1 correspond to the limits of energies

$$0.05\ \mathrm{eV} < \Delta E < 0.3\ \mathrm{eV}, \tag{12}$$

the lower limit being set by experimental difficulties and the upper one by the access of CV methods. These energy limits, as can be seen from the Figure, correspond to the following limits in temperature

$$20\ \mathrm{K} < T < 120\ \mathrm{K}\,. \tag{13}$$

These limits can be slightly modified if the $\frac{\nu}{\beta}$ values are higher or lower than those considered in the Figure. The range of $\frac{\nu}{\beta}$ was determined for this Figure by taking into account the known experimental data for σ_n [3, 14] and the usual values of β. This range covers 4 orders of magnitude. One can see in Fig. 1 that

$$2.3 \cdot 10^{-3} \frac{\mathrm{eV}}{\mathrm{K}} < \frac{\Delta E}{T} < 2.9 \cdot 10^{-3} \frac{\mathrm{eV}}{\mathrm{K}} \tag{14}$$

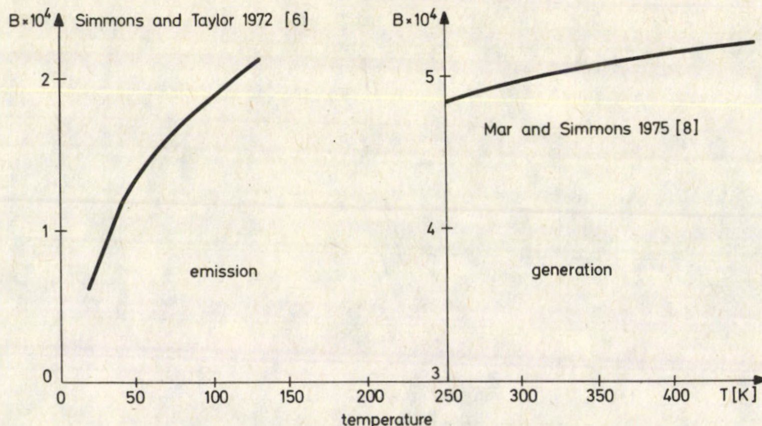


Fig. 2. The *B*-factor in Eq. (3) for both emission and generation processes. The solid line represents *B* as calculated in this work, the dotted lines correspond to the *B* values given by Simmons et al

or

$$\frac{\Delta E}{T}=(2.6\pm 0.3)\cdot 10^{-3}\,\frac{\mathrm{eV}}{\mathrm{K}}. \tag{15}$$

So the determination of the T_m value (the temperature where the current peak appears due to interface states situated at energy ΔE in the silicon band gap) leads to a value of

$$\Delta E=2.6\cdot 10^{-3}\, T_m\,(\mathrm{eV}) \tag{16}$$

with an accuracy

$$\delta(\Delta E)=0.3\cdot 10^{-3}\, T_m\,(\mathrm{eV}). \tag{17}$$

In the worst case, if $T_m=120$ K, this corresponds to

$$\delta(\Delta E)=0.036\ \mathrm{eV}. \tag{18}$$

Similar calculations were made by Mar and Simmons for the generation process as well [8]; we would only mention here that the constant C in an expression similar to (10) was determined erroneously. The temperature dependence of B can be omitted (Fig. 2) and then we obtain the (numerical) equation

$$E_c-E=10^{-4}\,T\left(1.98\,\lg\frac{\nu}{\beta}+4.6\right)-0.115. \tag{19}$$

The results are shown in Fig. 3, where the energy range

$$0.8\ \mathrm{eV}<E_c-E<1.1\ \mathrm{eV} \tag{20}$$

is shown by dotted lines. It can be seen that for a similar range of $\frac{\nu}{\beta}$ as in the emission

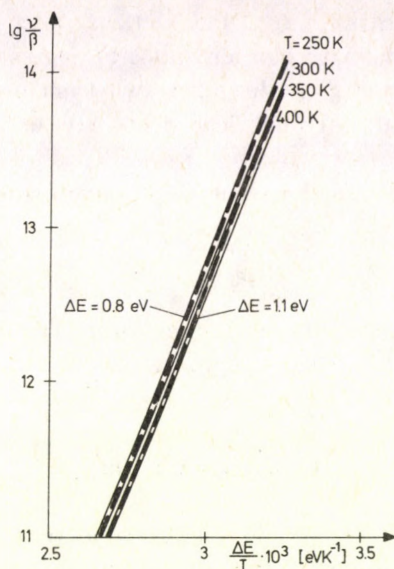


Fig. 3. The $\lg \frac{\nu}{\beta}$ versus $\frac{\Delta E}{T}$ relation in the case of the generation process in the TDRC method. The parameter is the temperature (K). Energy ΔE is given in eV-s. The diagram covers the full realistic range of ν, when $0.1\ \mathrm{Ks}^{-1} < \beta < 1\ \mathrm{Ks}^{-1}$. The dotted lines correspond to constant energies

process, the temperature limits are

$$250\ \mathrm{K} < T < 400\ \mathrm{K}\,, \tag{21}$$

and

$$\frac{\Delta E}{T} = (3.0 \pm 0.28) \cdot 10^{-3}\,\frac{\mathrm{eV}}{\mathrm{K}}\,. \tag{22}$$

In the worst case (at $350\ \mathrm{K} - 400\ \mathrm{K}$) we have

$$\delta(\Delta E) = 0.1\ \mathrm{eV}\,. \tag{23}$$

The temperature limits obtained mean that the generation process of MOS structures should be studied at temperatures making possible the motion of some ions.*

Our conclusion can be summarized in the following statement: the measurement of ν with an error of ± 2 orders of magnitude makes the determination of the energies of interface states possible with a satisfactory accuracy: the maximum error is ± 0.035 eV for the emission and ± 0.1 eV for the generation process.

* This may be a problem for the so-called TSIC measurements. Boudry and Stagg [20] observed a current peak at about 350 K and we think that a generation phenomenon could not be excluded in that measurement. The evaluation of those results [20] as a TSC peak provides a surface state level of $D = 5 \cdot 10^{11}\ \mathrm{cm}^{-2}$ density at the energy $\Delta E = (0.96 \pm 0.09)$ eV.

The $D(E)$ energy distribution of the interface states can be determined very simply on the basis of the above considerations by measuring a single $I-T$ curve, provided that typical lg v_n and lg v_p values are known from literature and assuming that the capture cross sections are not much dependent on the technology and the energy of the states.

Several experimental σ_n and σ_p values from the literature are shown in Fig. 4. It can be seen that

$$\delta(\lg v) \cong \pm 3 \tag{24}$$

which may cause too great errors in the determination of the energy. Thus we return to the evaluation procedure suggested by Simmons et al [6—9] with the results obtained above in view.

Experimental results

Experiments were carried out on four different types of samples: A, B and D groups were n-type silicon wafers of 2 Ohm cm resistivity and $\langle 111 \rangle$ orientation, the C group was p-type silicon of the same orientation and 10 Ohm cm resistivity.

The thermal oxides on the surface of the wafers in group A were 100 nm thick, grown at $T_{0x} = 1150$ °C in dry oxygen. Al dots of $A = 1.15 \cdot 10^{-2}$ cm^2 area produced by electron-beam evaporation were used as metal electrodes.

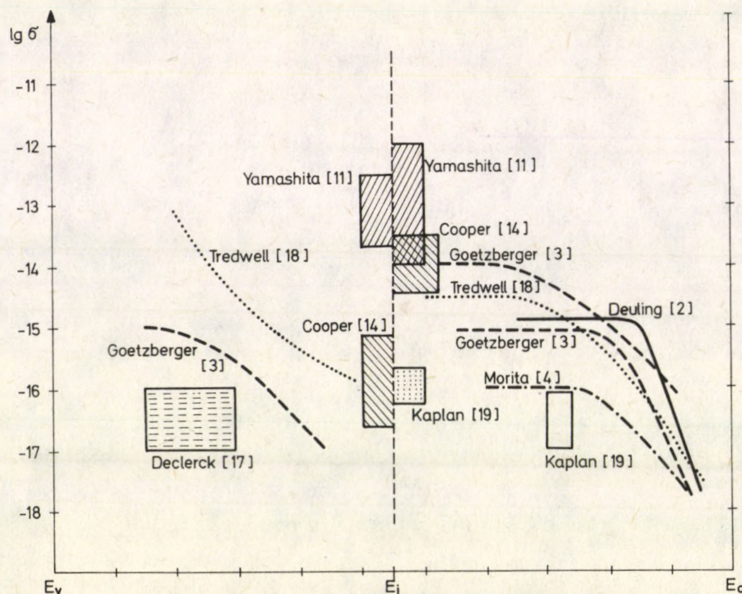


Fig. 4. Capture cross section data obtained by different authors. In the lower part of the gap ($E < E_i$) the data are for holes σ_p, in the upper one for electrons σ_n. The data were obtained in conductivity measurements [14, 2, 3, 19, 4, 17], and in DLTS [5] or TSSP [11] measurements

The samples in the B and C groups were oxidized in an atmosphere containing HCl (4 vol. % HCl) at $T_{0x} = 1100$ °C. The thickness of the oxide was 110 nm. The gate electrode was produced in essentially the same way as for the wafers of group A.

Oxide layers on samples of group D were grown together with those of group A. Instead of the metallic electrode, in this case poly — Si of $1.6 \cdot 10^{-3}$ cm^2 area was used.

Experiments were carried out by means of a complex measuring system described in a previous work [18]. The only modification was that the cryostat was smaller and the pressure of the liquid nitrogen was allowed to decrease to $1.25 \cdot 10^4$ Pa so that the temperature 63 K as the lowest temperature available could be reached. The temperature ramp for the TDRC measurement could be started at 70 K. Pt resistors (Degussa) were used for temperature measurement and control at two different points of the bulk sample holder made of copper. The accuracy of the temperature measurement was $\delta T < 0.1$ K.

The measuring system and the cryostat allowed TDRC, IDRC, quasi-static and high-frequency CV measurements to be carried out on the same sample under essentially the same conditions.

The high frequency CV measurements were made at 1 MHz with 25 mV_{pp} measuring voltage. The slow ramp meant $\alpha = 1$ mV s^{-1}. . .0.1 Vs^{-1} constant voltage ramps. The temperature programme had a linearity better than 1% in the range $\beta = (0.1 - 0.5)$ Ks^{-1} and $T = 70 - 300$ K. The currents were measured with a Keithley 610 C electrometer.

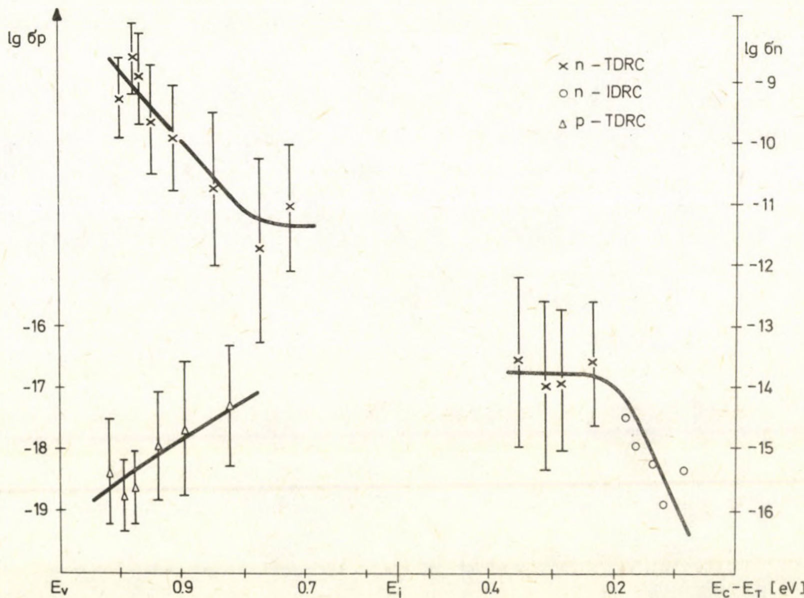


Fig. 5. σ_n versus $(E_c - E_t)$ and σ_p versus $(E_c - E_t)$ curves obtained in TSC measurements. The mean value and the standard deviation of 12 TDRC measurements are indicated in the Figure. The IDRC results are the mean values of two measurement series

The σ data summarized in Fig. 4 suggest that the capture cross section and thus ν depend on the energy of the level in the gap. The authors returned, therefore, to the evaluation procedure as described by Simmons et al [6—9].

Numerous $I-T$ curves (7—12) have been measured under similar circumstances, but with different β values. Using correspondent points ν values have been determined in a wide interval of the gap. σ_n and σ_p have been calculated using (7) and the following equations:

$$v_{\mathrm{th}}=\sqrt{\frac{3kT}{m}}, \tag{25}$$

$$N_c=5.4\cdot 10^{15}\cdot T^{3/2}, \tag{26}$$

$$N_v=2.43\cdot 10^{15}\cdot T^{3/2}. \tag{27}$$

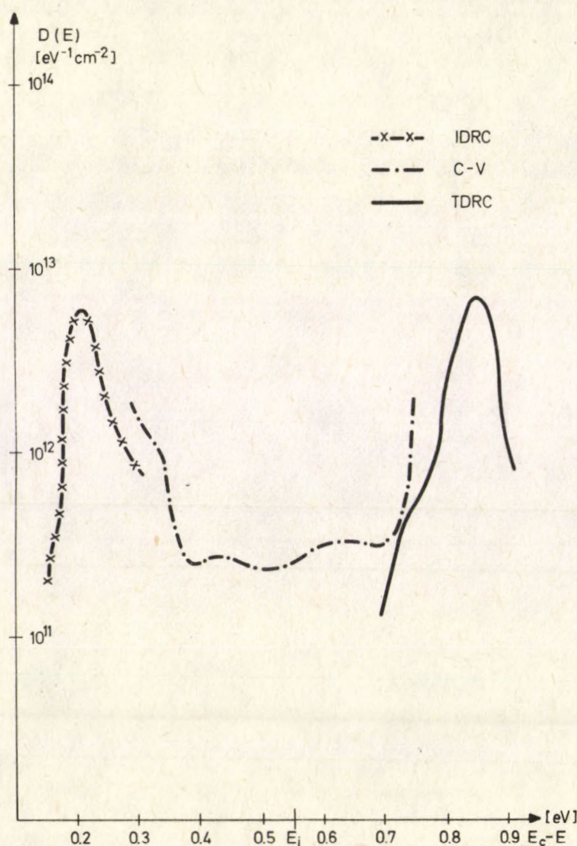


Fig. 6. Energy distribution of interface state density of an MOS capacitor based on a ⟨111⟩ oriented n-type Si wafer oxidized by dry oxygen at 1150 °C and furnished with Al gate (Sample type A). Different parts of the curve were obtained using different measurement methods

The results obtained on samples B and C are shown in Fig. 5 as a function of the energy $E_c - E_T$.

As we have seen above the achieved minimum temperature (70 K in TDRC) does not guarantee the measurement of the emission peak if the states lie close to the band edge. Thus for the determination of σ_n the IDRC method elaborated by Simmons and Wei [15] has been applied to complete the data.

The other two types of samples revealed less than $\delta(\lg \nu) = \pm 2$ deviation from the $\sigma - E$ function described here; in view of this, corresponding to the former conclusions, the $D(E)$ evaluation has been carried out using the σ values shown in Fig. 4.

The interface state distribution $D(E)$ of the $\langle 111 \rangle$ n−Si sample with the dry-grown 100 nm oxide (sample A) is shown in Fig. 6. In the energy range near the conduction band no considerable increase of the interface state density was found (the energy of the highest level that could be detected with the limited cooling possiblity of

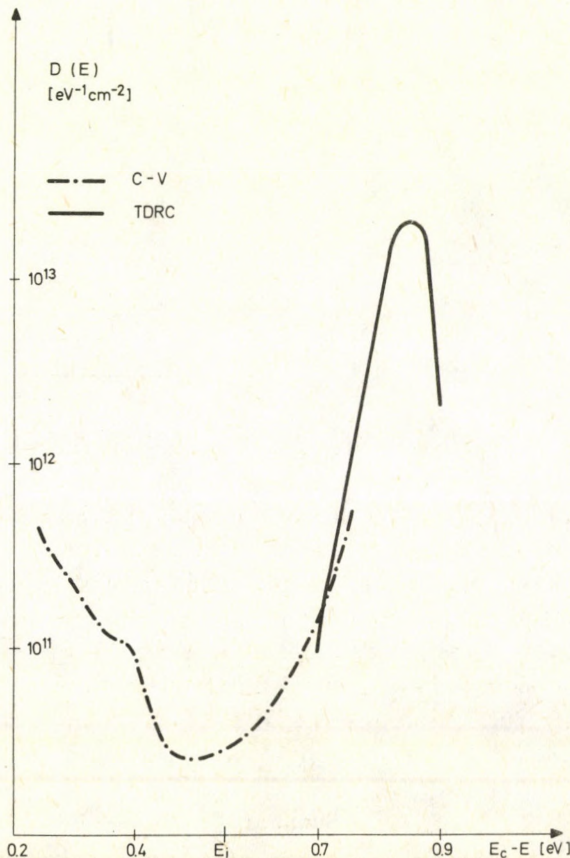


Fig. 7. Energy distributions of interface state densities of different MOS capacitors, as measured by CV and TDRC methods (different parts of the curves)
The evaluation of the TDRC curves in this case is based on results shown in Fig. 4
a) $\langle 111 \rangle$ oriented n-type Si, HCl grown oxide, Al gate (Sample B)

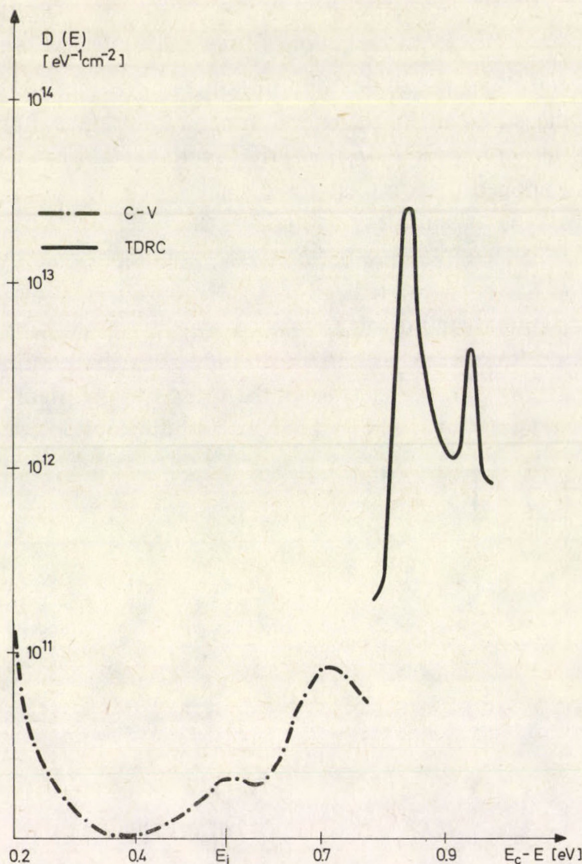


Fig. 7. b) ⟨111⟩ oriented p-type Si, HCl grown oxide, Al gate (Sample C)

our equipment was $E \cong E_c - 0.2$ eV), thus for the emission peak measurement we used the IDRC method. The density of the interface states close to the midgap was determined by means of the well-known combination of high frequency CV with the ramp CV method. The results of the latter were used exclusively for the determination of the energy of states in the middle range of the gap [16].

Similar results obtained by means of TDRC and CV measurements are shown in Fig. 7 for the other types of samples (B—D).

Summary

The TDRC method as suggested by Simmons and co-workers [6—9] provides data for the capture cross sections of interface states of MOS capacitors that can be quite inaccurate: the error can be several orders of magnitude, judging by analysis of

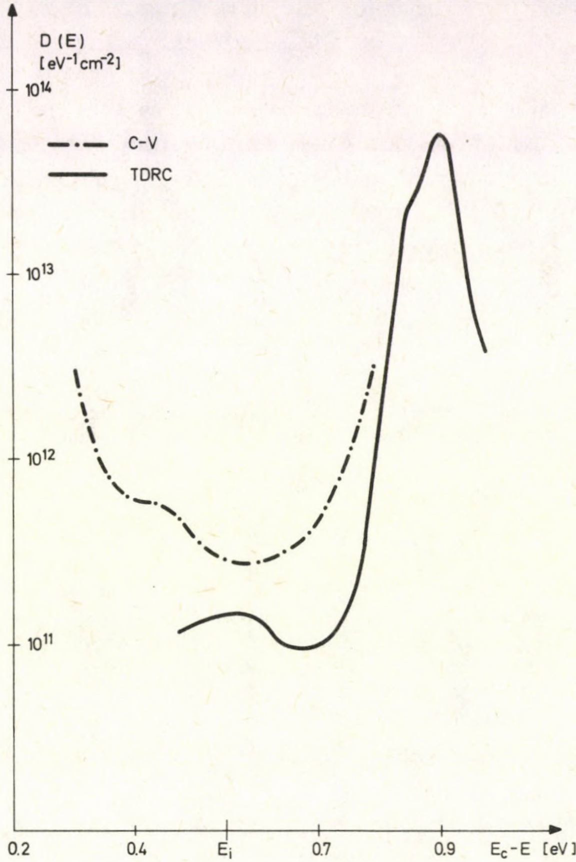


Fig. 7. c) ⟨111⟩ oriented n-type Si, oxidized in dry oxygen, poly-Si gate (Sample D)

the evaluation procedures. This inaccuracy does not, however, exclude quite accurate $D(E)$ data.

In spite of this problem the authors were able to obtain relatively correct capture cross section data as a function of the energy of the state (σ vs E). In the emission range this function is in good agreement with the data of Deuling et al [2], Goetzberger et al [3] and Morita et al [4]. The results published by Tredwell and Viswanathan [5] for n-type silicon are in good agreement with ours, too, but for the p-type a different shape of the energy-dependence has been found. The capture cross sections obtained for electrons in the lower part of the gap could not be compared with relevant literature data.

Our results seem to confirm the idea put forward by Goetzberger et al [3] and Morita et al [4]: that the capture cross section is higher when the density of states is low and decreases when the density increases. Since donor states are likely to be found in

the upper half of the gap and acceptor states in the opposite one, the obtained energy dependence of σ_n and σ_p seems to be in full accordance with this assumption in both cases.

On the one hand, the dramatic decrease of the capture cross sections near the band edges can cause—at least partly—the differences between the $D(E)$ results obtained by CV and thermal methods in this energy range. On the other hand, the number of donor and acceptor states can vary in the middle of the gap, and this can lead to a large scatter of the measured capture cross section data published by different authors.

Acknowledgments

The authors are indebted to Prof. Dr. J. Giber for his continuous interest and encouragement during this work, and to Mr. I. Krafcsik and Mr. I. Bársony for help in the design of the cryostat and their consultative assistance.

References

1. E. H. Nicollian and A. Goetzberger, Bell Syst. Techn. J., *46*, 1055, 1967.
2. H. Deuling, E. Klausmann and A. Goetzberger, Solid-State Electronics, *15*, 559, 1972.
3. A. Goetzberger, E. Klausmann and M. J. Schultz, CRC Crit. Rev. in Solid State Sci., *6*, 1, 1976.
4. M. Morita, K. Tsubouchi and N. Mikoshiba, Appl. Phys. Lett. *33*, 8/745, 1978.
5. W. E. Dahlke and D. W. Grewe, Solid-State El., *22*, 893, 1979.
6. J. G. Simmons and G. W. Taylor, Phys. Rev., *135*, 1619, 1972.
7. J. G. Simmons and G. W. Taylor, Solid-State El., *17*, 125, 1974.
8. H. A. Mar and J. G. Simmons, Phys. Rev., *B11*, 775, 1975.
9. L. S. Wei and J. G. Simmons, Solid-State El., *17*, 591, 1974.
10. K. H. Zaininger and G. Warfield, IEEE Trans. Electron Dev. *ED-12*, 179, 1965.
11. K. Yamashita, M. Iwamoto and T. Hino, J. Appl. Phys., *49*, 2866, 1978.
12. T. H. Ning, J. Appl. Phys., *49*, 5997, 1978.
13. A. F. Saunders and G. T. Wright, Electronics Letters, *6*, 207, 1970.
14. J. A. Jr. Cooper and R. J. Schwartz, Solid-State El., *17*, 641, 1974.
15. J. G. Simmons and L. S. Wei, Solid-State El., *17*, 117, 1974.
16. R. Castagné and A. Vapaille, Surf. Sci., *28*, 157, 1971.
17. G. Declerck, R. van Overstraten and G. Broux, Solid-State El., *16*, 1451, 1973.
18. I. Krafcsik and D. Marton, Solid-State El., *22*, 751, 1979.
19. G. Kaplan, Solid-State El., *24*, 89, 1981.
20. M. R. Boudry and J. P. Stagg, J. Appl. Phys., *50*, 942, 1979.

Acta Physica Hungarica 57 (1—2), pp 55—68 (1985)

LOW ENERGY ELECTRON DIFFRACTION BY THIN FILMS
PART I

K. STACHULEC

Technical University, 25-314 Kielce, Poland

(Received 10 July 1983)

A simple dynamical approach to the description of the low energy electron diffraction (LEED) intensity is presented using the quantum field methods developed in thin film theories [1, 2, 5]. The scattering amplitude of the electrons scattered by the metallic thin film is expressed in terms of the surface electron density modulated by lattice vibrations of the atoms at the surface.

The influence of the temperature on the thin film dynamics and then on the LEED intensity is treated in pseudoharmonic approximation [4, 5]. The pseudoharmonic approximation brings an essential correction to the temperature behaviour of LEED spectra, and the approximation can be useful to study the change of the lattice parameters of the film, in particular near the surface.

1. Introduction

Until now the chief motivation for the study of LEED intensities has been the possibility of determining crystal surface structure by intensity analysis in a way analogous to the determination of crystal bulk structure by X-ray diffraction. A common feature of any method of structure determination by diffraction is the possibility to calculate the diffraction intensity for a known structure. This problem, however, in the case of LEED is very difficult that has not yet been treated with the accuracy necessary for structure determination. The difficulties stem mainly from the circumstance that low-energy electrons in crystals suffer very strong scattering, bot elastic and inelastic.

Most attempts to account for the regularities in the energy dependence of the LEED intensity curves involve the calculation of the multiple scattering of electrons between atoms and between atomic layers of the crystal. The general theory of LEED which includes all orders of multiple scattering is called dynamical theory and is practically synonymous with band theory. The results of the dynamical theory of LEED are presented with reference to the band structure of the crystal. One approach to the dynamical theory of LEED intensity which is currently in wide use is nearly free electron treatment due orginally to Bethe [3]. In this theory the scattering power of a crystal is represented by Fourier components U_g of its potential. The zeroth term U_0, the inner potential, causes an overall shift of the band structure to lower energies, while

the higher order terms, U_g are responsible for the band gaps. Apart from the general displacement by U_0, the top edge of each band gap for an attractive potential coincides with a Bragg condition and the width of the gaps $2U_g$, where **g** is the reciprocal lattice vector included in the Bragg condition.

Recently considerable progress has been made in the understanding of the origin of the major features in LEED intensity data and the required dynamical theory has been developed to a point where it can make predictions that may be tested by experiments. However, in addition to being quite complex, the theory is still incomplete, and as a result it has not yet become a practical tool that can be used routinely in the interpretation of structures from experimental data.

In the present paper we give a modified approach to the description of the LEED intensity by means of the methods developed in thin film theories [1, 2]. A scattering sample is treated as a thin film evaporated on substrate. In this way we introduce into consideration of the LEED in a natural way the free surface of the sample which plays a very important role in the problem. We describe the temperature dependence of the surface effects considered in the low energy diffraction introducing a dynamical scattering potential which is a temperature and thickness dependent quantity. The dynamics of the thin film lattice particles are described using the anharmonic model of the thin film crystal in the pseudoharmonic approximation elaborated for bulk crystals in [4] and for thin films in [5]. Using the pseudoharmonic approximation one can explain the temperature dependence of the energy shift of the Bragg peaks, which are due to the change of the lattice distance between atoms and layers, as well as the temperature behaviour of the LEED intensity.

2. Heterogeneity of the scattering potential of a sample with a surface

We shall assume that the experiment has been sufficiently well prepared to consider a monoenergetic, collimated beam of electrons incident on a perfectly clean, well ordered surface of the sample. Our interest will centre on the elastically scattered electrons, because they produce almost all the structure in the diffraction pattern. We divide the scattering sample into a thin film and a substrate. By the thin film we will understand n monoatomic layers parallel to the surface and we number them by ν beginning with $\nu=1$ for the free surface of the film and finishing with $\nu=n$ for the atomic layer which is lying directly on the substrate. We denote the atomic positions inside a ν-th layer of the film by the two dimensional vector **J** and the distance between two atomic sites in the thin film by vector $\mathbf{R}_{\nu j}$.

The physical motivation of the model we chose is as follows. The surface atoms of any solid body are in a situation which is different from that of atoms situated in the inside of the material. The surface atoms feel the changes in the geometry of the neighbours surrounding them caused by the missing neighbours, by the spontaneous deformation of the lattice near the surface. As a result the scattering potential near the

surface must be different from that inside of a bulk material. It is to be expected that the changes of the electronic structure near the surface must cause some changes of all physical properties related to the surface. The redistribution of the electrons near the surface creates some new boundary conditions for lattice vibrations which must influence the temperature-dependence of the low energy electron diffraction intensity, because the low energy electrons in LEED experiments are mainly scattered by a field near the surface of a sample.

To describe the scattering potential $V(r)$ for LEED we suppose that in a static case it can be written as a sum of atomic potentials $V_0(\mathbf{r}-\mathbf{R}_{v\mathbf{j}})$ produced by all atoms of the considered film in the presence of the substrate, that is, we suppose that

$$V(\mathbf{r})=\sum_{v\mathbf{j}} V_0(\mathbf{r}-\mathbf{R}_{vj}). \tag{1}$$

We consider the atomic potential as a sum of the potentials produced by the electrons at the free atom orbitals κ, where κ denotes a set of the quantum numbers for free atom electrons. The potential of a free atom at rest $V_0(\mathbf{r})$ can be expressed by

$$V_0(\mathbf{r})=-\frac{Ze^2}{r}+\sum_{\kappa}\int\frac{e^2\rho_\kappa^0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r', \tag{2}$$

where (Ze^2/r) is the Coulomb potential of the nucleus, ρ_κ^0 denotes the electron density distribution of the κ-th orbital, which can be taken for a free atom in the form [6]:

$$\rho_\kappa^0(r)=\alpha_\kappa e^{-\beta_\kappa r}, \tag{3}$$

where α_κ and β_κ are the numerical parameters which can be found by means of the method presented in [6].

In thin films, however, different atoms are in different circumstances, as we mentioned above, so the electron density distribution of the κ-th orbital may change from one atom to another, particularly, in the direction perpendicular to the surface of the film. Taking into consideration the above statement we propose the following expression for the electron distribution in the $(v\mathbf{j})$-th atom of thin film, namely

$$\rho^0(\mathbf{r}-\mathbf{R}_{vj})=\sum_{\kappa}\langle n_{vj\kappa}\rangle\rho_\kappa^0(\mathbf{r}-\mathbf{R}_{vj}), \tag{4}$$

where $\langle n_{v\mathbf{j}\kappa}\rangle$ denotes the effective numbers of electrons at the κ-th orbitals of the $v\mathbf{j}$-th atom, which can be found in a self-consistent way in which the boundary conditions at the surface are taken into account [2]. In this way the static scattering potential for the electrons diffracted by thin film can be written in the general form [6]

$$V_0(r)=-\sum_{v\mathbf{j}}\left\{\frac{Ze^2}{|r-\mathbf{R}_{v\mathbf{j}}|}-\sum_{\kappa}e^2\langle n_{v\mathbf{j}\kappa}\rangle\int\frac{\alpha_\kappa e^{-\beta_\kappa|\mathbf{r}-\mathbf{R}_{v\mathbf{j}}-\mathbf{j}'|}}{|\mathbf{r}-\mathbf{R}_{v\mathbf{j}}-\mathbf{r}'|}d^3r'\right\}=$$

$$=-Ze^2\sum_{v\mathbf{j}\kappa}\alpha_\kappa\langle n_{v\mathbf{j}\kappa}\rangle\frac{e^{-\beta_\kappa|\mathbf{r}-\mathbf{R}_{v\mathbf{j}}|}}{|\mathbf{r}-\mathbf{R}_{v\mathbf{j}}|}, \tag{5}$$

where e stands for electron charge. It is known, however, that even at zero temperature the atoms of the sample take part in their temperature vibrations around their equilibrium position, and the influence of the temperature on the scattering potential of the real sample must be taken into consideration. In the present paper the influence of the temperature on the scattering potential we take into consideration defining the dynamic scattering potential $V_T(\mathbf{r})$ for ($\nu\mathbf{j}$)-th atom in a thin film as

$$V_T(\mathbf{r}-\mathbf{R}_{\nu\mathbf{j}})=\left(\frac{1}{2\pi B_{\nu\mathbf{j}}}\right)^{3/2}\int V_0(\mathbf{u}-\mathbf{R}^0_{\nu\mathbf{j}})e^{-\frac{|\mathbf{r}-\mathbf{R}^0_{\nu j}-\mathbf{u}|^2}{2B_{\nu j}}}d^3u\,, \tag{6}$$

where $\mathbf{R}^0_{\nu j}$ denotes the equilibrium position of the atom under consideration, $B_{\nu\mathbf{j}}$ denotes its mean square displacement from the equilibrium position and the function $T(r)=(2\pi B_{\nu j})^{-3/2}e^{-\frac{u^2}{2B_{\nu j}}}$ is used as the temperature-dependence distribution of the mass centre of the atom in thin film during its vibration around the equilibrium position [7]. Then the full temperature-dependent scattering potential for the film can be taken in the form [9]

$$V_T(r)=-Ze^2\sum_{\nu\mathbf{j}}\left(\frac{1}{2\pi B_{\nu j}}\right)^{3/2}\int\frac{e^{-\frac{|\mathbf{r}-\mathbf{R}^0_{\nu j}-\mathbf{u}|^2}{2B_{\nu j}}}}{|\mathbf{u}-\mathbf{R}^0_{\nu j}|}d^3u+$$

$$+\sum_{\nu\mathbf{j}\kappa}e^2\langle n_{\nu j\kappa}\rangle\int d^3r'\int d^3u\left(\frac{1}{2\pi B_{\nu\mathbf{j}}}\right)^{3/2}\alpha_\kappa\frac{e^{-\beta_\kappa|\mathbf{r}'-\mathbf{R}_{\nu\mathbf{j}}|}\,e^{-\frac{|\mathbf{u}-\mathbf{R}^0_{\nu j}-\mathbf{r}'|^2}{2B_{\nu\mathbf{j}}}}}{|\mathbf{u}'-\mathbf{R}_{\nu j}-\mathbf{r}'|}\,. \tag{7}$$

After the integration [9, 10] we obtain the following useful expression

$$V_T(\mathbf{r})=+Ze^2\sum_{\nu\mathbf{j}\kappa}\frac{\alpha_\kappa\langle n_{\nu j\kappa}\rangle}{|\mathbf{r}-\mathbf{R}^0_{\nu j}|}e^{\frac{\beta_\kappa^2B_{\nu\mathbf{j}}}{2}}F_\kappa(|\mathbf{r}-\mathbf{R}^0_{\nu j}|)\,, \tag{8}$$

where

$$F_\kappa(|\mathbf{r}-\mathbf{R}^0_{\nu j}|)=e^{-\beta_\kappa|\mathbf{r}-\mathbf{R}^0_{\nu j}|}\left[1-\Phi\left(\beta_\kappa\sqrt{\frac{\beta_{\nu\mathbf{j}}}{2}}-\frac{|\mathbf{r}-\mathbf{R}^0_{\nu\mathbf{j}}|}{\sqrt{2\beta_{\nu\mathbf{j}}}}\right)\right]-$$

$$-e^{+\beta_\kappa|\mathbf{r}-\mathbf{R}^0_{\nu j}|}\left[1-\Phi\left(\beta_\kappa\sqrt{\frac{\beta_{\nu\mathbf{j}}}{2}}+\frac{|\mathbf{r}-\mathbf{R}^0_{\nu\mathbf{j}}|}{\sqrt{2\beta_{\nu\mathbf{j}}}}\right)\right] \tag{9}$$

and

$$\Phi(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt\,. \tag{10}$$

This form of the scattering potential has interesting properties. At first this potential is finite at every lattice point $\mathbf{R}^0_{\nu\mathbf{j}}$, contrary to the effective potential of the free atoms which is infinite at the middle of the atom. An other thing we point out is that for $B_{\nu j}=0$ it comes back to the static form (5). In the next Section we use this form oı the dynamical potential to calculate the temperature dependence of the low energy electron amplitude.

3. Determination of the scattering amplitude for the LEED experiments

The scattering potential in the heterogeneous system derived in the last Section is used now to calculate the amplitude of the scattering of the electrons by thin film for the LEED experiments. To obtain an expression for the scattering amplitude there are possible approximation schemes which are essentially reduced to two alternatives. One consists of a power series expansion in the strength of the potential which can be generated as follows, starting from the integral equation for the wave function of the diffraction problem

$$\psi_k(\mathbf{r})=\left(\frac{1}{2\pi}\right)^{3/2} e^{i\mathbf{kr}}-\frac{1}{4\pi}\int\frac{e^{\pm i\mathbf{k}(\mathbf{r}-\mathbf{r}')}}{|\mathbf{r}-\mathbf{r}'|}V_T(\mathbf{r}')\psi_k^{\pm}(\mathbf{r}')d^3r'=$$

$$=\varphi_{\mathbf{k}}+\int G_0(\mathbf{r},-\mathbf{r}')V_T(r')\psi_{\mathbf{k}}^{\pm}(\mathbf{r}')d^3r' \tag{11}$$

and solving it by iteration for the $V_T(\mathbf{r}')$ potential we thus obtain, omitting the index $\pm$ for brevity, $\psi_{\mathbf{k}}^0=\varphi_{\mathbf{k}}$, where $\mathbf{k}$ denotes the wave vector of the incident electron and $\varphi_{\mathbf{k}}$ is simply its plane wave. Using this on the right-hand side of Eq. (11) we have:

$$\psi_k^{(1)}=\varphi_{\mathbf{k}}+\int G_0(\mathbf{r},\mathbf{r}')V_T(\mathbf{r}')\varphi_{\mathbf{k}}(\mathbf{r}')d^3r', \tag{12}$$

which, used again on the right-hand side of Eq. (11) yields

$$\begin{aligned}\varphi_{\mathbf{k}}^{(2)}&=\varphi_{\mathbf{k}}+\int G_0(\mathbf{r},\mathbf{r}')V_T(\mathbf{r}')\varphi_k(\mathbf{r}')d^3r'+\\&+\int G_0(\mathbf{r},\mathbf{r}')V_T(\mathbf{r}')G_0(\mathbf{r}',\mathbf{r}'')V_T(\mathbf{r}'')\varphi_{\mathbf{k}}(\mathbf{r}'')d^3r''d^3r'+\dots\end{aligned} \tag{13}$$

and so on.

The power series thus obtained is the Born series. It is obvious that using $\psi_{\mathbf{k}}^{(n-1)}$ in the following equation [10]

$$f=\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{k_R r}'}V_T(\mathbf{r}')\psi_{\mathbf{k}}d^3r, \tag{14}$$

which defines the scattering amplitude $f_k(\Omega)$, we obtain n-th order approximation for this quantity. In particular, the first Born approximation is

$$f_k(\theta)=\frac{m}{2\pi\hbar^2}\int e^{+i(\mathbf{k}-\mathbf{k}_R)r}V_T(r)d^3r, \tag{15}$$

where θ stands for the scattering angle m for the electron mass and $\mathbf{k}_R$ denotes the wave vector of the diffracted electron in the place of detection ($|\mathbf{R}|\to\infty$).

An alternative approach consists of analysing the problem in separate partial wave components [10]. In the present paper, however, we do not discuss this approach, limiting ourselves to pointing out only that in this approach our temperature-dependent potential $V_V(\mathbf{r})$ can be very useful because it has the properties required for the partial wave analysis method.

In this paper we would like to concentrate our consideration on the scattering amplitude calculated by means of the first Born approximation for the temperature dependent potential $V_T(r)$. Introducing the scattering vector $\mathbf{k}=\mathbf{k}'-\mathbf{k}$ and taking into consideration the spherical symmetry properties of the scattering potential $V_T(\mathbf{r})$ we can write

$$f(\mathbf{K})=\frac{Zem}{2\pi\hbar^2}\sum_{\mathbf{v}\mathbf{j}\kappa}\alpha_\kappa\langle n_{vj\kappa}\rangle e^{\frac{\beta_\kappa^2 B_{v\mathbf{j}}}{2}}\int\frac{F_\kappa(|\mathbf{r}-\mathbf{R}_{v\mathbf{j}}|)}{|r-R_{v\mathbf{j}}|}e^{-i\mathbf{k}\mathbf{r}}d^3r=$$

$$=\frac{Ze^2m}{2\pi\hbar^2}\sum_{\mathbf{v}\mathbf{j}\kappa}\alpha_k\langle n_{vj\kappa}\rangle e^{+\frac{\beta_\kappa^2 B_{v\mathbf{j}}}{2}-i\mathbf{K}\mathbf{R}_{v\mathbf{j}}^0}I_\kappa(\mathbf{K}), \tag{16}$$

where

$$I_\kappa(\mathbf{K})=\int d^3r\frac{F_\kappa(|\mathbf{r}|)}{r}e^{-i\mathbf{k}\mathbf{r}}. \tag{17}$$

Using the explicit form of the integrand function F_κ (Eq. 17) the integral can be calculated and the result is

$$I_\kappa(\mathbf{K})=\frac{4\pi}{K^2+\beta_\kappa^2}e^{-\frac{K^2B_{v\mathbf{j}}}{2}-\frac{\beta_\kappa^2B_{v\mathbf{j}}}{2}}. \tag{18}$$

Then the scattering amplitude takes the form

$$f(\mathbf{K})=\frac{2Ze^2m}{\hbar^2}\sum_{\mathbf{v}\mathbf{j}\kappa}\left(\frac{\alpha_\kappa}{K^2+\beta_\kappa^2}\right)\langle n_{v\mathbf{j}\mathbf{1}}\rangle e^{-i\mathbf{K}\mathbf{R}_{v\mathbf{j}}^0-\frac{K^2B_{v\mathbf{j}}}{2}}. \tag{19}$$

now taking into the mind the properties of the thin film we can use the equations

$$\langle n_{v\mathbf{j}\kappa}\rangle=\langle n_{v\kappa}\rangle \tag{20}$$

and

$$B_{v\mathbf{j}}=B_v, \tag{21}$$

which are the result of the fact that in the plane parallel to the surface the translation boundary conditions are fulfilled. Then we can write

$$f(\mathbf{K})=F_{\|}\sum_{v=1}^{n}f_v(\mathbf{K})e^{-W_v}, \tag{22}$$

where the quantities F, f_v, and W_v are defined as follows

$$F_{\|}=\sum_{\mathbf{j}}e^{-i\mathbf{K}_{\|}\mathbf{j}}, \tag{23}$$

$$f_v=\sum_\kappa\frac{2me^2Z}{\hbar^2}\alpha_\kappa\frac{\langle n_{v\kappa}\rangle}{K^2+\beta_\kappa^2}e^{K_\perp va}, \tag{24}$$

$$W_v=\frac{1}{2}K^2B_v=\frac{1}{2}\langle(\mathbf{K}\cdot\delta\mathbf{R}_{v\mathbf{j}}^0)^2\rangle, \tag{25}$$

where $\delta\mathbf{R}_{\nu\mathbf{j}}$ denotes the displacement of the ($\nu\mathbf{j}$)-th atom from its equilibrium position ($\delta\mathbf{R}_{\nu j}=\mathbf{R}_{\nu\mathbf{j}}-\mathbf{R}^0_{\nu\mathbf{j}}$).

In the above equation a denotes the lattice constant, $aN=L$ is the normalization length in the plane parallel to the surface of the film, while $\mathbf{K}_{\|}$ and $\mathbf{K}_{\perp}$ are the components of the scattering vector parallel and perpendicular to the surface plane of the film, respectively.

It is easy to see from Eq. (22) that the scattering amplitude of the film $f(\mathbf{K})$ is sum of the partial scattering amplitudes of the monoatomic layers f_ν defined by Eq. (24). Each of the last amplitudes is multiplied by its own factor e^{-W_ν} which we name the Debye—Waller factor, on the analogy of the bulk theory.

4. Electron density distribution in thin films

In the previous Section we have seen that our calculation of the LEED amplitude involves the determination of the electron density distribution across the thin film $\langle n_{\nu\kappa}\rangle$. In this Section we consider this problem in detail. To describe a behaviour of the electrons in thin films, usually one supposes some boundary conditions for the wave functions on the surfaces. These boundary conditions depend on the nature of a concrete problem chosen to be considered. The boundary conditions for thin films have been discussed in the case of magnetic [2], superconducting [15], and semiconducting problems [16]. In these discussions thin films have been treated mainly in the effective mass approximation. There are other papers [1, 2, 5] where the boundary conditions are introduced in the natural way by means of the second quantization method. In this case a Hamiltonian describing the behaviour of the electrons in thin films has the general form [2]

$$H=\sum_{\lambda\mu} t_{\lambda\mu} a^{t}_{2\mu} a_{\mu\kappa}+\sum_{\lambda\mu\rho\kappa}\frac{1}{2} I_{\lambda\mu\rho\kappa} a^{+}_{\lambda} a^{+}_{\mu} a_{\rho} a_{\kappa}\,, \tag{26}$$

where the indices λ, μ, ρ, κ represent the sets of the quantum numbers of the wave functions by means of which the matrix elements $t_{\lambda\mu}$ and $I_{\lambda\mu\rho\kappa}$ are calculated. By means of this Hamiltonian one can obtain any model used up to now in the theory of thin films [2]. Because of the complexity we expect, however, that the electron density distribution in thin films $\langle n_\nu\rangle$ can be in the first approximation sufficiently described by the following simplified form of Hamiltonian given by

$$H=\sum_{\nu j\kappa}\sum_{\nu' j'\kappa'} t_{\nu\mathbf{j}\kappa\nu'\mathbf{j}'\kappa'} c^{+}_{\nu j\kappa} c_{\nu'\mathbf{j}'\kappa'}. \tag{27}$$

The form of the latter thin film Hamiltonian means that we suppose as the sets of the basic second quantization function the localized atomic orbitals of the free atom

$$|\lambda\rangle=|\nu\mathbf{j}\kappa_{\nu\mathbf{j}}\rangle=|\nu\mathbf{j}\kappa\rangle\,, \tag{28}$$

where $\nu\mathbf{j}$ determine the position of the atom in thin film and $\kappa_{\nu j}$ its orbital quantum numbers of the free atom. The Hamiltonian (27) has been used many times in the theory of thin films [2]. We use it here, however, in its temperature-dependent form modified by the crystal potential $V_T(r)$. In this case the matrix elements $t_{\nu\mathbf{j}\kappa\nu'\mathbf{j}'\kappa}$ are changed by $t^T_{\nu\mathbf{j}\kappa\nu'\mathbf{j}'\kappa}$ and we define them as follows. The diagonal elements $t^T_{\nu j\kappa\nu j\kappa}$ are of the form

$$t^T_{\nu j\kappa\,\nu j\kappa}=E_\kappa+A^T_\kappa(\nu)\,, \tag{29}$$

where

$$A^T_\kappa(\nu)=\langle\nu\mathbf{j}\kappa|(V_T{}^{(r)}-V_0(\mathbf{r}-\mathbf{R}_{\nu\mathbf{j}}))|\nu j\kappa\rangle, \tag{30}$$

but nondiagonal elements are given by

$$t^T_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa}=-B^T_\kappa(\nu\nu')\qquad \kappa=\kappa' \tag{31}$$

and

$$B^T_\kappa(\nu\nu')=\langle\nu\mathbf{j}\kappa|\,\{V_T(r)-V_0(\mathbf{r}-\mathbf{R}_{\nu'\mathbf{j}'})\}\,|\nu'\mathbf{j}'\kappa\rangle\,. \tag{32}$$

In this model we are dealing with the matrix elements for the difference between the temperature-dependent crystal potential $V_T(r)$ and the static atom potential $V_0(\mathbf{r}-\mathbf{R}_{\nu\mathbf{j}})$ localized at the $\nu\mathbf{j}$-th site of the thin film lattice calculated by means of the two atomic orbitals of the free atom localized in the same site $\nu\mathbf{j}$ (30) and by means of the two atomic orbitals localized in the different site (32) of the crystal lattice. The other quantity which appears in Eq. (29) i.e. free atomic orbitals energy E^0_κ are the eigenvalues of the equation

$$\left[\frac{\rho^2}{2m}+V_0(\mathbf{r}-\mathbf{R}_{\nu\mathbf{j}})\right]|\nu\mathbf{j}\kappa\rangle=E^0_\kappa|\nu\mathbf{j}^\kappa\rangle\,. \tag{33}$$

Now we can use the Green function method to calculate the electron density distribution in thin films the Hamiltonian of which is

$$H=\sum_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}t^T_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}c^+_{\nu\mathbf{j}\kappa}c_{\nu'\mathbf{j}'\kappa'}\,. \tag{34}$$

The procedure is known from the the static potential case [2]. Defining as usual the following set of Green functions

$$G_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}=\langle\langle c_{\nu\mathbf{j}\kappa}|c^+_{\nu'\mathbf{j}'\kappa'}\rangle\rangle \tag{35}$$

and setting for them the equations of motion one has to solve the following sets of the equations

$$EG_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}=\delta_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}+\sum_{\nu''\mathbf{j}''\kappa''}t^T_{\nu\mathbf{j}\kappa\,\nu''\mathbf{j}''\kappa''}G_{\nu''\mathbf{j}''\kappa''\,\nu'\mathbf{j}'\kappa'}\,. \tag{36}$$

The method of the solution of these equations is given in [2] for the Hamiltonian (27) and we do not repeat it here. The only difference for our temperature-dependent Hamiltonian (34) is that the matrix elements are now temperature-dependent. Using then the method presented in [2] for the sought solution of Eq. (36) one obtains the following Green functions for the present problem

$$G_{\nu\mathbf{j}\kappa\,\nu'\mathbf{j}'\kappa'}=\sum_{\tau\mathbf{h}}\Gamma^{\mathbf{h}\kappa}_{\nu\tau}\Gamma^{\mathbf{h}\kappa}_{\nu'\tau}e^{i\mathbf{h}(\mathbf{j}'-\mathbf{j})}\frac{\delta_{\kappa\kappa'}}{E-E_{\tau\mathbf{h}}}, \tag{37}$$

where coefficients $\Gamma_{v\tau}^{\mathbf{h}\kappa}$ are the solutions of the following sets of equations

$$\sum_{\bar{v}\bar{\kappa}} (t_{v\kappa\bar{v}\bar{\kappa}}^{T} \Gamma_{\bar{v}\tau}^{\mathbf{h}\bar{\kappa}} - E_{\tau\mathbf{h}\kappa} \delta_{\bar{v}v} \delta_{\bar{\kappa}\kappa} \Gamma_{\bar{v}\tau}^{\mathbf{h}\bar{\kappa}}) = 0 , \tag{38}$$

where $E_{\tau\mathbf{h}\kappa}$ are given by

$$\sum_{\bar{v}\bar{\kappa}} t_{v\kappa\bar{v}\bar{\kappa}}^{T} \Gamma_{\bar{v}\tau}^{\mathbf{h}\bar{\kappa}} G_{\tau\mathbf{h}}^{\bar{\kappa}\kappa'} = E_{\tau\mathbf{h}\kappa} \Gamma_{v\tau}^{\mathbf{h}\kappa} G_{\tau h}^{\kappa\kappa'} , \tag{39}$$

$$t_{v\kappa\bar{v}\bar{\kappa}}^{T} = \sum_{d} t_{v\mathbf{j}\bar{v}\bar{\mathbf{j}}\bar{v}} e^{i\mathbf{h}(\mathbf{j}-\bar{\mathbf{j}})}, \tag{40}$$

and can be generally written in the form [1]:

$$\Gamma_{v\tau}^{\mathbf{h}\kappa} = A_{\tau}^{\mathbf{h}\kappa} \cos(\alpha_{\tau}^{T} v + \beta_{\tau}^{T}). \tag{41}$$

About the coefficients $\Gamma_{v\tau}^{\mathbf{h}\kappa}$ from now we suppose that they do not depend on the quantum number κ, i.e. we put

$$\Gamma_{v\tau}^{\mathbf{h}\kappa} = \Gamma_{v\tau}^{\mathbf{h}} = A_{\tau}^{\mathbf{h}} \cos(\alpha_{\tau}^{T\,v} + \beta_{\tau}^{T}) , \tag{42}$$

where $A_{\tau}^{\mathbf{h}}$ is the normalization constant and coefficients α_{τ}^{T} and β_{τ}^{T} are the temperature-dependent quantity which can be calculated numerically only. We suppose, however, that they can be used in the same analytical form as in the static potential case. The other relations between α_{τ}^{T} and β_{τ}^{T} are dependent on the crystallographic structure and the orienation of the sample and are the same as between α_{τ} and β_{τ} (see [1]).

By means of the Green function $G_{v\mathbf{j}\kappa\, v'\mathbf{j}'\kappa'}$ thus determined we can simply express the electron density distribution across the thin film, namely, we have

$$\langle n_{v\kappa} \rangle \equiv \langle n_{v\mathbf{j}\kappa} \rangle = \langle c_{v\mathbf{j}\kappa}^{t} c_{v\mathbf{j}\kappa} \rangle =$$

$$= \frac{i}{2\pi} \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} dE \sum_{\tau\mathbf{h}\kappa} \frac{G_{\tau\mathbf{h}\kappa}(E+i\varepsilon) - G_{\tau\mathbf{h}\kappa}(E-i\varepsilon)}{e^{\frac{E-E_F^{+}}{kT}} + 1} (\Gamma_{.v\tau}^{\mathbf{h}})^2 = \sum_{\tau\mathbf{h}\kappa} (\Gamma_{v\tau}^{\mathbf{h}})^2 f(E_{\tau\mathbf{h}\kappa}) , \tag{43}$$

where $f(E_{\tau\mathbf{h}\kappa})$ is the Fermi distribution function

$$f(E_{\tau\mathbf{h}\kappa}) = \frac{1}{e^{\frac{E_{\tau\mathbf{h}v} - E_F}{k_B T}} + 1} . \tag{44}$$

There $E_{\tau\mathbf{h}\kappa}$ denote the energy of the electrons in the thin film and E_F stands for the temperature-dependent Fermi energy of thin film, in κ-th band. The energy of the electron in thin crystal films is actually given approximately by the following expression

$$E_{\tau\mathbf{h}\kappa} = E_{\kappa}^{0} + A_{\kappa}^{T} \pm 2B_{\kappa}^{T} (\cos h_x a + \cos h_y a + \cos \alpha_{\tau} a) . \tag{45}$$

This equation describes the temperature-dependence of the band structure of thin films. This temperature-dependence of the band structure corresponds to the low energy electron diffraction spectra calculated in the dynamical theory of LEED.

Using Eq. (45) one can calculate the Fermi energy for each band which is created by the given free atomic orbitals κ. Now we limit our consideration to the one band approximation, i.e. we suppose that all free atomical orbitals κ in thin films are sufficiently localized in the energy space except the valence orbital which creates the film conduction subbands numbered by index τ. In this simple case the electron density distribution can be written by

$$\langle n_{\nu\kappa} \rangle = \begin{cases} \mu \\ \langle n_{\nu w} \rangle \end{cases}, \tag{46}$$

where μ denotes the number of electrons for κ-th orbital if it is not a valence orbital (κ includes the degeneracy of the orbital due to the spin σ). For the valence orbital we then have $\langle n_{\nu w} \rangle$, which we can write as

$$\sum_{\tau} \frac{S}{4\pi^2} (T^{\mathbf{h}}_{\nu\tau})^2 \int f(E^T_{\tau h}) d^2 h = \langle n_\nu \rangle . \tag{47}$$

The Fermi energy for the valence band can be calculated by means of the normalization condition of the Fermi distribution

$$2 \sum_{\tau h} f(E_{\tau h}) = nN^2, \tag{48}$$

where factor 2 stands for the spin degeneration of the electrons, nN^2 denotes the number of the valence electrons in the thin film, and the energy

$$E_{\tau \mathbf{h}} = \frac{\hbar^2 h^2}{2m} + \varepsilon_0 \tau^2 \qquad (\tau = 1, 2, \ldots) \tag{50}$$

is counted from the low energy edge of the valence band.

The normalization condition (48) leads to the equations

$$kT \sum_{\tau=1}^{M} \ln \left[\frac{1}{2} (1 + e^{\frac{E - \varepsilon_0 \tau^2}{kT}}) \right] - \frac{\pi n \hbar^2}{ma^2} 0 \tag{51}$$

$$M = \text{Integer} \left(E : \varepsilon_0 \right)^{1/2} \tag{52}$$

which must be solved numerically for energy E. Then for the effective number distribution of the valence electrons in the direction perpendicular to the surface of the film we obtain

$$\langle m_{\nu w} \rangle = \frac{2mkTa^2}{\pi \hbar^2} \frac{1}{n} \sum_{\tau=1}^{M} \frac{2 \cos^2 (\alpha^T_\tau \nu + \beta^T_\tau)}{\left(n + \frac{\sin (\alpha^T_\tau n) \cos [\alpha^T_\tau (n+1) + 2\beta^T_\tau]}{\sin \alpha^T_\tau} \right)} \ln \left[\frac{1}{2} (1 + e^{\frac{E_F - \varepsilon_0 \tau^2}{kT}}) \right] \tag{53}$$

$$M = \text{Integer} \left(\frac{E_F}{\varepsilon_0} \right)^{1/2} \tag{54}$$

Equations (51) with (53) and (54) determine the electron density distribution in thin films in the case of the one band approximation.

5. Distribution of the Debye—Waller factors for thin films in pseudoharmonic approximation

The calculation of the scattering amplitude for the LEED experiments interpretation by means of our formula (22) involves not only the determination of the electron density distribution in thin films, which we considered in the previous Section, but it also involves the calculation of the distribution of the Debye—Waller factors in thin films. This calculation can be done by means of a method presented in our previous paper [9] where we have discussed the Debye—Waller factors for thin films in harmonic and pseudoharmonic approximation, and for that reason in this Section we give only the essential relations and definitions which are necessary for future numerical calculations.

At first the Debye—Waller factor g_ν for a given monoatomic layer ν is defined by the equation

$$g_\nu = e^{-W_\nu}, \tag{55}$$

where the Debye—Waller coefficient W_ν is given by

$$W_\nu = \frac{1}{2}\langle(\mathbf{K}\cdot\delta\mathbf{R}_{\nu\mathbf{j}})^2\rangle \tag{56}$$

and $\mathbf{K}$ denotes the scattering vector $\mathbf{K}=\mathbf{K}'-\mathbf{K}$, while $\delta\mathbf{R}_{\nu\mathbf{j}}$ denotes as usual the displacement vector of the νj-th atom.

The phonons theory elaborated by Wojtczak and Zajac [1] allows to express the displacement $\delta\mathbf{R}_{\nu\mathbf{j}}$ by means of creation $a^t_{\mathbf{q}\lambda\mu}$ and annihilation operators $a_{\mathbf{q}\lambda\mu}$ and phonons amplitudes $T^{\mathbf{q}}_{\nu\lambda}$ as follows

$$\delta\mathbf{R}_{\nu\mathbf{j}} = \left(\frac{\hbar}{2MN^2n}\right)^{1/2} \sum_{\mathbf{q}\lambda\mu} \frac{e_{\mathbf{q}\lambda\mu}}{\sqrt{\omega_{\mathbf{q}\lambda\mu}}} T^{\mathbf{q}}_{\nu\lambda} e^{i\mathbf{q}\mathbf{j}}(a_{\mathbf{q}\lambda\mu}+a^t_{\mathbf{q}\lambda\mu}). \tag{57}$$

$\omega_{\mathbf{q}\lambda\mu}$ are the frequency eigenvalues of phonons, N^2 denotes the number of atoms in each atomic layer, n denotes the number of the atomic layers parallel to the surfaces of the film while $\mathbf{e}_{\mathbf{q}\lambda\mu}$ denotes the polarization vector of the phonon in $(\mathbf{q}\lambda\mu)$ state. By means of Eq. (57) the Debye—Waller coefficient W_ν (Eq. 56) can be expressed as

$$W_\nu = \frac{\hbar K^2}{4MN^2n} \sum_{\mathbf{q}\lambda} T^2_{\nu\lambda} \left\{ \sin^2\theta\cos^2\varphi \frac{\operatorname{cth}\left(\frac{\hbar\omega_{\mathbf{q}\lambda 1}}{2KT}\right)}{\omega_{\mathbf{q}\lambda 1}} + \right.$$

$$\left. +\sin^2\theta\sin^2\varphi \frac{\operatorname{cth}\left(\frac{\hbar\omega_{\mathbf{q}\lambda 2}}{2KT}\right)}{\omega_{\mathbf{q}\lambda 2}} + \cos^2\theta \frac{\operatorname{cth}\left(\frac{\hbar\omega_{\mathbf{q}\lambda 3}}{2KT}\right)}{\omega_{\mathbf{q}\lambda 3}} \right\} \Lambda \tag{58}$$

where the frequencies $\omega_{\mathbf{q}\lambda 1}=\omega_{\mathbf{q}\lambda 2}$ and $\omega_{\mathbf{q}\lambda 3}$ in the pseudoharmonic case are given by

$$\omega_{\mathbf{q}\lambda 1} = \alpha_T\omega^0_{\mathbf{q}\lambda 1}, \tag{59}$$

$$\omega_{\mathbf{q}\lambda 3} = \alpha_T\omega^0_{\mathbf{q}\lambda 3} \tag{60}$$

and $\omega^0_{\mathbf{q}\lambda\mu}$ denote the harmonic frequency and we approximate them by

$$\omega^0_{\mathbf{q}\lambda 1}=\omega_m\sqrt{\sin^2\frac{q_xa}{2}+\sin^2\frac{q_ya}{2}+\frac{A'}{A}\sin^2\frac{\alpha_\lambda}{2}}, \tag{61}$$

$$\omega^0_{q\lambda 3}=\bar{\omega}_m\sqrt{\sin^2\frac{q_xa}{2}+\sin^2\frac{q_ya}{2}+\frac{B'}{B}\sin^2\frac{\alpha_\lambda}{2}}, \tag{62}$$

$$\omega_m=\sqrt{\frac{4A}{M}};\qquad \bar{\omega}_m=\sqrt{\frac{4B}{M}}, \tag{63}$$

where A, A' and B, B' denote the force constants by means of which we describe the interaction between two atoms inside the film and on the surface. The renormalization coefficient α_T which is temperature-dependent and is defined by the rate of the pseudoharmonic and harmonic force contstants, can be calculated by means of the Siklós method [4, 5, 9, 8]. Supposing that it is the same for all branches μ one obtains

$$\alpha_T^2=1-\frac{36}{a^2}\left[\frac{1}{2n}\sum_{\nu=1}^{n}\langle(\delta\mathbf{R}_{\nu\mathbf{j}}-\delta\mathbf{R}_{\nu+1\mathbf{j}})^2+(\delta\mathbf{R}_{\nu\mathbf{j}}-\delta\mathbf{R}_{\nu-1\mathbf{j}})^2\rangle\right]. \tag{64}$$

The sum over $\mathbf{q}$, due to the periodic conditions which are fulfilled in the direction parallel to the plane of the film, can be replaced by an integral using the relation

$$\sum_{\mathbf{q}}(\ldots)=\frac{S}{4\pi}\int_{-\frac{\Pi}{a}}^{\frac{\Pi}{a}}\int_{\frac{\Pi}{a}}^{\frac{\Pi}{a}}dq_x dq_y(\ldots), \tag{65}$$

thus Eq. (58) becomes

$$W_\nu=\frac{\hbar K^2kT}{4MN^2m}\sum_{\lambda=1}\{I_1(\lambda)\sin^2\theta+I_3(\lambda)\cos^2\theta\}, \tag{66}$$

where

$$I_1(\lambda)=\frac{4N^2kT}{\pi\hbar\omega_m^2\alpha_T^2}\ln\frac{\mathrm{sh}\left(\frac{\hbar\omega_m\alpha_T}{2kT}\sqrt{2+\varepsilon_A\sin^2\frac{\alpha_\lambda}{2}}\right)}{\mathrm{sh}\left(\frac{\hbar\omega_m\alpha_T}{2kT}\sqrt{\varepsilon_A\sin^2\frac{\alpha_\lambda}{2}}\right)}, \tag{67}$$

$$I_3(\lambda)=\frac{4N^2kT}{\pi\hbar\bar{\omega}_m^2\alpha_T^2}\ln\frac{\mathrm{sh}\left(\frac{\hbar\bar{\omega}_m\alpha_T}{2kT}\sqrt{2+\varepsilon_B\sin\frac{2\alpha_\lambda}{2}}\right)}{\mathrm{sh}\left(\frac{\hbar\bar{\omega}_m\alpha_T}{2kT}\sqrt{\varepsilon_B\sin^2\frac{\alpha_\lambda}{2}}\right)}, \tag{68}$$

$$\varepsilon_A\equiv\frac{A'}{A};\qquad \varepsilon_B\equiv\frac{B'}{B} \tag{69}$$

and we have made the use of the fact that in our model $\omega_{\mathbf{j}\lambda 1}=\omega_{\mathbf{q}_{\lambda 2}}$. Introducing the following quantity

$$\eta_v(n)=\frac{1}{n}\sum_{\lambda=1}^{n} T_{v\lambda}^2 \ln \frac{\operatorname{sh}\left(\frac{\hbar\omega_m\alpha_T}{2kT}\sqrt{2+\varepsilon_A\sin^2\frac{\alpha_\lambda}{2}}\right)}{\operatorname{sh}\left(\frac{\hbar\omega_m\alpha_T}{2kT}\sqrt{\varepsilon_A\sin^2\frac{\alpha_\lambda}{2}}\right)}, \tag{70}$$

$$\gamma_v(n)=\frac{1}{n}\sum_{\lambda=1} T_{v\lambda}^2 \ln \frac{\operatorname{sh}\left(\frac{\hbar\bar{\omega}_m\alpha_T}{2kT}\sqrt{2+\varepsilon_B\sin^2\frac{\alpha_\lambda}{2}}\right)}{\operatorname{sh}\left(\frac{\hbar\bar{\omega}_m\alpha_T}{2kT}\sqrt{\varepsilon_B\sin^2\frac{\alpha_\lambda}{2}}\right)}, \tag{71}$$

then we can rewrite Eq. (66) in the form

$$W_v=\frac{K^2kT\eta_v(n)}{\pi M\omega_m^2\alpha_T^2}\left[\sin^2\theta+\frac{\omega_m^2}{\bar{\omega}_m^2}\frac{\gamma_v(n)}{\eta_n(n)}\cos^2\theta\right]=$$

$$=\frac{K^2kT\eta_v(n)}{4\pi A\alpha_T^2}\left[1+\xi_n(n)\cos^2\theta\right], \tag{72}$$

where

$$\xi_v(n)=\frac{\omega_m^2}{\bar{\omega}_m^2}\frac{\gamma_v(n)}{\eta_v(n)}-1=\frac{A}{B}\frac{\gamma_v(n)}{\eta_v(n)}-1 \tag{73}$$

and be called the anisotropy coefficients of the Debye—Waller factor for the v-th monoatomic layer of the film.

From the last form for the W_v-coefficient one can see the anisotropic character of the Debye—Waller factor as well as its position and the thickness-dependence.

6. Conclusion

In this paper we have presented the theoretical model of the description of the low energy electron diffraction by thin films using the field theory methods known in the literature. In the paper the expression of the scattering amplitude for the electron scattered by the film into the component amplitudes of the given monoatomic layers is a natural consequence of introducing the dynamical scattering potential which depends on temperature as well as on the position of the scattering atoms in the thin film. Using the field theory method allowed us to calculate the approximate electron distribution across the film as well as the distribution of the Debye—Waller factor of thin film, which are involved in the calculation of the component amplitudes of the electron scattered by the particular monoatomic layers.

Due to the pseudoharmonic approximation by means of which in our model the lattice dynamics of thin film atoms is treated, one can use this model to search for the

low energy electron diffraction problem in the harmonic as well as in the anharmonic case of thin films. In particular, it can be useful to study the effect of the change of the surface lattice constants, which, as is known from LEED experiments, changes the location of the suitable Bragg's peaks on the energetic scale.

Concluding we would like to point out that the theoretical approach to the LEED problem presented here is sufficiently complete to be treated numerically. We have done such a numerical investigation and present the results in a supplementary paper [17].

References

1. L. Wojtczak, Š. Zajac, Bulletin de l'Académie Polonaise des Sciences, *16*, 527, 1968.
2. L. Valenta and L. Wojtczak, Czech. J. Phys., *B30*, 1025, 1980.
3. H. Bethe, Ann. Physik, *87*, *55*, 1928.
4. T. Siklós, Acta Phys. Hung., *30*, 181, 1971; ibid., *30*, 193, 1971, ibid., *30*, 301, 1971, ibid., *34*, 327, 1973.
5. J. Zimnicki, Acta Phys. Hung., *30*, 27, 1975.
6. T. G. Strand and T. Tietz, Nuovo Cimento, *41B*, 89, 1966.
7. W. F. Dvoriankin, Kristalografija, *9*, 2, 1964.
8. L. Wojtczak and K. Stachulec, Phys. Stat. Sol., *b70*, K165, 1975.
9. K. Stachulec, Acta Phys. Hung., *54*, 267, 1983.
10. A. Stachulec, Acta Phys. Hung., *51*, 243, 1981.
11. A. Messiah, Quantum Mechanics, North-Holland Publishing Company, Amsterdam, 1965.
12. P. J. Estrup and E. G. McRae, Surface Sciences, *25*, 1, 1971.
13. A. Corciovei and A. Berinde, J. de Physique, *24*, 89, 1963.
14. J. B. Pendry, Low Energy Electron Diffraction, Academic Press, London and New York, 1974.
15. M. Green, Solid State Surface Science, New York, 1969.
16. T. J. Couttis, Electrical Conduction in Thin Metal Films, Elsevier Scientific Publishing Company, Amsterdam—London—New York, 1974.
17. K. Stachulec, Acta Phys. Hung., to be published.

Acta Physica Hungarica 57 (1—2), pp. 69—72 (1985)

THERMODYNAMIC PROPERTIES OF ANHARMONIC NOBLE METALS IN TERMS OF INTERATOMIC POTENTIALS FROM EXPERIMENTAL PHONON SPECTRA

C. MALINOWSKA-ADAMSKA

Institute of Physics, Technical University of Łódź
Łódz, Poland

(Received 27 July 1983)

A new model interatomic pair potential derived from the experimental phonon spectra is applied to the investigation of the thermodynamical properties of anharmonic noble metals.

The computed Debye temperature, Grüneisen parameter, lattice specific heat at constant volume and volume thermal expansion coefficient are compared with the available calorimetric data. The theory is in reasonably good agreement with the experimental data.

The knowledge of two-body interatomic potentials is necessary for the understanding of many of the physical properties of simple metals. Various models have been proposed to describe lattice energy as a function of the atomic positions in terms of some physical parameters [1]. These parameters are then fitted to experimental data, e.g., elastic constant, phonon spectra, threshold energy of radiation damage, channeling and shock wave data [2]. A procedure to directly obtain data on the interatomic potential in cubic lattices, using all the information contained in the phonon spectrum, was proposed by Brosens and co-workers [3]. The close agreement obtained between the Brosens' potential and one calculated from pseudopotential theory led Esterling and Swaroop [4, 5] to consider an application of this approach to the general class of simple, noble and transition metals [5].

The calculated pair potential [4]

$$\Phi_i(r)=\sum_{j=1}^{5} C_{j,i}(r-r_i)^{j-1}, \qquad i=1,2,\ldots N_s, \tag{1}$$

where r_i is the distance to the i-th nearest neighbour, $C_{j,i}$ are the polynomial coefficients, and N_s is the number of shells of neighbours) was found to be reliable and accurate on the basis of comparison made with experimental data for the elastic constant and bulk modulus at the temperature $T=0$ K.

Using the model potential (1), in this note we explore various thermodynamical properties like Debye temperature Θ_D, Grüneisen parameter γ, volume thermal expansion coefficient β and volume heat capacity C_v in the anharmonic approximation

which assumes short range interactions between atoms [6, 7]. We choose this simple approach in order to assess the validity of our potential model for qualitative description of the thermal properties of the cubic metals over a wider temperature range. The relevant expressions used here have been taken from [6—8].

So, the Debye temperature [8]

$$\Theta_D = n(5K_T r_e)^{1/2}/k_B . \tag{2}$$

The Grüneisen parameter [8]

$$\gamma = -\frac{r_e \Phi^{III}(r)}{6\Phi^{II}(r)} . \tag{3}$$

The volume thermal expansion coefficient [8]

$$\beta = -\frac{C_v}{2r_e N_A}\left\{\frac{\Phi^{III}(r)}{[\Phi^{II}(r)]^2}\right\} . \tag{4}$$

The molar heat at constant volume C_v in the low $(\Theta \ll \omega_L)$ and high $(\Theta \gg \omega_L)$ temperature limits are respectively given by [6, 7]

$$C_v = \frac{12}{5}\pi^4 R\left(\frac{T}{T_A}\right)^3 \quad \text{for} \quad \Theta \ll \omega_L , \tag{5}$$

and

$$C_v = 3R\left[1 - \Theta C_x - 0.04166\left(\frac{\omega_L}{\Theta}\right)^2\right] \quad \text{for} \quad \Theta \gg \omega_L . \tag{6}$$

The symbols appearing in (5) and (6) have the following meaning:

$$\Theta = k_B \cdot T ,$$

$$T_A = \Theta_D \alpha (1 + 8A\omega_L C_x) ,$$

$$\alpha = 1 - A\omega_L G ,$$

$$\omega_L = \hbar \sqrt{\frac{8\Phi^{II}(\bar{r}_e)}{m}} ,$$

$$C_x = 0.125(K - BG) ,$$

$$G = [\Phi^{III}(r)]^2/[\Phi^{II}(r)]^3 ,$$

$$K = \Phi^{IV}(r)/[\Phi^{II}(r)]^2 .$$

$\Phi^{II}(r)$, $\Phi^{III}(r)$ and $\Phi^{IV}(r)$ are the second-, third- and fourth-order spatial derivatives of the potential $\Phi_i(r)$ taken at the equilibrium point r_e which is determined by minimizing of the total free energy. $\bar{r}_e$ corresponds to the first minimum of the potential energy [7, 9]. The numerical coefficients are $A = 0.04257$, $B = 0.2992$. K_T, m and R are the isothermal bulk modulus, mass of atom and gas constant, respectively.

The values of Debye temperature Θ_D and Grüneisen parameter γ calculated for $N_s=3$ are given and compared with experimental data [10—12] in Table I. The temperature behaviour of the volume thermal expansion coefficient β and molar heat at constant volume are listed together with the observed results in Table II. The experimental values of C_v were obtained from the relation $C_v=C_p-\beta^2 K_T\, T/\rho$ using experimental data for all quantities on the right-hand side of the equation, in which C_p is the molar heat at constant pressure, and ρ is the density. The values of isothermal bulk modulus for copper were obtained from the ultrasonic measurements of Overton and Gaffney [13]. For silver and gold these are taken from the work of Neighbours and Alers [14]. The sources of other experimental data for noble metals selected for the present comparison are shown in Table III.

It is evident from Tables I and II that the values of Θ_D, β, γ and C_v calculated for the Esterling and Swaroop potential are in reasonably good agreement with their measured data. Bearing in mind that the comparison of calculated and observed

Table I

Debye temperature Θ_D and Grüneisen parameter γ of noble metals at $T=0$ K in terms of the Esterling and Swaroop pair potential

Solids	Θ_D[K]		γ	
	our value	experiment	our value	experiment
Cu	349.46	346.7	1.75	1.7
Ag	228.55	226.5	2.53	2.7
Au	166.44	164.8	2.15	2.3

Table II

Comparison of the experimental data for the volume thermal expansion coefficient β and molar heat at constant volume C_v for noble metals with those calculated in terms of the interatomic potentials from experimental phonon spectra

T [K]	Cu				Ag				Au			
	β*		C_v**		β		C_v		β		C_v	
	Ex	P	Ex	P	Ex	P	Ex	P	Ex	P	Ex	P
20	0.69	0.96	1.00	0.46	4.05	3.84	2.08	1.60	6.33	9.66	5.15	4.20
50	11.52	13.42	6.40	6.39	23.94	27.77	11.60	11.57	22.95	35.68	14.35	15.52
100	30.99	36.81	16.00	17.53	41.13	48.53	19.96	20.22	33.89	55.14	21.40	23.97
150	40.80	46.15	20.49	21.91	47.58	57.46	22.57	23.94	37.83	57.35	23.48	24.93
200	45.54	50.30	22.67	23.95	52.26	58.52	23.66	24.38	39.90	57.46	24.41	24.98
300	49.83	52.23	23.63	24.87	56.19	59.85	24.50	24.98	42.30	57.52	25.35	25.00

* β in units of $10^{-6}\left[\frac{1}{\mathrm{K}}\right]$; Ex — the experimental data, P — present calculations

** C_v in units of $\left[\frac{\mathrm{J}}{\mathrm{mol\,K}}\right]$

temperature variation of the molar heat at constant volume and volume thermal expansion coefficient is a demanding test it appears from Table II that the potential model derived from the experimental spectra is reasonably realistic for noble metals. A next improvement in the description of the anharmonic noble metals is the investigation effects of phonons in thermodynamic properties from the present model in the self-consistent (s.c.) theory which allows to take into account in the lower order perturbation theory all the higher order anharmonic terms in a s.c. manner [21]. Work in this direction is in progress.

Table III

Sources of experimental data for the volume thermal expansion coefficient β and molar heat at constant pressure C_p

Solids	Quantity	Source	Ref.
Cu	β	Hahn	[15]
	C_p	Martin	[16]
Ag	β	Fraser and Hollis Hallet	[17]
		Leksina and Novikova	[18]
	C_p	Meads, Forsythe and Giauque	[19]
Au	β	Fraser and Hollis Hallet	[17]
		Leksina and Novikova	[18]
	C_p	Geballe and Giauque	[20]

References

1. P. C. Gehlen, J. R. Beeler, Jr, and R. I. Jaffee (Ed.) Interatomic Potentials and Simulation of Lattice Defects, Plenum Press, New York, 1972.
2. I. M. Torrenz, Interatomic Potentials, Academic Press, New York, 1972.
3. F. Brosens, J. Cornelis, D. C. Wallance and J. T. Devreese, Phys. Stat. Sol., *(b)81*, 557, 1977.
4. D. M. Esterling and A. Swaroop, Phys. Stat. Sol., *(b)96*, 401, 1979.
5. A. Swaroop and D. M. Esterling, Phys. Stat. Sol., *(b) 96*, 691, 1979.
6. A. A. Maradudin, P. A. Flinn and R. A. Coldwell-Horsfall, Ann. Phys. (USA), *15*, 337, 360, 1961.
7. P. A. Flinn and A. A. Maradudin, Ann. Phys. (USA), *22*, 223, 1963.
8. R. K. Singh and D. K. Neb, Phys. Stat. Sol., *(b)112*, 735, 1982.
9. A. A. Maradudin, Phys. Stat. Sol., *2*, 1493, 1962.
10. G. A. Alers and J. R. Neighbours, Revs Mod. Phys., *31*, 675, 1959.
11. J. A. Reissland, Physics of Phonons, Izd. Mir, Moscow, 1975, p. 151, (in Russian).
12. F. N. D. D. Pereira, C. H. Barnes and G. M. Graham, J. Appl. Phys., *41*, 5050, 1970.
13. W. C. Overton, Jn. and J. Gaffney, Phys. Rev., *98*, 969, 1955.
14. J. R. Neighbours and G. A. Alers, Phys. Rev., *111*, 707, 1958.
15. T. A. Hahn, J. Appl. Phys., *41*, 5096, 1970.
16. D. L. Martin, Canad. J. Phys., *38*, 17, 1960.
17. D. B. Fraser and A. C. Hollis Hallet, Canad. J. Phys., *43*, 193, 1965.
18. I. E. Leksina and S. I. Novikova, Fiz. Tverd. Tela, *5*, 1094, 1963.
19. P. F. Meads, W. R. Forsythe and W. F. Giauque, J. Am. Chem. Soc., *63*, 1902, 1941.
20. T. H. Geballe and W. F. Giauque, J. Am. Chem. Soc., *74*, 2368, 1952.
21. N. M. Plakida and T. Siklós, Phys. Stat. Sol., *33*, 103, 1969.

Acta Physica Hungarica 57 (1—2), pp. 73—77 (1985)

STUDY OF DEFECT MORPHOLOGY IN THE METAL AND DIELECTRIC LAYERS OF MIS STRUCTURE BY SIMS AND SEM*

P. I. DIDENKO, V. G. LITOVCHENKO, R. I. MARCHENKO and G. F. ROMANOVA

Institute of Semiconductors, Academy of Sciences of the Ukrainian SSR
252028 Kiev, USSR

and

M. NÉMETH-SALLAY, I. C. SZÉP and I. CSEH

Institute for Technical Physics, Hungarian Academy of Sciences
1325 Budapest, Hungary

(Received in revised form 28 July 1983)

The results of complex studies of Si-based MIS layer structures using SIMS with the secondary ion energy distribution analysis and SEM have been analysed. By these methods morphology-type defects in metal layers, the MI interface and the adjacent dielectric layer can be revealed. These structural-impurity complexes at the interfaces MI and IS caused by the impurity and phase-inclusions may stimulate phase-structural transformations in MIS systems subjected to heat treatment. The effect of the conditions under which real MIS structures were fabricated (temperature, method of metallization, orientation of SI substrate, conditions during oxidation) has been studied. The physical and chemical processes in MIS layers affecting the development of morphology-type defects in MIS structures are discussed.

The main feature of any MIS layer structure is the sequence of contacts between phases of different properties: the monocrystalline semiconductor, a submicron layer of the amorphous or glass-like dielectric, and a thin polycrystalline metal layer. The formation of the interface boundaries in such systems is determined by a number of factors, such as the conditions of growth of individual layers (growth rate, temperature, ambient atmosphere), the reactions during the growth especially at the beginning, the extent to which the interatomic distances of the adjacent phases match each other, the thermodynamic stability of the newly created phases, the dimensions of the adjacent layers, etc. Under certain conditions the interaction between the layers may lead to the generation (or annihilation) of many kinds of defects in each layer, these defects being complicated structural-impurity complexes.

This report presents the results obtained by secondary ion mass spectroscopy (SIMS) and scanning electron microscopy (SEM) on the formation of transition layers

* This paper was presented at the School on Defect Complexes in Semiconductor Structures, Mátrafüred, Hungary, Sept. 1982.

with defect structures in the vicinity of the Al—SiO_2 interface in MOS structures. Some results have already been presented elsewhere [1].

It follows from analysis of the thermodynamic data that the formation of the Al—SiO_2 interface is determined mainly by the chemical interaction between Al and SiO_2 [2, 3, 4]. Since the binding energy between aluminium and oxygen is higher than that between silicon and oxygen (1577.6 kJ/mole and 805.54 kJ/mole, respectively), the oxidation of the aluminium and the reduction of the silicon dioxide occur simultaneously and as a result an interlayer of Al_2O_3 is formed [3, 4]. It is assumed that this interlayer hinders the further diffusion of the Al atoms into the SiO_2 layer and the reaction stops.

A study of the Al—SiO_2—Si system with the help of the SIMS method has shown that the structure of the Al—SiO_2 interface is complex and cannot be explained by the presence of a thin Al_2O_3 interlayer alone. The Al^+ and Si^+ yield profiles were investigated during ion-etching as was the relationship between complex ions and the secondary ion energy spectra which reflect the changes in the surface potential of various parts of the surface, emitting different ions (Fig. 1). The most pronounced

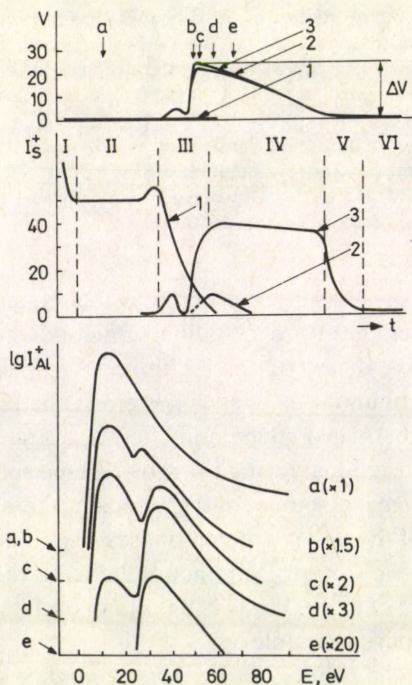


Fig. 1. Changes in external surface potential, secondary ion yield, and energy distributions as a result of the sputtering of Al—SiO_2—Si systems (thermally grown SiO_2, $d_{0x} = 150$ nm, Al deposited by thermal evaporation, $d_{Al} = 200$ nm). 1—external oxidized Al layer, 2—the main Al layer thickness, 3—transition layer between Al and SiO_2, 4—SiO_2 layer, 5—transition layer between SiO_2 and Si; a, b, c, d, e: the moments of Al^+ secondary ion energy distribution measurement

feature of this system—clearly expressed in all Al—SiO_2—Si systems studied—is the irregular morphology along the plane of the interface caused by the varying intensity of the reaction between Al and SiO_2 at different parts of the interface. By a subsequent application of SIMS, a selective etching of the surface takes place with the primary ions at the Al—SiO_2 interface and the appearance of uneven potential distribution on it, which in turn changes specifically the shape of Al^+ energy spectra, namely splits them into discrete maxima (Fig. 1).

The cross-section of the Al—SiO_2 layers etched wedge-like by an ion beam was also examined with the help of SEM. The existence of the uneven potential distribution could visually be observed and was of a peculiar lace-shape, built up of separate Al fragments on SiO_2 [1].

The following assumptions can be made about the nature of the irregular morphology of the Al—SiO_2 interface. The growth of a thin metal film (especially in the case of a good adhesion on amorphous substrate) begins with the formation of the nuclei at some particularly active surface centres. As the film substance accumulates around them, they partially coalesce and gradually greater grains are created. Up to thicknesses of about 10 nm the metal film is still discontinuous and the gaps between the grains are filled only after further deposition of Al. A similar process of Al film growth on quartz was observed experimentally in situ by electron microscope [5]. It is possible therefore to conclude that the interaction of the metal with the substrate at various stages of the deposition, when the nuclei are formed and when a thick layer grows, proceeds differently. However, it should be borne in mind that the process of film growth depends strongly on the presence of impurities both in the ambient (e.g. oxigen) and the metal. If a relatively weakly bound Al film (deposited on a cold substrate without subsequent annealing) is mechanically removed from the SiO_2 surface, the latter will still contain replicas of Al grains, created on it at the very beginning. Cases of the so-called columnar structures have also been reported [6], when the increase of the film thickness is accompanied by the replication of the nucleation network structure. In this case the growth proceeds mostly upward and the grains can remain separated for a long time.

The SIMS analysis was carried out on a set of samples prepared under various conditions, e.g. the Si substrate orientation was varied, the type of oxide ("dry", thermal, "HCl"), the substrate temperature at the metallization, etc. It was found that the development of the structural defects across the Al—SiO_2 interface depends very substantially on the conditions of its formation.

Indeed, the number of defects was less for "HCl" oxide or (100) orientation than for dry thermal oxide or (111) orientation. The structural inhomogeneity of the interface is also less in the case of metallization in good vacuum conditions.

Along with the development of the mosaic-like structure at the Al—SiO_2 interface some chemical reactions take place perpendicularly to the interface between Al and SiO_2. As is expected the oxygen goes over from SiO_2 to Al forming an Al_2O_3 layer at the Al—SiO_2 interface. Simultaneously Al atoms penetrate into the SiO_2 layer

creating aluminosilicate inclusions. The effective thickness of these inclusions is also determined by the conditions of MIS structure formation. The most important factors are the initial oxide structure (the interaction is stronger in the case of "HCl" oxides) and the thermal treatment. SIMS measurements show that at temperatures of 720—770 K the aluminosilicate layer thickness may reach about 50—80 nm. Two peculiarities must be noted:

1) The Al content within the aluminosilicate layer does not as a rule exceed 1—2% of the Si content and during the thermal treatments the layer broadens, but is not enriched with Al;

2) The Al penetration into SiO_2 does not disturb the dielectric properties of the layer, but sometimes even slightly improves some of them, as indicated by the measurements of the dielectric strength.

Finally, under the influence of Al, changes occur in the content of the SiO_2 itself, which can lead to the formation of new defects and to various phase-structural transformations of the oxide as a whole. According to SIMS data, metallization lowers the yield of SiO^+ secondary ions by about 20—25% across the whole thickness of the SiO_2 layer and by about 50% near the boundary with Al (Fig. 2). Simultaneously, the yield of $SiOH^+$ and SiH^+ group complex ions is also reduced; these are formed by

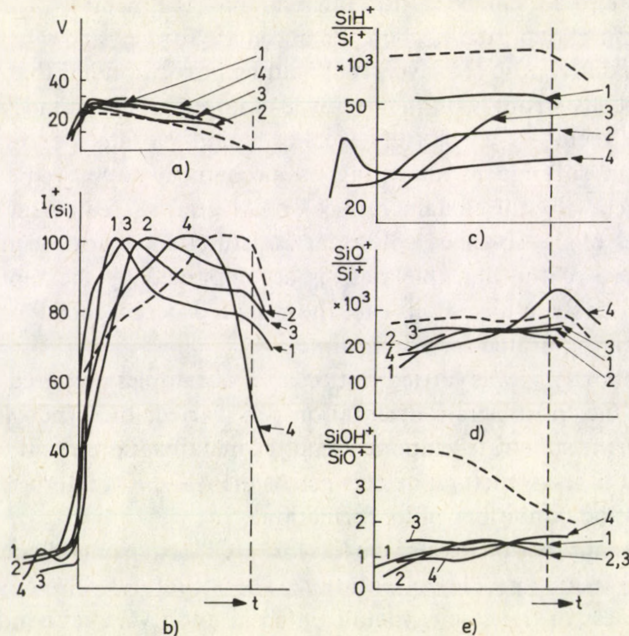


Fig. 2. Profiles of surface potential changes (a) and of secondary ion yield (b, c, d, e) during the sputtering of the oxide layer in the system Al—SiO_2—Si (substrate orientation (100), "HCl" SiO_2, Al deposited by thermal evaporation). 1—"cold" substrate, 2—"cold" substrate annealed in Ar atmosphere at 620 K for 30 min, 3—"hot" substrate, 4—"hot" substrate annealed in Ar atmosphere at 620 K for 30 min. Dotted line: the same values for the control sample

virtue of the presence of hydrogen or hydroxyls in SiO_2. Along with this the peak of $SiOH^+$ and SiH^+ yield is often observed at the actual SiO_2—Al interface. The changes in the yield of these ions and their profiles are also determined by the actual conditions of the MIS structure fabrication. For example, thermal treatments lead to smoother distribution of hydrogen containing ions across the SiO_2 layer thickness. However, the level of the yields in this case remains lower than for non-metallized samples.

The following assumptions can be made about the nature of the defects created in SiO_2 as a result of metallization. The removal of hydrogen and hydroxyls from SiO_2 must strengthen the silicon-oxygen skeleton, and lead to its definite ordering up to a formation of separate microcrystalline inclusions [7]. Subsequently, the breaking of the bonds at the interfaces between the microcrystallites and the amorphous matrix may occur as well as the appearance of non-bridging oxygen. However, a reduction in the relative yield of the secondary SiO^+/Si^+ ions below the level typical for stoichiometric SiO_2 ($18—20.10^{-3}$) indicates that due to the transition of the oxygen to aluminium, oxygen vacancies are created in the network of silicon–oxygen tetrahedra. It is possible that bonds of the Si—Si variety are created along with the vacancies. In that all of these defects are separated in space, they can probably exist simultaneously. If the penetration of Al into SiO_2 is associated with the replacement of Si by Al in the SiO_4 tetrahedron, then the difference in the ionic radii for Al and Si and the difference in their coordination numbers may also lead to the deformation of the whole initial structure of the SiO network and to the appearance of additional defects (namely modification of the coordination number and, probably, the trapping of hydrogen and hydroxyl groups near Al ions). Since the development of the structural defect complexes at the Al—SiO_2 interface may be connected with the differences in the degree of the reactions between Al and SiO_2 at different parts of the interface, it is natural to expect that different changes in SiO_2 take place at these parts.

Analysis of SIMS data has shown that the defects are impurity characteristics of SiO_2 surface regions where the nuclei were created, and they differed from those where the Al adhesion was weak. Thus, the irregular structure of the Al—SiO_2 interface induces a transformation of the morphology in the oxide layer. This property can be strengthened or weakened depending on the corresponding thermal treatments.

References

1. P. I. Didenko, V. G. Litovchenko, R. I. Marchenko, M. Németh-Sallay, G. F. Romanova and I. C. Szép, Ukrainian Phys. Journ., *26*, 1142, 1981.
2. D. M. Mattox, Thin Solid Films, *18*, 173, 1973.
3. K. Prabriputaloong and M. R. Pigot, J. Electrochem. Soc., *121*, 430, 1974.
4. Y. E. Strausser, E. J. Scheibner and J. S. Johannessen, Thin Solid Films, *52*, 203, 1978.
5. P. B. Barna, A. Barna and Z. Paál, Acta Phys. Hung., *49*, 77, 1980.
6. A. G. Dirks and H. J. Leamy, Thin Solid Films, *47*, 219, 1977.
7. V. G. Litovchenko and G. F. Romanova, Thin Solid Films, *81*, 27, 1981.

Acta Physica Hungarica 57 (1—2), pp. 79—98 (1985)

THE FOUNDATIONS OF NON-ASSOCIATIVE CLASSICAL FIELD THEORY

D. F. KURDGELAIDZE

Tbilisi State Medical Institute, Tbilisi, USSR

(Received 1 August 1983)

Non-associative classical field theory is set up on the basis of octonions. Algebra and projection operators of octonions, transformational properties of their elements regarding the Lorentz group are studied. Lorentz spinors and vectors have been determined through octonions. Lorentz invariants, Lagrangian function and equations of octonions field, their solutions, the Hamiltonian and spin vector have been obtained. A new non-associative quantum number, the associator, is introduced and its value for Lorentz spinors and tensors has been determined. Global and local symmetries of octonion fields equation and appropriate currents are studied, the problems of hydrogen atom and the possibilities of perturbation theory evolution are considered. It is indicated that the non-associative classical field theory in principle involves the basic results of Dirac theory along with new results.

Introduction

The application in physics of new methods based on mathematics not only widens the sphere of physical problems under question but can reveal new properties of matter as well. The transition from real numbers to complex ones allowed to describe the fields of both neutral and charged particles. The transition to quaternions (spinors) helped to introduce the notions of spin and subdivide physical fields into the so-called Bose and Fermi fields.

The complex numbers and quaternions may be obtained by means of the so-called "doubling". The doubling of quaternions produces octonions (Caley [1]). The complex numbers are commutative, quaternions are noncommutative, but both are associative, octonions are to be noncommutative and nonassociative ones. All these numbers form the so-called alternative algebrae (when two arbitrary elements give birth to associative algebra).

Hence, the next step for widening the scope of physical research methods can consist in the application of nonassociative algebra, octonions in particular, in physics. The algebra of numbers introduces in physics a new property connected with algebras nonassociativeness, the so-called "associator" — a value resembling spin. Due to it in the framework of alternative algebra there are now four basic types of fields: a) associative, commutative; b) associative, noncommutative; c) nonassociative, commutative; d) nonassociative, noncommutative. Octonions in elementary particles physics

were first used in the work [2] where the solution of the ordinary Dirac equation is written as octonion, its components being ordinary fields of fermions. The same idea was developed in the works [3] in which quarks were taken as Fermi fields in octonion form. The work [2] led to Hilbert space investigation with nonassociative multiplication [4]. Octonions have recently attracted the attention of physicists also in connection with elementary particles classification using exclusive Cartan groups, which are closely connected with octonions [5].

In contrast to the trends mentioned above the aim of this paper is to develop a consistent, classical, nonassociative field theory on octonion basis. The solution of this task requires the consideration of octonions algebra, projection operators, transformational properties of octonions in Lorentz transformations, spinors determination in nonassociative theory and the construction of respective invariants and Lagrange functions of octonions field, the solution of the octonion field equation, the investigation of global and local symmetries of these equations, of respective currents and charges, the investigation of Hamiltonian operators properties, of spin, of associator and the possibility of their application for particles description.

From the mathematical standpoint the new contribution consists not only in the construction of the nonassociative calculus but also in the fact that octonions have no matrix realization and the field theory as a whole should be developed on the basis of hypercomplex numbers.

§1. Octonion algebra and structure and projection operators

Octonion has 8 basic elements — e_0 unit ($e_0^2=1$) and elements — $e_1, e_2, \ldots, e_7$ satisfying the algebra:

$$\alpha, \beta, \gamma = 1, 2, 3, \ldots, 7; \qquad e_\alpha e_\beta = -\delta_{\alpha\beta} e_0 + E_{\alpha\beta\gamma} e_\gamma, \tag{1.1}$$

structural coefficients $E_{\alpha\beta\gamma}$ are completely antisymmetric

$$E_{123} = E_{561} = E_{624} = E_{435} = E_{471} = E_{572} = E_{673} = 1.$$

All others are equal to zero.

The nonassociativeness of algebra is characterized by the so-called associator

$$\Delta(e_\alpha e_\beta e_\gamma) = \frac{1}{2}\{(e_\alpha e_\beta)e_\gamma - e_\alpha(e_\beta e_\gamma)\}, \tag{1.2}$$

the triple multiplication is determined as

$$(e_\alpha e_\beta e_\gamma) = \frac{1}{2}\{(e_\alpha e_\beta)e_\gamma + e_\alpha(e_\beta e_\gamma)\}, \tag{1.3}$$

associator $\Delta(e_\alpha e_\beta e_\gamma)$ is completely antisymmetric.

In respect of these two operations 8 basic elements are divided into two subgroups, each consisting of four elements, quaternions-K for which

$$\Delta(e_\alpha e_\beta e_\gamma)=0, \qquad (e_\alpha e_\beta e_\gamma)\neq 0,$$

and as we will call them later associatorns-A for which we have

$$\Delta(e_\alpha e_\beta e_\gamma)\neq 0, \qquad (e_\alpha e_\beta e_\gamma)=0.$$

There are seven isomorphic quaternions and seven isomorphic associatorns.

Quaternions form a group relative to the triple multiplication

$$(e_n e_m e_k)=-E_{nmk}e_0+(-\delta_{nm}\delta_{kk'}+E_{nmm'}E_{m'kk'})e_{k'} \tag{1.4}$$

where n, m, k number three basic quaternion elements. Associatorn forms a group relative to associator

$$\Delta(e_\mu e_\nu e_\lambda)=E_{\mu\nu\lambda\rho}e_\rho, \tag{1.5}$$

$$E_{4567}=E_{2356}=E_{2347}=E_{1357}=E_{1346}=E_{1264}=E_{1245}=1.$$

All remaining $E_{\mu\nu\lambda\rho}$ are equal to zero. $E_{\mu\nu\lambda\rho}$ is antisymmetric. $(\mu, \nu, \lambda, \rho)$ number 4 basic elements of associatorn. Octonion Q can be presented as follows:

$$Q=K+A, \qquad K=\varphi_0 e_0+\varphi_n e_n, \qquad A=\varphi_\mu e_\mu; \tag{1.6}$$

$$n=1, 2, 3; \qquad \mu=4, 5, 6, 7.$$

The multiplication of quaternion by quaternion and associatorn by associatorn gives a quaternion. The multiplication of quaternion by associatorn gives associatorn. The conjugated octonion is

$$\bar{Q}=\varphi_0 e_0-\varphi_\alpha e_\alpha, \qquad \bar{Q}Q=\varphi_j^2 j(0,\alpha). \tag{1.7}$$

Hence, for conjugated quaternion and associatorn we have

$$\bar{K}=\varphi_0 e_0-\varphi_n e_n, \qquad \bar{A}=-\varphi_\mu e_\mu. \tag{1.8}$$

$\bar{Q}Q\equiv n(Q)$ is the norm of an octonion. If for three octonions we have the equation

$$Q_1 Q_2=Q_3, \tag{1.9}$$

then at $Q_1\neq 0$ the equation has the following solution

$$Q_2=\frac{1}{n(Q_1)}\bar{Q}_1 Q_3. \tag{1.10}$$

In order to isolate quaternion or associatorn parts from octonion new projection operators with the property $\hat{D}\hat{D}=\hat{D}$ are required. Any operator of the kind $\hat{D}_\alpha=\frac{1}{2}(e_0$

$+ie_\alpha)$, $\alpha=1, 2, \ldots 7$; is said to be projecting. Let us consider the influence of projection operator, $\hat{D}_7$, for example, upon octonion Q. We have in case of unidirectional action

$$(Q\hat{D}_7)=K(7)\hat{D}_7, \qquad (Q\hat{D}_7)=A(7)\hat{D}_7,$$
$$(\hat{D}_7Q)=\hat{D}_7\tilde{K}(7), \qquad (\hat{D}_7Q)=\hat{D}_7\tilde{A}(7);$$

where $Q=b_0e_0+b_\alpha e_\alpha$,

$$K(7)=(b_0-ib_7)e_0+(b_1+ib_4)e_1+(b_2+ib_5)e_2+(b_3+ib_6)e_3,$$
$$A(7)=(b_4-ib_1)e_4+(b_5-ib_2)e_5+(b_6-ib_3)e_6+(b_7+ib_0)e_7,$$
$$\tilde{K}(7)=(b_0-ib_7)e_0+(b_1-ib_4)e_1+(b_2-ib_5)e_2+(b_3-ib_6)e_3,$$
$$\tilde{A}(7)=(b_4+ib_1)e_4+(b_5+ib_2)e_5+(b_6+ib_3)e_6+(b_7+ib_0)e_7.$$

In the case of bidirectional action one we have

$$(\hat{D}_\alpha Q\hat{D}_\alpha)=K_0(\alpha)\hat{D}_\alpha, \qquad K_0(\alpha)=b_0-ib_\alpha. \tag{1.11}$$

§2. Octonion transformational properties

P.1. The specificity of nonassociative multiplication

In constructing nonassociative field theory one has to multiply systematically a certain number of basic elements obtaining an n-fold product of the kind $e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_n}$. After using the multiplication rules of nonassociative algebra similar products are reduced to the linear sum of primary basic elements. However, one can obtain different results depending on the way of performing this transition. In general, the product $e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_n}$ can be represented in different ways [4]:

$$e_\alpha e_\beta e_\gamma Q=\begin{cases}(e_\alpha(e_\beta(e_\gamma Q))),\\ (e_\alpha e_\beta e_\gamma Q),\\ ((e_\alpha e_\beta e_\gamma)Q),\end{cases} \tag{2.1}$$

where brackets (...) mean the transition to the main basis. This circumstance necessitates the introduction of the notion of "closed" and "open" products. Later on the product will be called "closed" and designated as $(e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_n})$; if the procedure of reducing multiple product to the main basis is performed only at these elements, no matter whether they are used separately or as comultipliers of a more complex product. The product will be called "open" and designated by $e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_n}$ if the procedure of reducing multiple product to the main basis is not applied. Both "closed" and "open" products allow two operations common for them: a) the inversion of comultipliers on

condition that $e_\alpha e_\beta = -e_\beta e_\alpha$, $\alpha \neq \beta$; b) the elimination of similar elements $e_\alpha e_\alpha = -1$ (it is not summed on α). In creating nonassociative field theory (unless otherwise specified) we will work with the "open" multiplication.

P.2. Octonion with "open" multiplication

Octonions with "open" multiplication have two evident subalgebras: 1) the first generated by quaternion basis, for example, $\{e_0, e_1, e_2, e_3, e_1e_2, e_1e_3, e_2e_4, e_1e_2e_3\}$ algebra possesses eight elements; 2) the second generated by associatorn basis $\{e_{\mu_1}, e_{\mu_1}e_{\mu_2}, e_{\mu_1}e_{\mu_2}e_{\mu_3}, e_{\mu_1}e_{\mu_2}e_{\mu_3}e_{\mu_4}\}$, $\mu = 4, 5, 6, 7$. This algebra has 16 basic elements (after the addition of the unit e_0) and in the framework of "open" multiplication it is isomorphic to algebra $\{\gamma_\mu\}$ of Dirac matrices. If $e_8 = e_4e_5e_6e_7$ is introduced, the general element of algebra can be represented in the form of "nonassociative undor"

$$Y_Q = \varphi_0 e_0 + \varphi_8 e_8 + \varphi_\mu e_\mu + \varphi_{\mu 8} e_\mu e_8 + \varphi_{\mu\nu} E_{\mu\nu}, \tag{2.2}$$

where

$$E_{\mu\nu} = \frac{1}{2}(e_\mu e_\nu - e_\nu e_\mu); \qquad \mu, \nu = 4, 5, 6, 7. \tag{2.3}$$

If in Y_Q one proceeds to the closed multiplication, octonion Q is obtained. In the general case of octonions of "open" multiplication the main basis elements generate algebra with $16 \times 8 = 128$ basic elements. The general element of this algebra is

$$T = \sum_{a=1}^{128} \Phi_a \Gamma_a, \qquad \Gamma_1 = e_0, \qquad \Gamma_2 = e_1, \ldots, \Gamma_{128} = e, \ldots e_7. \tag{2.4}$$

The application of octonion with "open" multiplication has the following advantages: first, associator can be explicitly introduced into the theory; secondly, in the octonion field theory one has to work with the projection operators which allow to obtain (at final stages) the numbers without transition to the main basis into multiple products. The specificity of nonassociative algebra is revealed in the specificity of projection operators.

P.3. Vectors in octonion calculus

A quantity X is called vector if in the transformations of coordinates $x'_\mu = L_{\mu\nu} x_\nu$, X is transformed according to the law $X' = SXS^+$, $SS^{-1} = 1$, where S are transformation operators. Then at $S^+ = S^{-1}$ we have rotation, at $S^+ = S$ Lorentz transformation. As octonion contains quaternion as a subgroup, the three-dimensional vector and its rotation can always be given on a quaternion basis. Giving the four dimensional vector in the form of

$$X = x_\mu e_{\alpha\mu}, \qquad X^2 = -x_\mu^2, \qquad x_\mu(x_n, it), \qquad n = 1, 2, 3,$$

and taking into account octonion properties one finds that X can be an associatorn only.

$$X = x_\mu e_\mu, \qquad \mu = 4, 5, 6, 7,$$

$$x_4 \Rightarrow x_1 = x, \qquad x_5 \Rightarrow x_2 = y, \qquad x_6 \Rightarrow x_3 = z, \qquad x_7 \Rightarrow x_4 = it. \tag{2.5}$$

Then S_4 can be both associatorn and quaternion.

A) in the case when S_4 is associatorn, i.e. $S_4 = \alpha_\mu e_\mu$, considering the transformation as "closed" product, one finds

$$x'_\nu = x_\mu \omega_{\mu\nu}, \qquad e'_\lambda = \omega_{\lambda\mu} e_\mu, \tag{2.6}$$

where

$$X' = (S_4 X S_4^{-1}) = x_\lambda e'_\lambda = x'_\nu e_\nu,$$

$$\omega_{\mu\beta} = \frac{1}{2} \{(\delta_{\mu\lambda'} \delta_{\lambda\beta} + \delta_{\mu\lambda} \delta_{\beta\lambda'}) - (E_{\gamma\lambda'\mu} E_{\gamma\lambda\beta} + E_{\gamma\lambda'\beta} E_{\gamma\lambda\mu})\} \alpha_\lambda \alpha_{\lambda'}. \tag{2.7}$$

It can be shown that $\omega_{\mu n} = 0$, $n = 1, 2, 3$.

In particular, at $\alpha_5 = \alpha_6 = 0$ one obtains the two dimensional vector transformation

$$X_2 = x_4 e_4 + x_7 e_7, \qquad S_t = \alpha_4 e_4 + \alpha_7 e_7, \qquad S_t = -S_t. \tag{2.8}$$

Then $S_4^a = e_4(S_t; \bar{S}_4^a = S_t)e_4$, i.e. the transformation

$$X' = e_4(S_t X \bar{S}_t)e_4,$$

$$\alpha_4^2 + \alpha_7^2 = 1, \tag{2.9}$$

gives a certain Lorentz transformation. Likewise, at $\alpha_7 = 0$ the three-dimensional vector transformation is obtained.

B) The case when S_4^b is a quaternion, i.e. $S_4^b \equiv K = \alpha_0 e_0 + \alpha_n e_n$, it is found

$$X' = (K X \bar{K}) = x'_\nu e_\nu = x_\mu e'_\mu,$$

$$x'_\nu = x_\mu \omega_{\mu\nu}, \qquad e'_\mu = \omega_{\mu\nu} e_\nu, \tag{2.10}$$

$$\omega_{\mu\nu} = \alpha_0^2 \delta_{\mu\nu} - 2\alpha_0 \alpha_n E_{n\mu\nu} - (\alpha_n E_{n\lambda\mu})(\alpha_m E_{m\lambda\nu}).$$

At $\alpha_0 = 0$, we have $\omega_{\mu\nu} = -\delta_{\mu\nu}$.

Thus we have

$$(S_4^a Q \bar{S}_4^a) = K + A',$$

$$(S_4^b Q \bar{S}_4^b) = K' + A, \qquad \alpha_0 = 0, \tag{2.11}$$

i.e. by the action of S_4^a on octonion only the associatorn part is transformed and by the action of $S_4^b(\alpha_0 = 0)$ only the quaternion part of octonion is changed.

P.4. Transformational properties of octonion in the case of "open" multiplication

Let us consider only transformation properties of nonassociative undor Y_Q. As it was mentioned above, the algebra of associatorn with the "open" multiplication is in fact isomorphic with Dirac algebra. Hence $\frac{1}{2}E_{\mu\nu}$ satisfies the commutation relations of Lorentz group momenta

$$-\frac{1}{2}[E_{\mu\nu}E_{\lambda\rho}]=\delta_{\mu\rho}E_{\nu\lambda}-\delta_{\nu\rho}E_{\mu\lambda}-\delta_{\mu\lambda}E_{\nu\rho}+\delta_{\nu\lambda}E_{\mu\rho}. \tag{2.12}$$

Then nonassociative undor components possess ordinary tensor dimensions, namely, φ_0 – scalar, φ_μ– vector, $\varphi_{\mu 8}$ – pseudovector, and $\varphi_{\mu\nu}$ – antisymmetric tensor of the second rank, φ_8 – pseudoscalar.

P.5. Spinors in nonassociative algebra

The quantities Ψ, $\bar{\Psi}$, which by the transformation of coordinates according to the law $x'_\mu=L_{\mu\nu}x_\nu$ and vectors X according to the law $X'=(SXS^{-1})$, are changed into

$$\Psi'=S\Psi, \qquad \bar{\Psi}'=\bar{\Psi}S^{-1}, \qquad SS^{-1}=1; \tag{2.13}$$

are called spinors. Then $\bar{\Psi}'\Psi'=\bar{\Psi}\Psi=\text{inv}$.

A) In the case of the algebra of associatorns $\{e_\mu\}$ with the "open" multiplication, spinors can be constructed on the basis of nonassociative undor just in the same way as they are constructed from undors in Dirac algebra [6]; we obtain

$$\begin{aligned}\Psi&=(\Psi_0e_0+\Psi_1e_4+\Psi_2e_6+\Psi_3e_4e_6)\hat{D}_7\hat{D}_{45},\\ \Psi^+&=\hat{D}_{45}\hat{D}_7(\Psi_0^*e_0+\Psi_3^*e_6e_4-\Psi_1^*e_4-\Psi_2^*e_6),\\ \bar{\Psi}&=i\Psi^+e_7=\hat{D}_{45}\hat{D}_7(\Psi_0^*e_0+\Psi_1^*e_4+\Psi_2^*e_6+\Psi_3^*e_6e_4),\\ e_\alpha^+&=-e_\alpha, \qquad (e_\alpha e_\beta)^+=e_\beta^+e_\alpha^+=-e_\alpha e_\beta, \qquad \alpha\neq\beta.\end{aligned} \tag{2.14}$$

Ψ_n-components are linearly expressed by nonassociative undor components. Hence $\bar{\Psi}\Psi=\text{inv}$, $\bar{\Psi}e_\mu\Psi$ vector, in particular,

$$\begin{aligned}\bar{\Psi}\Psi&=\hat{D}_{45}\hat{D}_7(\Psi_0^*\Psi_0-\Psi_1^*\Psi_1-\Psi_2^*\Psi_2+\Psi_3^*\Psi_3)\hat{D}_7\hat{D}_{45},\\ \bar{\Psi}e_7\Psi&=-i\Psi^+\Psi=\hat{D}_{45}\hat{D}_7(\Psi_0^*\Psi_0+\Psi_1^*\Psi_1+\Psi_2^*\Psi_2+\Psi_3^*\Psi_3)\hat{D}_7\hat{D}_{45}.\end{aligned} \tag{2.15}$$

B) To obtain the spinors from octonions is a more complicated problem. Let us consider octonion

$$Q=\varphi_je_j=\varphi_0e_0+\varphi_ne_n. \tag{2.16}$$

Let us act on it by projection operator;

$$\hat{\Pi}_c=\hat{D}_{14}\hat{D}_{25}\hat{D}_{36}\hat{D}_{70}, \qquad \hat{D}_{\alpha\beta}=\frac{1}{2}(1+ie_\alpha e_\beta). \tag{2.17}$$

Then we obtain

$$\Psi = Q\hat{\Pi}_c = (\Psi_1 e_4 + \Psi_2 e_5 + \Psi_3 e_6 + \Psi_4 e_7)\hat{\Pi}_c,$$
$$\Psi_1 = \varphi_4 - i\varphi_1, \qquad \Psi_2 = \varphi_5 - i\varphi_2, \qquad \Psi_3 = \varphi_6 - i\varphi_3, \qquad \Psi_4 = \varphi_7 - i\varphi_0. \tag{2.18}$$

Thus

$$\Psi^+ = -\hat{\Pi}_c(\Psi_1^* e_4 + \Psi_2^* e_5 + \Psi_3^* e_6 + \Psi_4^* e_7). \tag{2.19}$$

Let us determine the conjugated spinor by relations

$$\bar{\Psi} = \Psi^+ T, \qquad \bar{\Psi}\Psi = \text{inv}. \tag{2.20}$$

Then we find

$$T_1 = e_7 e_1 e_4, \qquad T_2 = e_7 e_2 e_5, \qquad T_3 = e_7 e_3 e_6, \qquad T_1^2 = T_2^2 = T_3^2 = 1. \tag{2.21}$$

Invariants will be

$$\bar{\Psi}^{(n)}\Psi = \Psi^+ T_n \Psi, \quad n = 1, 2, 3; \qquad \Psi^+ T_1 \Psi, \quad \Psi^+ T_2 \Psi, \quad \Psi^+ T_3 \Psi. \tag{2.22}$$

Vector will be

$$-\bar{\Psi}^{(n)} T_n e_7 e_\mu \Psi = -\Psi^+ T_n T_n e_7 e_\mu \Psi = -\Psi^+ e_7 e_\mu \Psi. \tag{2.23}$$

In particular, for density we have $-\bar{\Psi}^{(n)} T_n e_7 e_7 \Psi = \Psi^+ \Psi$. Spinor $\bar{\Psi}^{(n)}$ is essentially a nonassociative spinor. In Dirac theory there is no such 4-component spinor from the viewpoint of undor analogue calculus.

§3. Octonion field equations

P.1. Lagrangian of the octonion field

In constructing bilinear invariants of the octonion field it is necessary to know the transformation properties of the function. In particular, if we consider octonion spinors, then, as has been shown, it is necessary to consider two cases:

A) associatorn with "open" multiplication, when

$$\Psi = (\Psi_0 e_0 + \Psi_1 e_4 + \Psi_2 e_6 + \Psi_3 e_4 e_6)\hat{D}_7 \hat{D}_{45}, \; \bar{\Psi} = -i\Psi^+ e_7. \tag{3.1}$$

Then the invariants are

$$I_0 = \bar{\Psi}\Psi, \qquad I_1 = \bar{\Psi} e_\mu P_\mu \Psi, \qquad P_\mu = -i\frac{\partial}{\partial x_\mu}. \tag{3.2}$$

Then the Lagrangian of the free octonion field will be given by

$$L_a = -\frac{1}{2}\left\{\bar{\Psi} e_\mu \frac{\partial \Psi}{\partial x_\mu} - \frac{\partial \bar{\Psi}}{\partial x_\mu} e_\mu \Psi + 2iK_0 \bar{\Psi}\Psi\right\}, \qquad K_0 = \text{const}. \tag{3.3}$$

Thus we obtain for the field equations

$$e_\mu \frac{\partial \Psi}{\partial x_\mu} + iK_0 \Psi = 0, \qquad \frac{\partial \bar{\Psi}}{\partial x_\mu} e_\mu - iK_0 \bar{\Psi} = 0. \tag{3.4}$$

B) The case of essential nonassociative spinor, when

$$\Psi = Q\hat{\Pi}_c, \qquad \bar{\Psi}^{(n)} = \Psi^+ T_n,$$

then the invariants will be

$$I_0 = \bar{\Psi}^{(n)} \Psi = \Psi^+ T_n \Psi, \qquad I_2 = -\bar{\Psi}^{(n)} T_n e_7 e_\mu P_\mu \Psi = -\Psi^+ e_7 e_\mu P_\mu \Psi. \tag{3.5}$$

We have for the Lagrangian function

$$\begin{aligned} L_b^{(n)} &= -\frac{1}{2}\left\{ \bar{\Psi}^{(n)} T_n e_7 e_\mu \frac{\partial \Psi}{\partial x_\mu} - \frac{\partial \bar{\Psi}}{\partial x_\mu} T_n e_7 e_\mu \Psi + 2iK_0 \bar{\Psi}^{(n)} \Psi \right\} = \\ &= -\frac{1}{2}\left\{ \Psi^+ e_7 e_\mu \frac{\partial \Psi}{\partial x_\mu} - \frac{\partial \Psi^+}{\partial x_\mu} e_7 e_\mu \Psi + 2i\Psi^+ T_n \Psi \right\}. \end{aligned} \tag{3.6}$$

Then we find for field equations

$$e_7 e_\mu \frac{\partial \Psi}{\partial x_\mu} + iK_0 T_n \Psi = 0, \qquad \frac{\partial \Psi^+}{\partial x_\mu} e_7 e_\mu - iK_0 \Psi^+ T = 0. \tag{3.7}$$

Hence, we get two versions of the theory: a theory with Lagrangian L_a and a theory with Lagrangian $L_b^{(n)}$.

In this paper we will consider the theory with Lagrangian L_a.

P.2. Operators of energy (Hamiltonian), of spin and associator

From the field equation with Lagrange function L_a for the energy operator we find

$$\begin{gathered} \hat{H} = -i(\xi \nabla) - iK_0 e_7, \\ \xi_1 = ie_4 e_7, \qquad \xi_2 = ie_5 e_7, \qquad \xi_3 = ie_6 e_7. \end{gathered} \tag{3.8}$$

Then it can be written as

$$\left(i\frac{\partial}{\partial t} - \hat{H} \right) \Psi = 0, \qquad \frac{\partial}{\partial x_7} \Rightarrow -i\frac{\partial}{\partial t}, \qquad x_7 \Rightarrow it.$$

For the energy-momentum tensor we obtain:

$$\begin{gathered} T_{\mu\nu} = \frac{1}{2}\left(\frac{\partial \bar{\Psi}}{\partial x_\mu} e_\nu \Psi - \bar{\Psi} e_\nu \frac{\partial \Psi}{\partial x_\mu} \right), \qquad \mu, \nu = 4, 5, 6, 7. \\ x_4 \Rightarrow x_1, \qquad x_5 \Rightarrow x_2, \qquad x_6 \Rightarrow x_3, \qquad x_7 \Rightarrow x_4 = it, \end{gathered} \tag{3.9}$$

$T_{\mu\nu}$ is nonsymmetric. By symmetrization we find

$$Q_{\mu\nu}=T_{\mu\nu}+\eta_{\mu\nu}, \qquad \eta_{\mu\nu}=\frac{\partial}{\partial x_\lambda} f_{\mu[\nu\lambda]},$$
$$f_{\mu[\nu\lambda]}=\frac{1}{2}(\bar{\Psi}e_\mu e_\nu e_\lambda \Psi). \tag{3.10}$$

Thus for spin tensor we have:

$$S_{[\mu\nu]\lambda}=f_{\mu[\nu\lambda]}-f_{\nu[\mu\lambda]}=-\frac{1}{2}\bar{\Psi}e_\mu e_\nu e_\lambda \Psi.$$

Then we have for spin vector

$$S_\rho=-\frac{i}{6}E_{\mu\nu\lambda\rho}\frac{1}{2}\bar{\Psi}e_\mu e_\nu e_\lambda\Psi, \qquad S_\mu=\frac{1}{2}\bar{\Psi}\sigma_\mu\Psi,$$
$$\sigma_4=ie_5e_6\Rightarrow\sigma_1, \qquad \sigma_5=ie_6e_4\Rightarrow\sigma_2, \tag{3.11}$$
$$\sigma_6=ie_4e_5\Rightarrow\sigma_3, \qquad \sigma_7=-ie_4e_5e_6e_7=-ie_8.$$

In an analogous way associator is introduced

$$C_\rho=-\frac{i}{6}E_{\mu\nu\lambda\rho}\frac{1}{2}\bar{\Psi}\Delta(e_\mu e_\nu e_\lambda)\Psi, \qquad C_\mu=\frac{1}{2}\bar{\Psi}\xi_\mu\Psi,$$
$$\xi_4=ie_4e_7\Rightarrow\xi_1, \qquad \xi_5=ie_5e_7\Rightarrow\xi_2, \tag{3.12}$$
$$\xi_6=ie_6e_7\Rightarrow\xi_3, \qquad \xi_7=-i\Rightarrow\xi_4.$$

Hamiltonian $\hat{H}$ can now be written as.

$$\hat{H}=i\{e_8(\boldsymbol{\sigma}\nabla)-K_0e_7\}, \qquad \boldsymbol{\xi}=-e_8\boldsymbol{\sigma}.$$

P.3. Algebra of spin and associator operators

In case of "opened" multiplication, as mentioned above, the values $E_{\mu\nu}$ satisfy commutation relations of Lorentz group momenta $I_{\mu\nu}$. If a three-dimensional spin vector is introduced, as usual, [7]

$$\sigma_l=-\frac{i}{4}E_{lmn}E_{nm} \tag{3.13}$$

(a change is introduced in the designation of indices: $4\to1$, $5\to2$, $6\to3$) we obtain

$$\sigma_4=ie_5e_6\Rightarrow\sigma_1, \qquad \sigma_5=ie_6e_4\Rightarrow\sigma_2, \qquad \sigma_6=ie_4e_5\Rightarrow\sigma_3. \tag{3.14}$$

By the analogy a three-dimensional associator vector is introduced

$$\xi_n=-\frac{i}{2}E_{7n}$$

(A change is also introduced in designations: 4→1, 5→2, 6→3). Then we find from the momenta algebra

$$[\sigma_n \sigma_m] = iE_{nml}\sigma_l, \qquad [\xi_n \xi_m] = iE_{nml}\sigma_l,$$
$$[\xi_n \sigma_m] = iE_{nml}\xi_l. \tag{3.15}$$

Now new operators are introduced in a convective way

$$X_n = \frac{1}{2}(\sigma_n + \xi_n), \qquad Y_n = \frac{1}{2}(\sigma_n - \xi_n) \tag{3.16}$$

and we obtain

$$[X_n X_m] = iE_{nml}X_l, \quad [Y_n Y_m] = iE_{nml}Y_l, \quad [X_n Y_m] = 0. \tag{3.17}$$

Table I

	Scalar	Vector	Symmetric tensor of the 2nd rank	Antisymmetric tensor of the 2nd rank	Dirac spinors
S_3^0	0	1	2	1	1/2
C_3^0	0	0	0	1	+1/2
j	0	1/2	1	1 0	1/2 0
K	0	1/2	1	0 1	0 1/2

Then the original algebra of Lorentz generators can be given as a sum of two isomorphic algebras for the projection of quantitative motion momenta. The respective representations are classified according to proper values [7]

$$(X_3, Y_3) \Rightarrow (m_j, m_k), \qquad \begin{array}{l} m_j = -j, -j+1, \ldots\ 0, \ldots, j; \\ m_k = -k, -k+1, \ldots\ 0, \ldots, k. \end{array} \tag{3.18}$$

The representations are generally characterized by a pair of numbers $\tau \simeq (j, k)$. Here $2j = |S_3^0 + C_3^0|$, $2k = |S_3^0 - C_3^0|$,

$$S_3^0 = |j+k|, \qquad C_3^0 = |j-k|, \qquad \tau \sim (j, k) \sim (S_3^0, C_3^0); \tag{3.19}$$

where S_3^0 and C_3^0 are maximum values of the third component projection of spin and associator, respectively.

Hence, for different tensors we have (see Table I).

P.4. Spin and associator equations and their solutions

Let us introduce the differential operator of spin and associator.

$$\hat{S} = \frac{(\boldsymbol{\sigma}\nabla)}{ik}, \qquad \hat{C} = \frac{(\boldsymbol{\xi}\nabla)}{ik}. \tag{3.20}$$

Taking into account that $\xi_n = -e_8\sigma_n$, we find $[\hat{S}\hat{C}]=0$ the corresponding equations and solutions will be; for spin

$$\left(\frac{(\sigma\nabla)}{ik} - S\right)\Psi_k = 0, \qquad \Psi_k = \sum_{\beta\varkappa}^{\pm} \hat{D}^{\beta}_{\sigma_3} \frac{1}{2}(\beta + \varkappa S)\Psi^{\varkappa}_{\beta},$$
$$\hat{D}^{\beta}_{\sigma_3} = \frac{1}{2}(1+\beta\sigma_3) = \frac{1}{2}(1+i\beta e_4 e_5); \tag{3.21}$$

for associator

$$\left(\frac{(\xi\nabla)}{ik} - C\right)\Psi_k = 0, \qquad \Psi_k = \sum_{\delta\varkappa}^{\pm} \hat{D}^{\delta}_{\xi_3} \frac{1}{2}(\delta + \varkappa C)\Psi^{\varkappa}_{\delta},$$
$$\hat{D}^{\delta}_{\xi_3} = \frac{1}{2}(1+\delta\xi_3) = \frac{1}{2}(1+i\delta e_6 e_7). \tag{3.22}$$

Due to operators commutativity $\hat{S}$ and $\hat{C}$ the combined solution of these equations will be

$$\Psi_k = \sum_{\varkappa,\beta,\delta}^{\pm} \hat{D}^{\beta}_{\sigma_3}\hat{D}^{\delta}_{\xi_3} \frac{1}{2}(\beta + \varkappa S)\frac{1}{2}(\delta + \varkappa C)\Psi^{\varkappa}_{(\beta,\delta)}. \tag{3.23}$$

At $K_0 = 0$, i.e. for massless fields $\hat{H}$, $\hat{S}$ and $\hat{C}$ commute

$$[\hat{H}\hat{C}] = [\hat{H}\hat{S}] = [\hat{S}\hat{C}] = 0, \qquad K_0 = 0; \tag{3.24}$$

i.e. the state in this case can be simultaneously characterized by the quantum numbers of energy, spin and associator. At $K_0 \neq 0$, i.e. for the fields having rest mass

$$[\hat{H}\hat{S}] = [\hat{S}\hat{C}] = 0, \qquad [\hat{H}\hat{C}] \neq 0, \qquad K_0 \neq 0, \tag{3.25}$$

i.e. the fields with rest mass can be characterised only by two quantum numbers either by energy and spin or by associator and spin.

P.5. Hamiltonian equation and its solution

The Hamiltonian equation

$$\left(i\frac{\partial}{\partial t} - \hat{H}\right)\Psi = \left\{i\frac{\partial}{\partial t} + i((\xi\nabla) + K_0 e_7)\right\}\Psi = 0. \tag{3.26}$$

In case of the plane wave $K_x = K_y = 0$, we have

$$(\underline{E} - K\xi_3 + i\varkappa K_0 e_7)\Psi^{\varkappa}_k = 0, \qquad K_z = K$$
$$\Psi = \sum_{\varkappa}^{\pm} \Psi^{\varkappa}(K)\exp(i\varkappa(Kz - \underline{E}t)) \tag{3.27}$$

with the solution

$$\Psi^{\varkappa}(K)=\sum_{\varepsilon}^{\pm}\left\{\left(\hat{\Theta}_3-\frac{iK}{\varepsilon\underline{E}_0+\varkappa K_0}\hat{\Theta}_1\right)\Psi_3^{\varkappa}(\varepsilon)+\left(\hat{\Theta}_4+\frac{iK}{\varepsilon\underline{E}_0+\varkappa K_0}\hat{\Theta}_2\right)\Psi_4^{\varkappa}(\varepsilon)\right\},$$

$$\hat{\Theta}(\gamma\delta)\equiv\hat{D}_7^{\gamma}\hat{D}_{\xi_3}^{\delta},\qquad \hat{D}_7^{\gamma}=\frac{1}{2}(1+i\gamma e_7),\qquad \underline{E}=\varepsilon\underline{E}_0, \tag{3.28}$$

$$\hat{\Theta}_1\equiv\hat{\Theta}(++),\qquad \hat{\Theta}_2=\hat{\Theta}(+-),\qquad \hat{\Theta}_3=\hat{\Theta}(-+),\qquad \hat{\Theta}_4=\hat{\Theta}(--),$$

$$\underline{E}_0=\sqrt{K^2+K_0^2}.$$

The solution of the combined Hamiltonian and spin equations

$$(\underline{E}+e_8\varkappa SK+ie_7\varkappa K_0)\Psi^{\varkappa}(K)=0,\qquad (\sigma_3-\varkappa S)\Psi^{\varkappa}(K)=0; \tag{3.29}$$

is

$$\Psi^{\varkappa}=\sum_{\beta\delta}^{\pm}\left\{\hat{\Pi}(\alpha\beta\gamma)-\frac{\underline{E}+\alpha\beta K}{\gamma\varkappa K_0}\hat{\Pi}(-\alpha\beta\gamma)\right\}\Psi_{(\alpha\beta\gamma)}^{\varkappa}\,\frac{1}{2}(\beta+\varkappa S),$$
$$\hat{\Pi}(\alpha\beta\gamma)=\hat{D}_{e_8}^{\alpha}\hat{D}_{\sigma_3}^{\beta}\hat{D}_7^{\gamma}. \tag{3.30}$$

In the case $K_0=0$ the solution of the equations

$$\{i(\xi\nabla)+\partial_7\}\Psi_k=0,\qquad \left(\frac{(\sigma\nabla)}{ik}-S\right)\Psi_k=0,\qquad \left(\frac{(\xi\nabla)}{ik}-C\right)\Psi_k=0. \tag{3.31}$$

is

$$\Psi_{(k)}=\sum_{\beta\delta\varkappa}^{\pm}\hat{D}_{\sigma_3}^{\beta}\hat{D}_{\xi_3}^{\gamma}\frac{1}{2}(\beta+\varkappa S)\frac{1}{2}(\delta+\varkappa C)\frac{1}{2}(\delta-\varepsilon)\Psi_{(\varepsilon,\beta,\delta,\gamma)}^{\varkappa} \tag{3.32}$$

P.6. Nonassociative Bose field

As shown above, the associator $C\neq 0$ for spinor field and the field of the antisymmetric tensor of the second rank. Giving the octonion equation does not define the type of the field. For this it is necessary also to indicate the transformation properties of the solution of equation.

In case of octonion Bose field the solution of the equation

$$\left(e_\mu\frac{\partial}{\partial x_\mu}+iK_0\right)\Psi=0,\qquad \bar{\Psi}\left(e_\mu\frac{\partial}{\partial x_\mu}-iK_0\right)=0 \tag{3.33}$$

should be sought in the form:

$$\Psi=\left(e_\mu\frac{\partial}{\partial x_\mu}-iK_0\right)U,\qquad \bar{\Psi}=\bar{U}\left(e_\mu\frac{\partial}{\partial x_\mu}+iK_0\right), \tag{3.34}$$

and we obtain

$$\left(-\frac{\partial^2}{\partial x_\mu^2}+K_0^2\right)U=0. \tag{3.35}$$

If the solution is sought for in the general form

$$\Psi = \varphi_0 e_0 + \varphi_\alpha e_\alpha, \qquad \bar{\Psi} = \bar{\varphi}_0 e_0 - \varphi_\alpha e_\alpha. \tag{3.36}$$

After its substitution in the equation and the transition to the main basis (by means of transition to the "closed" multiplication) it is obtained

$$\begin{gathered} \nabla \Phi_4 + i \frac{\partial \boldsymbol{\varphi}}{\partial t} - \text{rot} \quad = iK_0 \boldsymbol{\varphi}, \\ \boldsymbol{\varphi}(\varphi_1, \varphi_2, \varphi_3), \qquad \boldsymbol{\varphi}(\varphi_4, \varphi_5, \varphi_6), \qquad \Phi_4 \equiv \varphi_7. \end{gathered} \tag{3.37}$$

If we introduce

$$\mathbf{E} = i \frac{\partial \boldsymbol{\varphi}}{\partial t} + \nabla \Phi_4, \qquad \mathbf{H} = \text{rot}\, \boldsymbol{\varphi},$$

we find

$$\mathbf{E} + \mathbf{B} = iK_0 \boldsymbol{\varphi}, \qquad \varphi_0 = \frac{1}{iK_0} \partial_\mu \varphi_\mu. \tag{3.38}$$

At $\boldsymbol{\varphi} = 0$ we have

$$\Psi = \left(e_0 - \frac{1}{iK_0} e_\mu \partial_\mu \right) \varphi_0, \qquad \bar{\Psi} = \bar{\varphi}_0 \left(e_0 + \frac{1}{iK_0} e_\mu \partial_\mu \right). \tag{3.39}$$

One can introduce new hypercomplex numbers, generalizing Kemmer—Duffin matrices for nonassociative algebra. For this consider two independent (commutational) associator bases e_j^a, e_j^b, $j = 0, 1 \ldots 7$ and introduce a new basis

$$\begin{gathered} \Lambda_j = \frac{1}{2}(e_j^a + e_j^b), \qquad \Lambda_0 = \frac{1}{2}(e_0^a + e_0^b), \qquad \Lambda_\alpha = \frac{1}{2}(e_\alpha^a + e_\alpha^b), \\ 2(\Lambda_\alpha \Lambda_\alpha) = -1 + U_\alpha, \qquad U_\alpha = e_\alpha^a e_\alpha^b, \qquad \Lambda_\alpha^3 = -\Lambda_\alpha, \\ (\Lambda_\alpha \Lambda_\beta) = -U_\beta \Lambda_\alpha, \qquad (\Lambda_\alpha \Lambda_\beta \Lambda_\gamma) + (\Lambda_\gamma \Lambda_\beta \Lambda_\alpha) = \Phi_{\alpha\beta\gamma}, \\ \dot{\Phi}_{\alpha\beta\gamma} = -6(\delta_{\alpha\beta} \Lambda_\gamma + \delta_{\gamma\beta} \Lambda_\alpha) - 4\delta_{\alpha\beta} \Lambda_\beta + 2(E_{\alpha\beta\gamma} E_{\gamma\lambda\delta} + E_{\gamma\beta\delta} E_{\alpha\lambda\delta}) \Lambda_\lambda. \end{gathered} \tag{3.40}$$

Then the equation for octonion Bose field can be written as

$$\left(\Lambda_\mu \frac{\partial}{\partial x_\mu} + iM_0 \right) \Psi = 0, \qquad \bar{\Psi} \left(\Lambda_\mu \frac{\partial}{\partial x_\mu} - iM_0 \right) = 0, \qquad \bar{\Psi} = \Psi^+ \eta_7, \tag{3.41}$$

$$\eta_\mu = 2\Lambda_\mu^2 - 1.$$

P.7. Generalization of field equation for Riemannian geometry

1) Connection between associator and torsion

Let us consider the case of flat space with torsion tensor $\Phi_{\mu\nu\lambda}$ [8]. Then taking into account the associator the covariant spinor derivative will be written as

$$D_\lambda \Psi = \left\{\frac{\partial}{\partial x_\lambda} - \Phi_{\mu\nu\lambda}(a_1 e_\mu e_\nu + a_2 e_\rho \Delta(e_\rho e_\mu e_\nu))\right\}\Psi. \tag{3.42}$$

Then from Lagrangian

$$L = \frac{1}{2}\left\{\bar{\Psi} e_\mu \frac{\partial \Psi}{\partial x_\mu} - \frac{\partial \bar{\Psi}}{\partial x_\mu} e_\mu \Psi + 2iK_0 \bar{\Psi}\Psi - \frac{1}{2}\Phi_{\nu\mu\lambda}\bar{\Psi} \times \right.$$
$$\left. \times e_\lambda (a_1 e_\mu e_\nu + a_2 e_\rho \Delta(e_\rho e_\mu e_\nu))\Psi - b\Phi_{\mu\nu\lambda}\Phi_{\mu\nu\lambda}\right\}, \tag{3.43}$$

we find the equation and the relation between torsion tensor of flat space with spin and associator

$$e_\mu \frac{\partial \Psi}{\partial x_\mu} + iK_0 \Psi + \Phi_{\nu\mu\lambda}(a_1 e_\nu e_\mu e_\lambda + a_2 e_\nu e_\rho \Delta(e_\rho e_\mu e_\lambda))\Psi = 0,$$
$$\Phi_{\mu\nu\lambda} = \frac{1}{8b}\{a_1 \bar{\Psi} e_\nu e_\mu e_\lambda \Psi + a_2 \bar{\Psi} e_\nu e_\rho \Delta(e_\rho e_\mu e_\lambda)\Psi\}. \tag{3.44}$$

However, another approach to this problem can be proposed. At present there are three antisymmetric tensors in theory.

Let us carry out their linear sum

$$F_{\mu\nu\lambda} = \alpha\Phi_{\mu\nu\lambda} + \beta\bar{\Psi} e_\mu e_\nu e_\lambda \Psi + \gamma\Delta(e_\mu e_\nu e_\lambda)\Psi. \tag{3.45}$$

The term $F_{\mu\nu\lambda}F_{\mu\nu\lambda}$ is added to the free Lagrangian and we obtain

$$L = \frac{1}{2}\left\{\bar{\Psi} e_\mu \frac{\partial \Psi}{\partial x_\mu} - \frac{\partial \bar{\Psi}}{\partial x_\mu} e_\mu \Psi + 2iK_0 \bar{\Psi}\Psi\right\} + bF_{\mu\nu}F_{\mu\nu},$$
$$e_\mu \frac{\partial \Psi}{\partial x_\mu} + iK_0 \Psi = 0, \tag{3.46}$$
$$\Phi_{\mu\nu\lambda} = -\frac{1}{2}\{\beta\bar{\Psi} e_\mu e_\nu e_\lambda \Psi + \gamma\bar{\Psi}\Delta(e_\mu e_\nu e_\lambda)\Psi\};$$

i.e. the influence of torsion is compensated by that of spin and associator.

2) Generalization of octonions for curved space

If octonion local basis (depending on the point) is designated as $E_\alpha(x)$, then provided $E_7(x) = ie_7(x)$ its associator part $E_\mu(x)$, $\mu = 4, 5, 6, 7$ forms 4-vectors of Lorentz group.

In this case $x_\mu(x_n, t)$ are real and signature is equal to $(---+)$. Then we have

$$(E_\mu E_\nu)+(E_\nu E_\mu)=-2g_{\mu\nu}E_0. \tag{3.47}$$

The field equation for nonassociative spinor is as follows

$$\left\{E_\mu(x)\left(\frac{\partial}{\partial x_\mu}+\Gamma_\mu\right)+iK_0\right\}\Psi=0, \tag{3.48}$$

where Γ_μ-spinor connection should be expressed by E_μ and their derivatives. In the general case one can introduce seven-dimensional space metric

$$(E_\alpha E_\beta)+(E_\beta E_\alpha)=-2g_{\alpha\beta}E_0. \tag{3.49}$$

§4. Gauge symmetries of Lagrangian, currents and conservation laws

P.1. Global gauge symmetries

The Lagrangian of free massless octonion field is invariant under the transformation (in algebra with "open" multiplication)

$$\Psi'=S\Psi, \qquad \Psi'=\bar{\Psi}S^{-1}, \qquad S^{-1}e_\mu S=e_\mu, \qquad S^{-1}S=1. \tag{4.1}$$

These requirements are met by:

$$\begin{gathered} 1.\ S_1=e^{i\alpha}, \qquad 2.\ S_2=\prod^{n} e^{\alpha_n e_n}, \qquad 3.\ S_3=\prod^{m,n} e^{i\alpha_{[nm]}e_n e_m}, \\ 4.\ S_4=\prod^{n,m,l}\exp(\alpha_{[nml]}e_n e_m e_l), \qquad n,m,l=1,2,3. \end{gathered} \tag{4.2}$$

$\alpha_{[nm]}$ and $\alpha_{[nml]}$ are antisymmetric.

In the case of infinitesimal transformations they can be combined into one transformation

$$S=\exp\{i\alpha+\alpha_n e_n+i\alpha_{[nm]}e_n e_m+\alpha_{[nml]}e_n e_m e_l\}. \tag{4.3}$$

We see that the exponent contains an undor constructed upon quaternion bases. In the case of rest mass there is only S_1 and S_3 symmetry.

P.2. Local gauge symmetries of Lagrangian

In case of octonion algebra with "open" multiplication the non-associativeness of algebra is implicit and the conditions of availability of local gauge symmetries are formulated as usual. For the Lagrangian of massless octonion field taking into account the interactions with gauge fields

$$L=\frac{1}{2}\{\bar{\Psi}\overleftarrow{D}_\mu e_\mu\Psi-\bar{\Psi}e_\mu\vec{D}_\mu\Psi\}, \qquad \begin{aligned}\overleftarrow{D}_\mu&=\overleftarrow{\partial}_\mu+B_\mu\\ \vec{D}_\mu&=\vec{\partial}_\mu+B^+_\mu\end{aligned} \tag{4.4}$$

the conditions of local symmetry take the form

$$\Psi'=S(x)\Psi, \qquad \Psi'=\bar{\Psi}\bar{S}(x),$$

$$\bar{S}e_\mu S=e_\mu, \qquad \bar{S}B'_\mu e_\mu S+\bar{S}_\mu e_\mu S=B_\mu e_\mu, \tag{4.5}$$

$$\bar{S}e'_\mu B^+_\mu S+\bar{S}e_\mu S_\mu=e_\mu B^+_\mu, \qquad S_\mu=\frac{\partial S}{\partial x_\mu}.$$

Having chosen

$$B_\mu=iA^\alpha_\mu Q_\alpha, \qquad B^+_\mu=-iA^{+\alpha}_\mu Q^+_\alpha. \tag{4.6}$$

Q_α-octonion basis, Lcan be written in the form

$$L=L_0+L^+_1, \qquad L_1=I^\alpha_\mu B^\alpha_\mu, \qquad I^\alpha_\mu=\frac{i}{2}\Psi(Q_\alpha e_\mu+e_\mu Q^+_\alpha)\Psi. \tag{4.7}$$

Then, we obtain

1. $$S_1=S_1(x), \qquad A'_\mu=A_\mu+\partial_\mu\alpha, \qquad I^0_\mu=i(\bar{\Psi}e_\mu\Psi)$$
$$Q^0=\int I^0_7 dV=\int\Psi^+\Psi dV.$$

2. $$S_2=S_2(x), \qquad A_\mu=A^n_\mu e_n, \qquad A'_\mu=A_\mu+i\partial_\mu\alpha-[\alpha A'_\mu], \qquad \alpha=\alpha_n e_n$$
$$I^n_\mu=i(\bar{\Psi}e_n e_\mu\Psi), \qquad Q^n=\int I^n_7 dV=i\int\Psi^+e_n\Psi dV,$$
$$\partial_\mu I^n_\mu=2iK_0(\bar{\Psi}e_n\Psi).$$

3. $$S_3=S_3(x), \qquad A^{nm}_\mu e_n e_m=A_\mu, \qquad A_\mu=A'_\mu-\partial_\mu\alpha+i[A'_\mu\alpha],$$
$$\alpha=\alpha_{nm}e_n e_m, \qquad I^{nm}_\mu=i(\bar{\Psi}e_\mu e_n e_m\Psi)\ n\neq m, \qquad \partial_\mu I^{nm}_\mu=0,$$
$$Q^{(nm)}=\int I^{nm}_7 dV=\int\Psi^+e_n e_m\Psi dV. \tag{4.8}$$

P.3. Octonions in the electromagnetic field

The solution of octonion field equation in the presence of electromagnetic field

$$(e_\mu\vec{D}_\mu+iK_0)\Psi=0, \qquad \bar{\Psi}(\overleftarrow{D}_\mu e_\mu-iK_0)=0,$$
$$\vec{D}_\mu=\vec{\partial}_\mu-ieA_\mu, \qquad \overleftarrow{D}_\mu=\overleftarrow{\partial}_\mu+ieA_\mu; \tag{4.9}$$

will be sought in the form:

$$\Psi=(e_\nu\vec{D}_\nu-iK_0)\varphi, \qquad \bar{\Psi}=\bar{\varphi}(\overleftarrow{D}_\nu e_\nu+iK_0), \tag{4.10}$$

we obtain

$$\{(-\vec{D}^2_\mu+K^2_0)e_0-\frac{i}{2}F_{\mu\nu}e_\mu e_\nu\}\varphi=0,$$
$$\varphi\{(-\overleftarrow{D}^2_\mu+K^2_0)e_0+\frac{i}{2}F_{\mu\nu}e_\mu e_\nu\}=0, \tag{4.11}$$

$$\vec{D}^2_\mu=\vec{\partial}^2_\mu-A^2_\mu-2iA_\mu\vec{\partial}_\mu, \qquad \overleftarrow{D}^2_\mu=\overleftarrow{\partial}^2_\mu-A^2_\mu+2i\overleftarrow{\partial}_\mu A_\mu.$$

Take into consideration that

$$iF_{\mu\nu}e_\mu e_\nu = -\{(\boldsymbol{\xi}\mathbf{E}) + i(\boldsymbol{\sigma}\mathbf{H})\} = \boldsymbol{\sigma}(e_8\mathbf{E} - i\mathbf{H}).$$

According to this relation it follows that the magnetic moment and electric moment of the electron are connected with spin $\boldsymbol{\sigma}$ and associator $\boldsymbol{\xi}$ respectively

$$\boldsymbol{\mu} = -\frac{eh}{2mc}\boldsymbol{\sigma}, \qquad \mathbf{d} = \frac{ieh}{2mc}\boldsymbol{\xi}. \tag{4.12}$$

If the solution is sought in the form

$$\varphi = \sum_\lambda^{\pm} D_8^\lambda \varphi^\lambda, \qquad D_8^\lambda = \frac{1}{2}(1 + ie_8), \tag{4.13}$$

$$\varphi^\lambda = \{f_0e_0 + f_1e_4 + f_2e_5 + f_3e_4e_5\}\hat{D}_{45}, \tag{4.14}$$

we find (in the framework of "open" multiplication)

$$\{(-\vec{D}_0^2 + K_0^2) - in_x^\lambda\}f_3^\lambda - \{in_y^\lambda - n_z^\lambda\}f_0^\lambda = 0, \tag{4.15}$$

$$\{(-\vec{D}_0^2 + K_0^2) - in_x^\lambda\}f_3^\lambda - \{in_y^\lambda - n_z^\lambda\}f_0^\lambda = 0, \quad \mathbf{n}^{(\pm)} = (\mathbf{H} \pm i\mathbf{E})$$

and identical equations for f_1^λ, f_2^λ, $(f_0^\lambda \to f_1^\lambda,\ f_3^\lambda \to f_2^\lambda)$.

P.4. Relativistic hydrogen atom

If the solution of equations

$$\left\{e_7\left(\frac{\partial}{\partial x_4} - ieA_4\right) + e_6\frac{\partial}{\partial z} + e_5\frac{\partial}{\partial y} + e_4\frac{\partial}{\partial x} + iK_0\right\}\Psi = 0 \tag{4.16}$$

are sought in the form

$$\Psi = e^{-iK_4x_4}(f_0e_0 + f_1e_4 + f_2e_5 + f_3e_4e_5)\hat{D}_{56}\hat{D}_7, \tag{4.17}$$

we obtain ordinary known combined equations of the first order for four functions [6] $f_1 \equiv \Psi_3$, $f_2 \equiv \Psi_2$, $f_3 \equiv \Psi_4$, $f_4 \equiv \Psi_1$.

P.5. Charge conjugation function

From octonion equation for the charge $+e$

$$(e_\mu D_\mu^+ + iK_0)\Psi = 0, \qquad D_\mu^\pm = \partial_\mu \pm ieA_\mu, \qquad A_\mu(A_n, iV), \tag{4.18}$$

by means of the complex conjugation

$$(e_7 D_7^- - e_6 D_6^- - e_5 D_5^- - e_4 D_4^- - iK_0)\Psi^* = 0, \qquad (4.19)$$
$$e_\mu^+ = -e_\mu, \qquad x_\mu(x_n, it)$$

multiplying the equation on the left $e_4 e_5 e_6 \equiv -e_8 e_7$ we find

$$(e_\mu D_\mu^- + iK_0)\Psi^c = 0, \qquad \Psi^c = e_4 e_5 e_6 \Psi^* = e_7 e_8 \Psi^*. \qquad (4.20)$$

§5. Perturbation theory in octonion field theory

In the case of octonion field equation, taking into account the interactions with the external electromagnetic field

$$\left(e_\mu \frac{\partial}{\partial x_\mu} + iK_0\right)\Psi = iee_\mu A_\mu \Psi. \qquad (5.1)$$

If the solution is sought in the form

$$\Psi = \Psi_0 + \Psi_1 + \ldots + \Psi_n, \qquad (5.2)$$

in the n-th approximation we find:

$$\left(e_\mu \frac{\partial}{\partial x_\mu} + iK_0\right)\Psi_{(n)} = iee_\mu A_\mu \Psi_{(n-1)}. \qquad (5.3)$$

Then taking

$$Q_1 \equiv e_\mu \frac{\partial}{\partial x_\mu} + iK_0, \qquad Q_2 \equiv \Psi_{(n)}, \qquad Q_3 \equiv iee_\mu A_\mu \Psi_{(n-1)}$$

we find for the solution

$$\Psi_{(n)} = \frac{e_\mu \dfrac{\partial}{\partial x_\mu} - iK_0}{-\dfrac{\partial^2}{\partial x_\mu^2} + K_0^2} iee_\mu A_\mu \Psi_{(n-1)}. \qquad (5.4)$$

As it is seen, the specificity of nonassociative algebra is not explicit in the construction of the perturbation theory, hence all the laws of Dirac theory and of Feinmann rules type in particular for the construction of matrix elements are valid, the only difference being that values γ_μ are replaced by e_μ.

Since the matrix element has the general structure of the form

$$M \sim \Psi^+(\quad)\Psi \sim \hat{\Pi}(\quad)\hat{\Pi},$$

where $\hat{\Pi}$ is the projection operator of the type $\hat{\Pi} \sim D_8^\alpha D_{45}^\beta$, and in the brackets there will ultimately be the value of the general structure

$$a_0 e_0 + a_8 e_8 + a_\mu e_\mu + a_{\mu 8} e_\mu e_8 + a_{\mu\nu} \underline{E}_{\mu\nu}, \tag{5.5}$$

we find

$$M \sim \hat{\Pi}(a_0 - (\alpha a_8 + \beta a_{45}))\hat{\Pi}. \tag{5.6}$$

Since for obtaining final results there is no need in transition to the main basis with application of nonassociative multiplication rules, so the final results of Dirac theory in the formulation concerned remain unchanged.

Consequently, the classical octonion field theory in the given version completely encompasses the Dirac theory and adds a new quantum number, the associator, and possesses an additional symmetry. In particular, if octonion Q is provided with the unitary index "l" of the group SU (3), then the general vector $M = Q^l \lambda_l$ where λ_l are Gell-Mann matrix will be the so-called matrix M_8^3 and symmetry E_6 can be included into the theory.

The author owes acknowledgements to Prof. D. Ivanenko for valuable remarks.

References

1. А. Г. Курош, Лекции по общей алгебре, Госфизматиздат, Москва, 1962 г., стр. 262.
2. A. Pais, Phys. Rev. Lett., *7*, 291, 1961.
3. M. Günydin and F. Gürsey, Phys. Rev. D, *9*, 3387, 1974; R. Casalbaumi, G. Domokos and S. Kövesi-Domokos, Nuvo Cimento, *31A*, 423, 1976.
4. H. H. Goldstine and L. P. Horwitz, Proc. Nat. Acad. Sci. *48*, 1134, 1962.
5. F. Gursey, P. Romand and P. Sikivie, Phys. Rev. D, *10*, 2167, 1975.

F. Gursey, M. Serdarolu, Nuvo Cimento, *65A*, 337, 1981.

6. А. Зоммерфельд, Строение атома и спектры, Т. II. Госфизматиздат, Москва, 1956 г., стр. 203.
7. А. И. Ахиезер, В. Б. Берестецкий, Квантовая электродинамика, Госфизматиздат, Москва, 1976 г., стр. 182.
8. Д. Иваненко, в сборнике, Проблемы физики: классика и современность. «Мир», Москва, 1982 г. стр. 127.

В. И. Родичев, ЖЭТФ, *40*, 1456, 1961.

Acta Physica Hungarica 57 (1—2), pp. 99—111 (1985)

CORRELATION PROPERTIES AND THE STRUCTURAL INSTABILITY OF THE TWODIMENSIONAL ANHARMONIC CRYSTAL. THE VIRIAL APPROACH

V. K. PERSHIN and I. S. GERSHT

Ural Polytechnical Institute, 620002 Sverdlovsk, USSR

(Received in revised form 9 August 1983)

In terms of the virial theorem for an arbitrary two-dimensional lattice a self-consistent system of equations is developed that allows its dynamics and the temperature evolution to the stability limit to be investigated in the nearest neighbour approximation. Contrary to the traditional approach the relative correlation functions of the longitudinal and transversal displacements of particles and the force constants corresponding to them are introduced. The topology of the virial surface and the evolutional trajectory of the lattice state, the temperature behaviour of the longitudinal and transversal force constants of the lattice are discussed. The irregular growth of the adiabatic elastic modulus with temperature, and the change of sign of its derivative observed experimentally in cubic crystals are explained. It is shown that there exists a temperature range where the system stability is achieved by the redistribution of the kinetic energy between the longitudinal and transversal displacements of particles. It is established that the instability of the lattice relative to its longitudinal oscillations is the main cause of its structural re-arrangement in the high temperature range in spite of the more rapid development of the transversal oscillations with temperature increase.

1. Introduction

The balance between the kinetic energy of particles and the energy of their coupling in the lattice is destroyed during the structural transitions in crystals [1] which is why the virial theorem establishing the relation between these most important characteristics of the system is of great significance regarding the stability of the crystalline state [2]. The advantage of the use of the virial equation was proved by its having given the best practical results for weakly interacting systems [3]. In spite of this, however, the application of the virial theorem for investigating the properties of crystals and in particular the mechanisms of their structural instability is still at its very beginning.

In the present paper the mechanisms of the formation of the two-dimensional lattice instability due to temperature increase are discussed in terms of a self-consistent pseudoharmonic approximation based on the virial theorem. The successive introduction of the correlation functions for different types of displacements gives the possibility to establish the importance of each of them for instability formation.

According to the Landau—Peierls theorem the system of atoms in the plane forms a crystalline lattice at $T=0$ K only. Long-range order of the lattice is destroyed by long-wave fluctuations at any non-zero temperature. Strictly speaking, such a system is not a crystalline one and it can be considered as such only within the range of distance R [4]

$$R=a\exp(\pi mu^2/2kT),$$

where a is the average distance between the particles; u is the average sound velocity and m is the mass of the atom.

The absence of long-range order does not rule out the possibility of structural instability in such systems, a fact confirmed by high temperature expansion [5, 6], the correlation functions technique [7], computer simulation [8, 9] and the self-consistent phonons method [10, 11].

Further, for simplifying the calculations the case of the high temperature limit will be considered only where the theorem of the equipartition of kinetic energy over the degrees of freedom of particles is fulfilled. If the mass of atoms is large enough then the lowest temperatures can be involved in this classical range [12].

2. Self-consistent system of equations

Let us consider a combination of the atoms interacting with each other by the pair central forces in the plane. The Hamiltonian of such a system has the usual form

$$H=K+U,$$

where

$$K=\sum_j\left(-\frac{\hbar^2}{2m}\nabla_j^2\right),\qquad U=\frac{1}{2}\sum_{k,1}\Psi(|\mathbf{r}_k-\mathbf{r}_1|).$$

Let us write the virial theorem for an arbitrary degree of freedom of the atom "0" located at the origin of the coordinate system

$$2m\langle(\dot{u}_0^i)^2\rangle=\left\langle u_0^i\frac{\partial U}{\partial u_0^i}\right\rangle,\qquad i=1,2;$$

where u_0^i is the i-component of the translational displacement of the atom from the equilibrium position. For the high temperature limit, by expanding the potential energy U in this relation into the infinite series over displacements and presenting the multidimensional correlation functions in the pseudo-harmonic approximation [13] we get

$$kT=-\sum_{i',\mathbf{n}}x_{0\mathbf{n}}^{ii'}\Phi_{0\mathbf{n}}^{ii'},\qquad i,i'=1,2,;\tag{1}$$

where $x_{0\mathbf{n}}^{ii'}=\langle(u_0^i-u_\mathbf{n}^i)(u_0^{i'}-u_\mathbf{n}^{i'})\rangle$, $\mathbf{n}$ is the radius-vector of a certain lattice site and

$$\Phi_{0\mathbf{n}}^{ii'}=\left\langle\frac{\partial^2 U}{\partial u_0^i\,\partial u_\mathbf{n}^{i'}}\right\rangle=\frac{\partial^2}{\partial u_0^i\,\partial u_\mathbf{n}^{i'}}\exp\left[-\frac{1}{4}\sum_{\substack{jj'\\ \mathbf{pp}'}}x_{\mathbf{pp}'}^{jj'}\frac{\partial^2}{\partial u_\mathbf{p}^j\,\partial u_{\mathbf{p}'}^{j'}}\right]U\tag{2}$$

is the force constant defined by the functions $x_{0\mathbf{n}}^{ii'}$ included in Eq. (1). Each of these functions can be determined by the spectral theorem

$$x_{0\mathbf{n}}^{ii'} = \frac{1}{mN} \sum_{\mathbf{k},j} \frac{e_{\mathbf{k}j}^{i} e_{\mathbf{k}j}^{i'}}{\omega_{\mathbf{k}j}} (1 - e^{-i\mathbf{kn}}) \operatorname{cth} \frac{\omega_{\mathbf{k}j}}{2kT}, \tag{3}$$

where N is the number of particles; $\mathbf{k}$ is the quasi-wave vector; j is the number of the oscillation branch.

The force constants (2) are calculated for the equilibrium state of the lattice. Therefore, so that the system of equations (1)—(3) becomes closed to all unknown parameters it should be added to the conditions of the lattice equilibrium.

The frequencies of the excitations of the crystal, $\omega_{\mathbf{k}j}$, included in (3) and the components of the eigenvectors corresponding to them, $e_{\mathbf{k}j}^{i}$, are determined by the relation

$$\operatorname{Det} \| D_{ii}(\mathbf{k}) - m\omega_{\mathbf{k}j}^{2} \delta_{ii'} \| = 0, \tag{4}$$

$$m\omega_{\mathbf{k}j}^{2} e_{\mathbf{k}j}^{i} - \sum_{i'} D_{ii'}(\mathbf{k}) e_{\mathbf{k}j}^{i'} = 0. \tag{5}$$

These parameters are also the functions of $\{x\}$ since the elements of the dynamical matrix $D = \| D_{ii'}(\mathbf{k}) \|$,

$$D_{ii'}(\mathbf{k}) = \sum_{\mathbf{n}} \Phi_{0\mathbf{n}}^{ii'} e^{i\mathbf{kn}} \tag{6}$$

depend on the pseudo-harmonic force constants (2). Then, from Eq. (4) and (5), for an arbitrary lattice we have

$$m\omega_{\mathbf{k}j}^{2} = \frac{1}{2} (D_{11} + D_{22} + (-1)^{j+1} \sqrt{(D_{11} + D_{22})^{2} - 4(D_{11} D_{22} - D_{12}^{2})}, \tag{7}$$

$$e_{\mathbf{k}j}^{1} = \frac{D_{12}}{\sqrt{D_{12}^{2} + (D_{11} - m\omega_{\mathbf{k}j}^{2})^{2}}}; \qquad e_{\mathbf{k}j}^{2} = -\frac{D_{11} - m\omega_{\mathbf{k}j}^{2}}{\sqrt{D_{12}^{2} + (D_{11} - m\omega_{\mathbf{k}j}^{2})^{2}}}. \tag{8}$$

Since $D_{12} = 0$ due to reflection symmetry for rectangular and square lattices $e_{\mathbf{k}j}^{i} = \delta_{ij}$. Going over to the high temperature limit $\hbar\omega \ll kT$ in (3), integrating over the k-space and using also relations (7) and (8) we get

$$x_{0\mathbf{n}} = \frac{kT}{\pi^{2}} S \int_{(\mathbf{k})} \frac{1}{\Delta(\mathbf{k})} A D A^{T} \operatorname{Sin}^{2} \frac{\mathbf{kn}}{2} d^{2}k, \tag{9}$$

where $x_{0\mathbf{n}} = \| x_{0\mathbf{n}}^{ii'} \|$ is the second order translational matrix; A is the $\pi/2$-rotation matrix; S is the area of the unit cell; $\Delta(\mathbf{k}) = D_{11} D_{22} - D_{12}^{2}$.

The combination of relations (1)—(9) defines the self-consistent system of equations describing the dynamic state of a two-dimensional lattice for an arbitrary number of coordination spheres and in the particle interaction approximation. Let us note that the virial equations (1) are represented by the linear combinations formed accordingly of the elements of the correlation matrix (9) where the force constants (2)

are used as coefficients. Thus, they are not independent. Because of their clear physical meaning, however, the virial equations have a significant advantage over the spectral ones which is precisely why the virial theorem is applied for interpreting the mechanisms of the instability formation.

3. Square lattice in the approximation of the nearest neighbours

According to Eq. (1) the two-dimensional lattice with the arbitrary angle θ between its axes is described by two virial equations in the nearest neighbour approximation. Each of them has four constants $\Phi^{ii'}_{0\mathbf{n}}$ (i is the fixed index; α is the number of the nearest neighbour) depending on the full set of correlations $\{x^{k1}_{0\mathbf{n}_\alpha}\}$. The problem is essentially simplified for the case of the square lattice. Due to its symmetry we have

$$\Phi^{12}_{0\mathbf{n}_\alpha}=\Phi^{21}_{0\mathbf{n}_\alpha}=0\,, \qquad x^{12}_{0\mathbf{n}_\alpha}=x^{21}_{0\mathbf{n}_\alpha}=0\,.$$

The consequence of the isotropy of the square lattice is the degeneration of the force constants

$$\Phi^{11}_{0\mathbf{n}_1}=\Phi^{22}_{0\mathbf{n}_2}=\Phi_1\,, \qquad \Phi^{11}_{0\mathbf{n}_2}=\Phi^{22}_{0\mathbf{n}_1}=\Phi_2\,,$$

of the correlation functions

$$x^{11}_{0\mathbf{n}_1}=x^{22}_{0\mathbf{n}_2}=x_1\,, \qquad x^{22}_{0\mathbf{n}_1}=x^{11}_{0\mathbf{n}_2}=x_2$$

and of the two virial equations into one. With the notations introduced before the latter can be written in the form

$$kT=-x_1\,\Phi_1-x_2\,\Phi_2\,. \tag{10}$$

The correlation functions x_i describe the longitudinal ($i=1$) and transversal ($i=2$) displacements of particles and characterize the extent of corresponding oscillations in the crystal at the given temperature.

To form a closed system one of the following equations

$$x_i=-\frac{2kT}{\Phi_i}\,\mathrm{arc\,sin}\,G_i\,; \qquad i=1,2 \tag{11}$$

is to be added to Eq. (10), where $G_i=\sqrt{\Phi_i/(\Phi_1+\Phi_2)}$. These values are obtained by calculating the elements of the dynamic matrix (6) in the nearest neighbour approximation

$$D_{ii'}(\mathbf{k})=-4\left(\Phi^{ii'}_{0\mathbf{n}_1}\sin^2\frac{\mathbf{kn}_1}{2}+\Phi^{ii'}_{0\mathbf{n}_2}\sin^2\frac{\mathbf{kn}_2}{2}\right)$$

and putting them into (9). Because the equation $\Phi_2=0$ is strictly valid in the harmonic case in consideration of the first coordination sphere only [14] one can suppose that in the pseudo-harmonic approximation the unequality

$$|\,\Phi_2\,|\ll|\,\Phi_1\,| \tag{12}$$

is fulfilled. Taking into account the resulting approximative relation arcsin $G_2 \approx G_2$ after corresponding transformations we get:

$$\frac{x_1}{x_2} = \frac{\pi}{2}\left(\frac{\Phi_2}{\Phi_1}\right)^{1/2} - \frac{\Phi_2}{\Phi_1}. \tag{13}$$

The system of equations (10), (13) will be closed and self-consistent together with the condition of the equilibrium of the crystalline lattice only if

$$\partial\langle\Psi\rangle/\partial l = 0, \tag{14}$$

where $\langle\Psi\rangle$ is the averaged isotropic pair interaction potential; l is the equilibrium parameter of the lattice in the self-consistent approximation.

Now we solve the system (10), (13), (14) for the Morse potential [15]

$$\Psi(r) = D(e^{-2g\left(\frac{r}{r_0}-1\right)} - 2e^{-g\left(\frac{r}{r_0}-1\right)}), \tag{15}$$

where D is the depth of the potential well; r_0 is the equilibrium distance in harmonic approximation; $g = \alpha r_0$, $\alpha > 0$ are the potential parameters. The interaction potential corresponds to the conditions $\Psi(r_-/r_0) = 0$, $\partial^2\Psi(r_+/r_0)/\partial^2 r = 0$, where $r_\pm/r_0 = 1 + \ln^2/g$. From this consideration and from the analysis of the dependence of the half-width of the potential well on g it follows that the interaction forces in the lattice are of more long-range order now.

The averaged interaction potential $\langle\Psi_{0\mathbf{n}}\rangle$ corresponds to the equilibrium state and is calculated by the equation [13]

$$\langle\Psi_{0\mathbf{n}}\rangle = \exp\left\{\frac{1}{2}\sum_{ii'} x_{0\mathbf{n}}^{ii'} \nabla_{\mathbf{r}_n}^{i} \nabla_{\mathbf{r}_n}^{i'}\right\}\Psi(\mathbf{r}_n) \tag{16}$$

in the pseudo-harmonic approximation. Putting (15) into (16) and confining ourselves to the terms of order not less than $x_{0\mathbf{n}}^{ii'}/l_{\mathbf{n}}^2$ we get the expression for the anisotropic averaged potential of an arbitrary lattice

$$\langle\Psi_{0\mathbf{n}}\rangle = D(\mathbf{r}/r)\sum_{p=1}^{2}(-1)^{p+1}p\exp\{-\alpha_p(l_{\mathbf{n}} - r_0) - \\ -\sum_{i,j}\frac{\alpha_p}{2 l_{\mathbf{n}}}\left(\delta_{ij} - \frac{\alpha_p n^i n^j}{l_{\mathbf{n}}} - \frac{n^i n^j}{l_{\mathbf{n}}^2}\right)x_{0\mathbf{n}}^{ij}\}, \tag{17}$$

where $\alpha_1 = 2\alpha(\mathbf{r}/r)$, $\alpha_2 = \alpha(\mathbf{r}/r)$, $|\mathbf{n}| = 1$. For a square lattice $\left(\sum_i n_\alpha^i n_{\alpha'}^i = \delta_{\alpha\alpha'} l^2,\ D(\mathbf{r}/r) \equiv D,\ \alpha(\mathbf{r}/r) \equiv \alpha\right)$ in the coordinate system $n_\alpha^i = \delta_{i\alpha} l$, we have

$$2\rho = \rho_0 + \sqrt{\rho_0^2 - 2x_\perp}, \tag{18}$$

$$\Phi_{\parallel} = (1 - x_\perp/\rho^2)\exp(-x_{\parallel}), \tag{19}$$

$$\Phi_\perp = (x_\perp - x_{\parallel})\exp(-x_{\parallel})/\rho^2, \tag{20}$$

where

$$\rho=\alpha l\,,\qquad \rho_0=g+3x_{\|}/2\,,\qquad x_{\|}=\alpha^2x_1\,,\qquad x_{\perp}=\alpha^2x_2\,,$$

$$\Phi_{\|}=-\Phi_1/2D\alpha^2\,,\qquad \Phi_{\perp}=-\Phi_2/2D\alpha^2\,.$$

Relations (19) and (20) confirm the correctness of the supposition expressed above by unequality (12). Putting them into Eq. (13), squaring the two parts of the equation obtained, up to first order of the parameter $(x_{\perp}/\rho_0^2)$ we can transform (13) to

$$x_{\perp}^3+ax_{\perp}^2+bx_{\perp}+c=0\,. \tag{21}$$

The following notations are introduced in (21)

$$a=-x_{\|}\left(1+\frac{8}{\pi^2}\right)<0,\qquad b=x_{\|}^2\left(\frac{20}{\pi^2}\right),\qquad c=-x^2\left(\frac{^2\rho_0^2}{\pi}\right)<0\,. \tag{22}$$

The real solution of Eq. (21) can be written [16] as:

$$x_{\perp}=f(x_{\|})=\frac{1}{3}\left(1+\frac{8}{\pi^2}\right)x_{\|}+\left(\frac{2}{\pi}g\right)^{2/3}x_{\|}^{2/3}. \tag{23}$$

Now, putting (19) and (20) into the virial equation (10) we get

$$\frac{kT}{2D}=\frac{F(x_{\|},x_{\perp})}{2D}=x_{\|}\left(1-\frac{x_{\perp}}{\rho^2}\right)e^{-x_{\|}}+x_{\perp}\frac{x_{\perp}-x_{\|}}{\rho^2}e^{-x_{\|}}. \tag{24}$$

The notations $f(x_{\|})$ and $F(x_{\|},\,x_{\perp})$ have been introduced in Eqs (23), (24) for later convenience.

Using (23) in Eq. (24) and confining ourselves to the terms up to first order of the parameter $x_{\|}/g^2$ we get

$$\frac{kT}{2D}=\frac{F(x_{\|},f(x_{\|}))}{2D}=x_{\|}e^{-x_{\|}}(1+Q_g(x_{\|}))\,, \tag{25}$$

where

$$Q_g=\sum_{p=1}^{3}c_p\left(\frac{x_{\|}}{g^2}\right)^{p/3},$$

$$c_1=(2/\pi)^{4/3}\approx0.55,\qquad c_2=-\frac{4}{3}\left(1-\frac{4}{\pi^2}\right)\left(\frac{2}{\pi}\right)^{2/3}\approx-0.59\,,$$

$$c_3=-\frac{5}{9}\left(1+\frac{8}{\pi^2}\right)\left(1-\frac{8}{5\pi^2}\right)\approx-0.8\,.$$

The value of the correlation function x_c corresponding to the temperature of the dynamic instability of the lattice, T_c, is determined by the extremum condition for the function $F(x_{\|},\,f(x_{\|}))$

$$\frac{dF}{dx_{\|}}=(1+Q_g(x_{\|}))\,(1-x_{\|})+x_{\|}\frac{dQ_g}{dx_{\|}}=0\,. \tag{26}$$

The estimations show that $Q_g(x_{||}) \ll 1$ in the range of the physically reasonable values of g and $x_{||}(2 \leq g \leq 10,\ 0 < x \approx 1)$. Taking this into account we shall take the solution to equation (26) in the form $x_c = 1 + \Delta_g$, $\Delta_g \ll 1$. This gives

$$x_{||_c} = 1 + \sum_{p=1}^{3} p c_p g^{-2p/3}. \tag{27}$$

Putting (27) into Eq. (23) we have

$$x_{\perp_c} = \frac{1}{3}\left(1 + \frac{8}{\pi^2}\right) + \left(\frac{2}{\pi} g\right)^{2/3} + Q_g, \tag{28}$$

where

$$Q_g = \left[\frac{1}{3}\left(1 + \frac{8}{\pi^2}\right) + \frac{2}{3}\left(\frac{2}{\pi} g\right)^{2/3}\right] \sum_{p=1}^{3} \frac{1}{3} p c_p g^{-2p/3}.$$

Now, using (25), we can write the real expression for the instability temperature, T_c,

$$T_c = T_A\left(1 + \sum_{p=1}^{3} c_p g^{-2p/3}\right),$$

where $T_A = 2D/ke$ is the instability temperature of the lattice without taking the relative transversal correlations [18] into account. The contribution to the critical values of the correlation functions which is bound with the second and the third term in relations (27) and (28), respectively, is negative at $g \geq 2$ and the modulus of its magnitude decreases with increasing g. The inequality $T_c \leq T_A$ is valid for $g \leq g_0 \approx 2.6$ where g_0 is determined by the equation $T_c(g) = T_A$. For all the other values of g the reverse inequality is valid.

It is mentioned that the estimation of the parameters $x_{||_c}$, $x_{\perp_c}$, T_c is mathematically equivalent to the problem of the estimation of the conventional extremum of the function $F = F(x_{||}, x_\perp)$ if the coupling equation $x_\perp - f(x_{||}) = 0$ is available [17—19].

4. Topology of the virial surface and the evolution trajectory

In Fig. 1 the virial surface of the square lattice $F = F(x_{||}, x_\perp)$ is shown. The radius-vector $\mathbf{r}$ of any point located on this surface and the unit tangent vector $\boldsymbol{\tau}$ of the surface can be represented in the form

$$\mathbf{r} = \mathbf{i}x_{||} + \mathbf{j}x_\perp + \mathbf{k}F(x_{||}, x_\perp), \tag{29}$$

$$\boldsymbol{\tau} = \frac{d\mathbf{r}}{ds} = \frac{1}{\xi}\left(\mathbf{i} + \mathbf{j}\frac{dx_\perp}{dx_{||}} + \mathbf{k}\frac{dF}{dx_{||}}\right), \tag{30}$$

where the parameters x_j change within the range from 0 to $x_{jc}(j = ||, \perp)$; at the given dependence $x_\perp = x_\perp(x_{||})$ s is the length of the part of the trajectory on the surface from a certain fixed point to the given one; $\xi = dx_{||}/ds = [1 + (dx_\perp/dx_{||})^2 + (dF/dx_{||})^2]^{1/2}$.

The curve DA (Fig. 1) is described by the radius-vector (29) at $x_{\|}=1-x_\perp^2 g^{-2}(1-2x_\perp g^{-2})^{-1}$, $F=(1-2x_\perp g^{-2})e^{-x_{\|}}$. At $g^2 \gg 1$ this curve turns into the nearly plane parabola $F(x_\perp)=(1-2x_\perp g^2+x_\perp^2 g^{-2})e^{-1}$ located on the surface $x_{\|}=1-x_\perp^2 g^{-2}$ perpendicular to the plane $x_{\|}ox_\perp$.

At the values $x_{\|}^B=1-x_\perp^B(1-x_\perp^B g^{-2})^{-1}$, $x_\perp^B=1-g^2/4-(1-g^4/4)^{1/2}$ the curve has a minimum corresponding to the saddle-point of the surface with which the existence of the virial well is bound. At $g^2 \gg 1$ $x_{\|}^B=x_\perp^B=1$, $F^B=e^{-1}(1-g^{-2})$. Here and henceforth

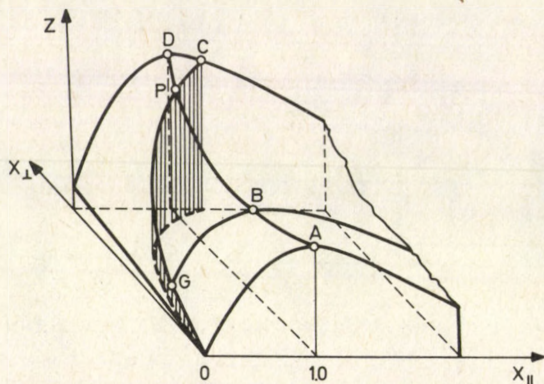


Fig. 1. Virial surface of the square lattice ($z=F(x_{\|}, x_\perp)/2D$)

alphabetical symbols corresponding to the points on the virial surface are used as indices.

Putting $x_\perp=0$ into (29) we get the virial for a square lattice with longitudinal displacements. It completely coincides with the virial for the one-dimensional lattice with the interaction between the nearest particles each having one degree of freedom along the crystalline axis. The evolution trajectory of the system is a plane curve — it is the part 0A of the virial where the point $A=x_{\|}^A=1, x_\perp^A\equiv 0, F^A=e^{-1}$ corresponds to the maximum of the virial over which the radius-vector moves as the temperature changes. The vector τ determined by relation (30) is directed to the virial increase over the whole curve except point A. At point A the situation is changed and vector τ takes the direction of the virial slope.

At $x_A \neq 0$ the evolution trajectory of the lattice state is the part of the curve of the intersection of the virial surface $F(x_{\|}, x_\perp)$ and the surface corresponding to the coupling equation $x_\perp=f(x_{\|})$. Vector τ has positive components for points of the latter surface and it indicates that the system has a tendency to change state with increasing temperature.

5. Importance of longitudinal and transversal displacements of particles in the formation of the structural instability

In Fig. 2 the temperature dependence of the functions $x_{||}$ and $x_{\perp}$ for different values of g are shown. From the Figure it is clear that $x_{||}$ and $x_{\perp}$ are less dependent on the temperature in systems where the potential is of more long range order. For the

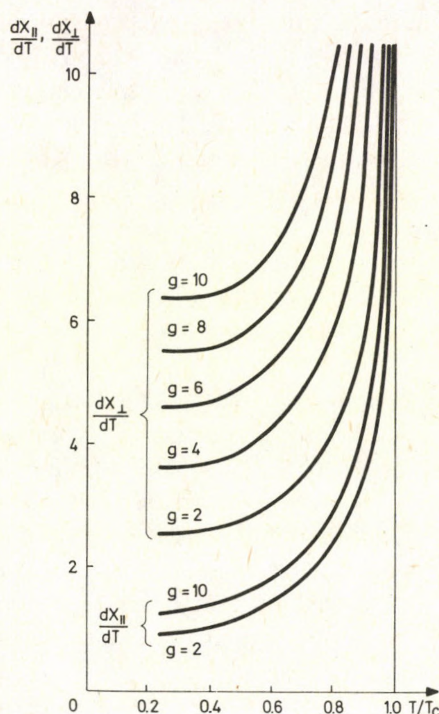


Fig. 2. Dependence of the correlation functions of the displacements $x_{||}$ and $x_{\perp}$ on T/T_c at different values of parameter g. The dotted line is the range of $x_{||}$- and $x_{\perp}$-values located beyond the hump of the virial surface

fixed value of g the range of change of $x_{\perp}$ exceeds significantly that of $x_{||}$ and the unequality

$$S_g = \frac{x_{\perp}}{x_{||}} = \frac{1}{3}\left(1 + \frac{8}{\pi^2}\right) + \left(\frac{2}{\pi} g\right)^{2/3} x_{||}^{-1/3} > 1 \tag{31}$$

is fulfilled at any temperature $T \leq T_c$.

From Fig. 2 and unequality (31) it follows that the $x_{\perp}$-oscillations are more developed than the $x_{||}$-ones in the high temperature range of the crystalline state. This is also confirmed by the fact that the growth rate of the $x_{\perp}$-fluctuations exceeds the corresponding growth rate of $x_{||}$-fluctuations as temperature increases.

Differentiating Eqs. (25) and (23) with respect to temperature we get

$$\frac{dx_{\|}}{dT}=\frac{k}{2D}\left(\frac{dF}{dx_{\|}}\right)^{-1}.$$

$$\frac{dx_{\perp}}{dT}=\frac{k}{2D}\left[\frac{1}{3}\left(1+\frac{8}{\pi^2}\right)+\frac{2}{3}\left(\frac{2}{\pi}g\right)^{2/3}x_{\|}^{-1/3}\right]\left(\frac{dF}{dx_{\|}}\right)^{-1}. \tag{32}$$

Since at T_c Eq. (26) is fulfilled identically, then from (32) it follows that at $T\to T_c\, dx_{\|}/dT$, $dx_{\perp}/dT\to\infty$. Hence, the approach to the instability point is accompanied by the sharp growth of the $x_{\|}$- and $x_{\perp}$-fluctuations of the displacement or particles from the equilibrium position. The instability of the lattice is due not to one vibrational movement but to $x_{\|}$- and $x_{\perp}$-oscillations simultaneously. However, each of them plays a different role in instability formation.

To substantiate the latter statement we consider the temperature dependences of the force constants $\Phi_{\|}$ and $\Phi_{\perp}$ (Fig. 3) obtained on the basis of relations (19) and (20). According to Fig. 3a, $\Phi(T)$ is a decreasing monotonic function of the temperature. Such a dependence is due to both $x_{\perp}$ and $x_{\|}$ leading to a decrease in the value of $\Phi_{\|}(T)$ as the temperature increases. The $x_{\|}$-correlation functions give the main contribution to the process of the lattice softening over the longitudinal force constant.

The dependence $\Phi_{\perp}(T)$ shown in Fig. 3b differs essentially from $\Phi_{\|}(T)$. As the temperature increases the growth of $\Phi_{\perp}(T)$ occurs first, then this function reaches its maximum value and starts decreasing in the pre-transitional range. It follows from Eq. (20) that such behaviour of the function $\Phi_{\perp}(T)$ is linked with the temperature competition of the stabilization and destabilization processes due to the temperature evolution of the $x_{\|}$ and $x_{\perp}$ correlation functions. $x_{\|}$ correlation functions give a destabilizing contribution both to the longitudinal and transversal force constants.

Preliminary calculations show that the functions $\Phi_j(T)$ for simple cubic crystals are similar to those shown in Fig. 3. According to [13] the relation between the elastic modulations c_{ijkl} and the force constants of the lattice is determined by the relations

$$\tilde{c}_{ijkl}=\frac{1}{l^3}\sum_{\mathbf{n}}\Phi^{ij}_{0\mathbf{n}}n_k n_l\,,$$

$$2\tilde{c}_{ijkl}=c_{ikjl}+c_{iljk}\,.$$

For the cubic crystals we have

$$c_{11}=\frac{c_{iiii}}{2D\alpha^2}=\frac{\Phi_{\|}}{l}; \qquad c_{12}=\frac{c_{iijj}}{2D\alpha^2}=\frac{\Phi_{\perp}}{l}.$$

Specifically, these equations allow one to interpret the anomalous growth of c_{12} and the sign change of the derivative dc_{12}/dT[20]. We note that in cubic structures of NaCl type the elastic moduli change with the temperature in accordance with the dependences obtained above [21].

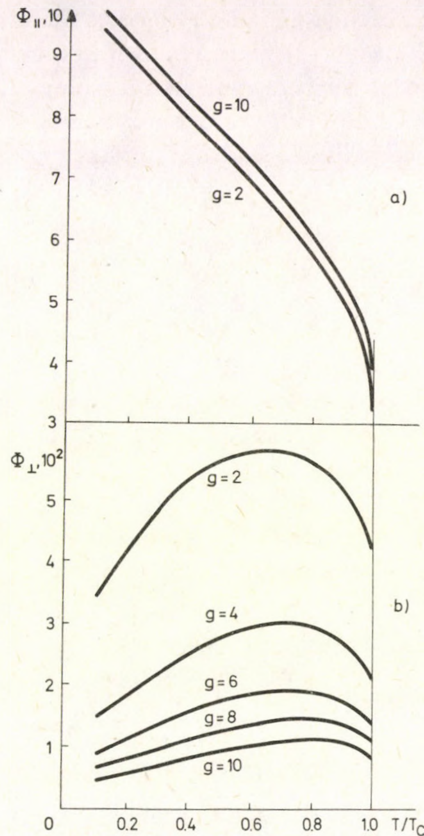


Fig. 3. Dependence of (a) longitudinal and (b) transversal force constants of the square lattice on the given temperature T/T_c at different values of parameter g

Let us consider the mechanism of the formation of the system instability more thoroughly. We write the infinitesimal change of the kinetic energy of the degree of freedom and of the virial in the form

$$KdT = dF = dF_{x_{\parallel}} + dF_{x_{\perp}}$$

where

$$dF_{x_{\parallel}} = \frac{\partial F}{\partial x_{\parallel}} dx_{\parallel}, \qquad dF_{x_{\perp}} = \frac{\partial F}{\partial x_{\perp}} dx_{\perp} .$$

Since $F(x_{\parallel}, x_{\perp})$ is a function of $x_{\perp}$ that increases monotonically up to point C on the virial surface, then $dF_{x_{\perp}} > 0$ over the whole of the evolution trajectory. The second differential $dF_{x_{\parallel}}$ is positive up to the crossing point P only. It changes its sign at P so the $x_{\perp}$-subsystem accepts the thermal energy at any point of the evolution trajectory up to the instability point, and the $x_{\parallel}$-subsystem up to point P only. Beyond this, the $x_{\parallel}$-subsystem transfers a part of its energy to the $x_{\perp}$-subsystem — which also accepts the

thermal energy from the thermostat in this range of the evolution trajectory since $dF>0$. The condition $dF_{x_{\parallel}}=-dF_{x_{\perp}}$ corresponds to the extremum of the evolution trajectory which is reached at temperature T_c. At this temperature the increase in the kinetic energy of the lattice does not occur since due to the latter equality the $x_{\perp}$-subsystem is not able to accept the thermal energy of the thermostat which has already been used for the destruction of the lattice $\left(\frac{dE}{dT}=\frac{dK}{dT}+\frac{dU}{dT}\to\infty\right.$, at $T\to T_c$; E is the internal energy$\Big)$. Further temperature increase is not possible without structural change of the system because the $x_{\parallel}$-subsystem is not able to transfer to the $x_{\perp}$-subsystem the energy excess destroying it ($dF_{x_{\parallel}}>dF_{x_{\perp}}$).

Thus, the existence of the crystalline lattice derives from the presence of the connection between the $x_{\parallel}$ and $x_{\perp}$ subsystem at $T>T_p$ and the re-distribution of the kinetic energy in this temperature range. The instability of the lattice relative to the longitudinal oscillations of particles due to the limitation of the energy capacity of the $x_{\parallel}$ subsystem and to the impossibility of the re-distribution of the kinetic energy between the $x_{\parallel}$ and $x_{\perp}$ subsystems at $T>T_c$ is the main cause of its structural rearrangement.

6. Conclusion

The first approach to the study of the structural instability of a crystal was outlined on the basis of the virial theorem of statistical mechanics in [1]. This approach is general enough. If the virial of the system could be calculated exactly merely for a certain type of interaction, then it would indicate an exact solution of the problem of the phase transition. The difficulty is that, as in the case of any other approach, the calculation of the virial relations can be carried out in certain approximations only. Therefore, the application of the pseudo-harmonic approximation to the analysis of the stability conditions, i.e. a certain procedure of the uncoupling of the multidimensional correlation functions of displacements, is not a matter of principle in qualitative considerations. Without restriction of the generality, it allows us to write a virial of the system in the real analytical form and to determine a complete set of the correlation functions describing a crystal in this approximation. Each of these correlation functions can be calculated by the use of the spectral theorem. Such a procedure allows one to add a system of virial equations up to the closed one for any number of variables in the system. The virial relations are valid for any aggregated substance since suppositions about the structural arrangement of the neighbours are not used in their derivation. But they undergo a destruction of the change type at the boundary of the temperature range of the existence of the given state. The latter is due to statistical averaging which accounts for the main structural-dynamic peculiarities of a state investigated in the description of the system with the given arrangement of particles.

In principle, due to its generality, the virial theorem allows one to describe not only an old state but also a new one arising beyond the point of the structural instability, and to build up a multiphase theorem. On its basis the model ideas about the microscopic mechanisms of the phase transitions for systems with different interatomic interaction can be consistently developed. Up-to-date achievements in the application of the virial equation of state to the interpretation of experimental data from the point of view of the interatomic interactions for macromolecules, polymers and dense gases [3, 22] suggest that the virial approach may be utilized for investigating the properties of different states of a substance including a crystalline one.

Acknowledgement

The authors are grateful to N. M. Plakida, Vl. K. Pershin, S. A. Skopinov and L. A. Fishbein for their interest in the paper, for useful discussions of the results, and for their critical comments.

References

1. N. M. Plakida, Report JINR R4-3930, Dubna, 1968.
2. L. D. Landau and E. M. Lifshits, Mechanics, Nauka, Moscow, 1958, p. 35.
3. E. A. Mason and E. M. Spurling, The Virial Equation of State. International Encyclopedia of Physical Chemistry and Chemical Physics, Vol. 2, Ed. J. S. Rowlinson, Springer-Verlag, Berlin, 1967.
4. V. P. Mineev, Zh. exper. teor. Fiz., *67*, 1894, 1974.
5. H. E. Stanley, Phys. Rev. Lett., *17*, 913, 1966; *20*, 589, 1968.
6. T. A. Kaplan, Phys. Rev. Lett., *44*, 760, 1980.
7. V. L. Berezinski, Zh. exper. teor. Fiz., *59*, 907, 1970.
8. J. H. Williams and G. Joyce, J. Chem. Phys., *59*, 741, 1973.
9. K. Parlinsky, A. C. Mitus, R. Sikora and T. Wasiutinsky, J. Chem. Phys., *67*, 5366, 1977.
10. N. M. Plakida and T. Siklós, Phys. Lett., *26A*, 342, 1968.
11. J. C. Raich and R. D. Etters, J. Chem. Phys., *55*, 3901, 1971.
12. A. I. Anselm, Osnovy statisticheskoj fiziki i termodinamiki, Nauka, Moscow, 1973, p. 171.
13. N. M. Plakida, in: Statisticheskaya fizika i kvantovaya teoriya polya, Nauka, Moscow, 1973, p. 206.
14. G. Liebfried and N. Breuer, Point Defects in Metals. Introduction to the Theory, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
15. T. Matsubana and K. Kamiga, Progr. Theor. Phys., *58*, 767, 1977.
16. G. Korn and T. Korn, Spravochnik po matematika, Nauka, Moscow, 1977, p. 43.
17. H. L. Pina, Portugaliae Physica, *10*, 197, 1979.
18. V. K. Pershin and Vl. K. Pershin, Phys. Stat. Sol., *92b*, 9, 1979.
19. V. K. Pershin, Vl. K. Pershin and L. A. Fishbein, Acta Phys. Hung., *48*, 281, 1980.
20. G. Liebfried and W. Ludwig, Theory of Anharmonic Effects in Crystals, Eds. F. Seitz and D. Turnbull, Academic Press Inc., New York—London, 1961.
21. V. C. Shrivastava, Phys. Rev., *B21*, 2602, 1980.
22. C. Domb and P. J. Ratcliffe, J. Phys. A: Math. and Gen., *14*, L373, 1981.

Acta Physica Hungarica 57 (1—2), pp. 113—118 (1985)

THE EXPECTATION VALUE OF THE INELASTIC LOSS IN HEAVY ATOM COLLISIONS FOR keV ENERGIES

I. NAGY, J. LÁSZLÓ and J. GIBER

Physical Institute of the Technical University
1521 Budapest, Hungary

(Received 22 August 1983)

Inelastic energy losses in single collisions are calculated according to the statistical theory of Firsov. The interaction between atoms is described by a potential including additional terms for exchange and correlation energy. A new form of the electronic stopping cross-section is determined.

Theory

The determination of the inelastic energy loss is a fundamental question in many fields, e.g. in ion penetration and slowing down in solids, ion reflection, collision of gas atoms with an ion beam. The Firsov statistical model [1] is used for the description of the collisions of heavy atoms in the keV energy range [2, 3]. If the problem of the binary collision event is not restricted to the approximation where the trajectory seems to be rectilinear, the Kishinevsky model [4] serves as an appropriate description.

An essential question is the choice of the shape of the interaction potential. The theories and the computer codes (e.g. [5]) apply a potential with the Molière screening function as the best approximative formula [6]. However, according to the calculations of Biersack and Ziegler [7], the Molière potential underestimates the real values at larger separations between atoms. Therefore in our model we use the interaction potential determined in [7], whose exchange and correlation energy terms can result in significant contributions in the case of larger distances from the atom. These contributions become essential at lower energy collisions. The Biersack potential is not universal any more, as the screening function depends on the atomic numbers due to these terms, in contrast to the usual Thomas—Fermi description and its approximations so far published [8, 9]. If the distances are expressed in terms of the Thomas—Fermi radius of the bombarding atom ($a_{\mathrm{TF}} = 0.8853\, a_0 Z_1^{-1/3}$), and the values of the atomic constants are substituted, the general formula of the energy loss is the following:

$$\varepsilon(\mathrm{eV}) = K \left\{ \int_{x_0}^{\infty} \left[1 - \frac{A}{\underline{E}_1} \frac{\Phi(x)}{x} \right] \left[1 - \frac{A}{\underline{E}_1} \frac{\Phi(x)}{x} - \left(\frac{b}{x} \right)^2 \right]^{-1/2} \int_x^{\infty} \frac{\Phi^*(u)}{u} du \right\} dx \,, \quad (1)$$

where we have used the notations

$$A(\text{eV})=30.724\frac{m_1+m_2}{m_2}Z_1^{4/3}Z_2,$$

where Z_1 and Z_2 are the atomic numbers, m_1 and m_2 the corresponding masses, b the impact parameter and $\underline{E}_1$ the bombarding energy in eV.

Then

$$\Phi(x)=\{[\varphi_1(x)-Z_1^{-2/3}\varphi_2(x)]\,[\varphi_1(tx)-Z_2^{-2/3}\varphi_2(tx)]\}^{1/2},$$

$$\Phi^*(u)=[\varphi_1(u/2)-Z_1^{-2/3}\varphi_2(u/2)]\,[\varphi_1(tu/2)-Z_2^{-2/3}\varphi_2(tu/2)],$$

$$t=(Z_2/Z_1)^{1/3}$$

and

$$K=0.04875(Z_1+Z_2)^2Z_1^{-1/3}(E_1/m_1)^{1/2}.$$

The Biersack screening functions:

$$\varphi_1(y)=0.09\exp(-0.19y)+0.61\exp(-0.57y)+0.3\exp(-2y)$$

$$\varphi_2(y)=0.07\exp[-(0.295/7y)^2-(0.295y/4)-(0.295y/7)^2]$$

and y can be $u/2$; $tu/2$ and x. The closest approach (x_0) is the root of the following equation:

$$1-\frac{A}{\underline{E}_1}\frac{\Phi(x)}{x}-\left(\frac{b}{x}\right)^2=0. \tag{2}$$

Karpuzov discussed the general loss formula of Kishinevsky in detail [10] considering the possible forms of the screening functions. Our calculations show that there is only one function for which the double integral procedure defined in the general formula (1) is analytical up to the end. This is one of the approximative forms of Firsov [11]:

$$\varphi_1(y)=\alpha/y \qquad \text{and} \qquad \varphi_2(y)\equiv 0.$$

Considering that the models till now expressed the linear dimensions in terms of the Thomas—Fermi radius of the two atoms ($a'_{\text{TF}}=0.8853a_0(Z_1^{1/2}+Z_2^{1/2})^{-2/3}$), the final shape of the loss with this screening is:

$$\varepsilon(\text{eV})=K'\alpha^2\frac{\pi}{2}(B+2b^2)(B+b^2)^{-3/2}, \tag{3}$$

where

$$K'=0.04875(Z_1+Z_2)^2(Z_1^{1/2}+Z_2^{1/2})^{2/3}(E_1/m_1)^{1/2},$$

$$B=\alpha\,30.724\,Z_1Z_2(Z_1^{1/2}+Z_2^{1/2})^{2/3}\frac{m_1+m_2}{m_2}(E)^{-1}.$$

b is in a'_{TF} units and in the case of the Firsov approximative form $\alpha=0.45$. This result is really valid for the class of the $V(r)\sim r^{-2}$ type interaction potentials.

Let us discuss the behaviour of the closed formula (3).

a) It is easy to see that the expression (3) as an $\varepsilon(b)$ function has a maximum at $b=(B/2)^{1/2}$

b) If the bombarding energy (E_1) is sufficiently small and b is not too large, then $\varepsilon \sim E_1$.

If E^1 is sufficiently large, we can get back the characteristic energy dependence of the Lindhard theory [12], where $\varepsilon \sim (E_1)^{1/2}$. The features mentioned in a) and b) are valid in the cases of more exact screening functions if not with the same numerical values.

c) Let us mention here that the usual electronic stopping formula:

$$S_e = 2\pi \int_0^\infty \varepsilon(b) b\, db$$

would prove infinite for an application according to (3). Carrying out the integral procedure till an upper limit b_H its limes, if $b_H \to \infty$ is:

$$\lim_{b_H \to \infty} \{S_e = K'\alpha^2\pi^2[(3B+2b_H^2)(B+b_H^2)^{-1/2} - 3(B)^{1/2}]\} = \infty .$$

Thus, the interaction potential has to slow down faster than by r^{-2}. This condition is fulfilled either by the Molière or by the Biersack potential.

Results and discussion

We performed the integral procedure expressed in (1) on several materials over a wide energy range. Figs 1a and 1b show the results obtained on Al—Al interaction vs the impact parameter. For a comparison we plotted the results for Al—Al, Cu—Cu and Au—Au interactions together at a certain bombarding energy ($E_1 = 10$ keV) in Fig. 2 as a function of the common impact parameter calculated in nm units.

We examined the development of the elastic and inelastic loss through a series of collisions, i.e. in a cascade where the impact parameter was randomly chosen at each collision. The results of this work and the detailed atomic number dependence of the inelastic loss will be published soon.

We applied a new assumption for the calculation of the electronic stopping power. Supposing a random structured target we considered the probability density function of the impact parameter values according to Sanders and Roosendaal [13].

The distribution function is the following:

$$f(b) = \frac{32}{3}\pi N^2 \int_0^\infty (x^2+b^2)^2 \exp\left[-\frac{4}{3}\pi N(x^2+b^2)^{3/2}\right] dx\,, \tag{4}$$

where x and b are in $\frac{\text{nm}}{10}$, N is the target density in $\left(\frac{\text{nm}}{10}\right)^{-3}$. The probability density

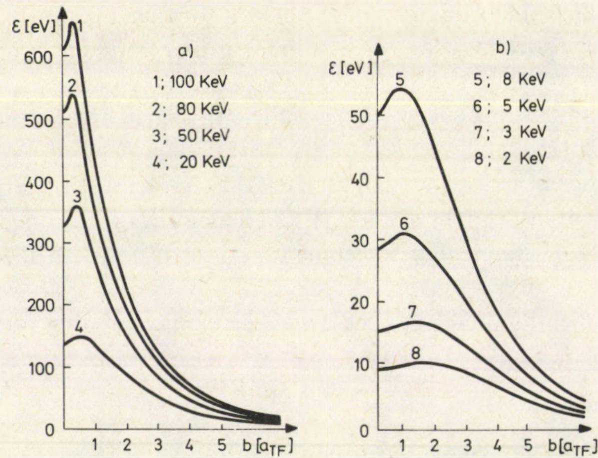


Fig. 1. The inelastic energy loss for Al—Al interactions at different bombarding energies as a function of the impact parameter a) above 10 keV; b) below 10 keV

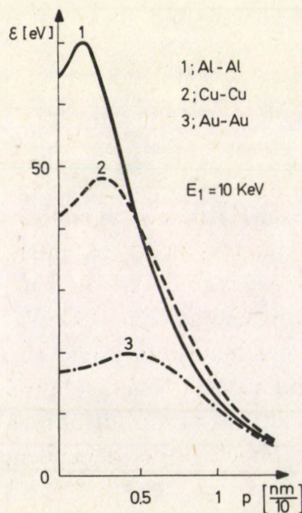


Fig. 2. The inelastic energy loss for Al—Al; Cu—Cu; Au—Au interactions at $E_1 = 10$ keV energy vs the real impact parameter p

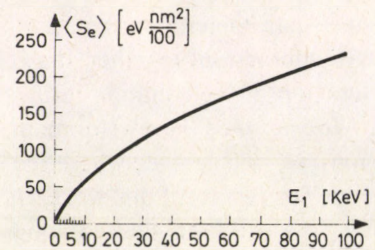


Fig. 3. The expectation values of the electronic stopping $\langle S_e \rangle$ vs the bombarding energy E_1

function is normalized to 1. We determined the expectation value of the electronic stopping with the help of (4):

$$\langle S_e \rangle = 2\pi \int_0^\infty bf(b)\, \varepsilon(b)\, db \tag{5}$$

at fixed E_1 energies.

The dimension of $\langle S_e \rangle$ will be $\mathrm{eV}\,\frac{\mathrm{nm}^2}{100}$, if b is in a_{TF} unit. The result of the calculation is shown in Fig. 3 for an Al—Al interaction as a function of the bombarding energy (E_1).

The curve of the function is remarkable. In the low energy range $\langle S_e \rangle \sim E_1$, then with increasing energy the feature of $\langle S_e \rangle \sim (E_1)^{1/2}$ becomes more effective. A similar characteristic behaviour was found by Blume et al [14] in the evaluation of their experiment bombarding an Au target by light ions.

This electronic stopping cross section can be connected to the usual form of the stopping power:

$$\left(\frac{dE}{dx}\right)_e = N\langle S_e \rangle , \tag{6}$$

where N is the target density.

The dependence of the electronic stopping on the bombarding energy mentioned above shows a close connection to other fields, too. Svenson et al [15] measured the number of ejected electrons due to ion bombardment vs bombarding energy. The target was Ag, the bombarding ions were Se^+ and Te^+. The energy dependence of the number of ejected electrons shows the same behaviour, as the $\langle S_e \rangle$ vs energy function calculated by us. This conclusion allows us to suppose a deterministic connection between the phenomena.

Finally, we have to mention that the models used according to the original conception of Firsov, i.e. the models where the trajectories are rectilinear, follow simply from formula (1), if b and E_1 are sufficiently large. In these cases the analytical shape of the screening functions characterizes the inelastic loss unambiguously, i.e. considering exponential screenings the integral exponent appearing in the inner integral can be approximated by the values of the arguments at corresponding places. For example, in the case of the Molière potential the behaviour will be the following:

$$\varepsilon \sim (E_1)^{1/2} \exp(-0.3\, x_0)$$

at approaches (x_0) far from the target atom. Here we used the fact that the dominant term of the Molière screening is the $\exp(-0.3x)$ one.

References

1. O. B. Firsov, Sov. Phys. JETP, *36*, 1076, 1959.
2. A. H. El-Hosby and J. F. Gibbons, Phys. Rev., *173*, 454, 1968.
3. C. P. Bhalla, J. N. Bradford and G. Reese: in Atomic Collisions Phenomena in Solids, ed.: H. W. Thompson, North-Holland, Amsterdam, 1970.
4. L. M. Kishinevsky, Bull. Acad. Sci. USSR (Phys. Ser.), *26*, 1433, 1962.
5. O. S. Oen and M. T. Robinson, Nucl. Instrum. Methods, *132*, 647, 1976.
6. G. Molière, Z. Naturforsch., *2a*, 133, 1947.

7. J. P. Biersack and J. F. Ziegler, Nucl. Instrum. Methods, *194*, 93, 1982.
8. I. M Torrens, Interatomic Potentials, Acad. Press, New York, 1972.
9. P. Gombás, Die Statistische Theorie des Atoms und ihre Anwendungen, Springer, Wien, 1949.
10. D. S. Karpuzov, Appl. Phys., *24*, 121, 1981.
11. O. B. Firsov, Sov. Phys. JETP, *34*, 308, 1958.
12. J. Lindhard, M. Scharff and H. E. Schiott, K. Dan. Vidensk. Selsk. Mat.—Fis. Medd., *33*, No. 14, 1963.
13. J. B. Sanders and H. E. Roosendaal, Nucl. Instrum. Methods, *132*, 267, 1976.
14. R. Blume, W. Eckstein and H. Veerbeck, Nucl. Instrum. Methods, *194*, 67, 1982.
15. B. Svensson, G. Holmen and J. Linnors, Nucl. Instrum. Methods, *194*, 429, 1982.

Acta Physica Hungarica 57 (1—2), pp. 119—123 (1985)

RADIATIVE SELF-INTERACTION IN CLASSICAL FIELD THEORY

P. KÁLMÁN

Department of Experimental Physics, Institute of Physics, Technical University
1521 Budapest, Hungary

(Received 13 October 1983)

The electron and the electromagnetic field are described in terms of classical field theory. In order to secure that the observed rest mass would also contain the electromagnetic part a constraint is used during variation. As a consequence, radiative self-interactions arise which produce spontaneous emission.

Several authors have dealt with the interaction of quantum mechanical systems and radiation from different points of view, treating the problem in semiclassical manner [1]. Besides quantum electrodynamics (QED) these semiclassical theories also have importance from practical and general points of view. They are very useful in laser physics [2] and nonlinear optics [3] and can help understand what specific features of QED might be modified.

Our aim is to make clear what kind of results the mass renormalization — an inherent part of QED — can produce without canonical quantization and thus to attempt to construct a classical theory of electromagnetic self-interaction.

We describe the electron by the complex bispinor field $\psi(x)$ and the electromagnetic field by the four-vector field $A^\mu(x)$. The total Lagrange density $\mathscr{L}_T$ has four parts

$$\mathscr{L}_T = i\hbar c\, \bar{\psi}\gamma^\mu \partial_\mu \psi - Mc^2 \bar{\psi}\psi - J^\mu A_\mu - \frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}. \tag{1}$$

The first and second terms describe the free electron without any interaction, the third term gives the interaction between the charged matter and the electromagnetic field and the fourth one accounts for the free electromagnetic field. The four-current J^μ is the sum of the electron current $J_e^\mu = -e\bar{\psi}\gamma^\mu\psi$ and the current J_s^μ of the surroundings. The constant M describes the electron mass of non-electromagnetic origin [4].

The corresponding Euler—Lagrange equations are obtained by means of Hamilton's variational principle from the Lagrange function L which is the integral of (1) over the whole volume. In this way we get the coupled Dirac and Maxwell equations

$$[i\hbar c\gamma^\mu \partial_\mu - Mc^2 + e\gamma^\mu A_\mu]\psi = 0, \tag{2}$$

$$\Box A^\mu = -4\pi J^\mu, \tag{3}$$

while the Lorentz condition is imposed on the electromagnetic four-potential. Since one part of $A^{\mu}(x)$ is produced by the electron current, a self-interaction arises that strongly influences the solution of the Dirac equation. The statical part of the self-interaction completely destroys the energy spectrum of the hydrogen atom. This is thought to be one of the main problems of the semiclassical theory. We can, however, get rid of this type of self-interaction.

From (1) the energy density of a free and stationary electron is

$$H_T=\psi^+(-i\hbar c\,\boldsymbol{\alpha}\nabla+\beta Mc^2)\psi+J^{\mu}_e A_{\mu e}+\frac{1}{16\pi}F_{\mu\nu e}F^{\mu\nu}_e, \tag{4}$$

where the indices e refer to the electron. The total energy is the volume integral of (4). It appears obvious that this total energy ought to be equal to the observed rest mass m of the electron when it is free and at rest i.e.

$$\int[\bar{\psi}\psi Mc^2+J^{\mu}_e A_{\mu e}+\frac{1}{16\pi}F_{\mu\nu e}F^{\mu\nu}_e]d^3x=\int\bar{\psi}\psi mc^2 d^3x. \tag{5}$$

Equation (5) plays the role of a constraint that leads, however, to nonsensical equations according to which the electron behaves like a chargeless particle. This difficulty can be avoided if the electromagnetic mass is described separately by introducing a further vector field b^{μ} and an antisymmetrical tensor $H^{\mu\nu}=\partial^{\mu}b^{\nu}-\partial^{\nu}b^{\mu}$ in the constraint. With these new variables it has the following form

$$\int[\bar{\psi}\psi Mc^2+J^{\mu}_e b_{\mu}+\frac{1}{16\pi}H_{\mu\nu}H^{\mu\nu}-\bar{\psi}\psi mc^2]d^3x=0. \tag{6}$$

As b^{μ} and M describe the electromagnetic and non-electromagnetic parts of the mass we consider them as dynamical variables.

We deduce new Euler—Lagrange equations adding the left side of (6) multiplied by λ to the original Lagrange function L where λ is an indeterminate multiplier and then vary the action constructed from this new Lagrangian. The modified equations are

$$[i\hbar c\,\gamma^{\mu}\partial_{\mu}+(\lambda-1)Mc^2-\lambda mc^2+e\gamma^{\mu}(A_{\mu}-\lambda b_{\mu})]\psi=0, \tag{7}$$

$$\square A^{\mu}=-4\pi J^{\mu}, \tag{8}$$

$$\square b^{\mu}=-4\pi J^{\mu}_e, \tag{9}$$

$$\lambda=1. \tag{10}$$

The Lorentz condition is required again for the fields A^{μ} and b^{μ}. Equation (10) is obtained by the variation of M.

We can solve Eqs (8) and (9) with the aid of appropriate Green functions. The field A^{μ} is built up as usual in the classical theory, with the retarded Green function $D_{\rm ret}=2\Theta(x_0)\delta(s^2)$, where δ and Θ denote the Dirac delta and the step functions, furthermore the abbreviation $s^2=x_0^2-\mathbf{x}^2$ is used. But if we want to give b^{μ}, then we face

the problem how to separate the own field of the electron, which builds up its electromagnetic mass, from the radiation field which leaves the electron and exists as a free one in far-away space regions. It is obvious that the electron self-energy must not contain the energy of such a radiation even if it was radiated by the electron. This can be achieved if the advanced Green function $D_{adv} = 2\Theta(-x_0)\delta(s^2)$ is used to determine b^μ. Putting (10) into (7), denoting by A_s^μ and A_e^μ the different parts of A^μ originating from the surroundings and the electron respectively ($A^\mu = A_s^\mu + A_e^\mu$) and introducing $C^\mu = A_e^\mu - b^\mu$ we get

$$[i\hbar c\,\gamma^\mu\partial_\mu - mc^2 + e\gamma^\mu A_{\mu s} + e\gamma^\mu C_\mu]\psi = 0. \tag{11}$$

The variable M vanishes in (11) and the observed rest mass m remains only. Using the Green functions

$$C^\mu = \int [D_{ret} - D_{adv}] J_e^\mu d^4x. \tag{12}$$

As the difference $D_{ret} - D_{adv} = D$ fulfils the homogeneous equation $\Box D = 0$, and so does C^μ, the statical self-interaction vanishes in (11), too, and the remaining one is of radiative type only.

We can rewrite (11) in the form

$$i\hbar c\,\partial_0\psi = (H_0 + H_I)\psi, \tag{13}$$

where $H_0 = -i\hbar c\,\boldsymbol{\alpha}\nabla + \beta mc^2 - e\gamma^0\gamma^\mu A_{\mu s}$ and $H_I = -e\gamma^0\gamma^\mu C_\mu$. Denoting the eigenvalues and eigenfunctions of H_0 by E_n, φ_n the time-dependent solution ψ is looked for in the form $\psi = \sum_n c_n\varphi_n$. Introducing $\sigma_{sr} = c_s c_r^*$ we obtain from (13)

$$i\hbar\,\partial_t\sigma_{sr} = \sum_j (H_{I,sj}\sigma_{jr} - H_{I,jr}\sigma_{sj}) + (E_s - E_r)\sigma_{sr}, \tag{14}$$

where $H_{I,ij}$ denotes the i,j matrix element of H_I. Expanding into power series both the retarded and advanced solutions one gets C^μ as a series of increasing time derivatives of σ_{lk}-s which leads to the following form of $H_{I,ij}$

$$H_{I,ij} = \sum_{n=0}^{\infty} \frac{2\hbar\alpha}{(2n+1)!} \int\int \bar{\varphi}_i\gamma^\mu\varphi_j \frac{r^{2n}}{c^{2n}} \sum_{k,l} \bar{\varphi}_k\gamma^\mu\varphi_l d^3x d^3y \frac{\partial^{2n+1}}{\partial t^{2n+1}}\sigma_{lk}, \tag{15}$$

where α is the fine structure constant and $r = |\mathbf{x} - \mathbf{y}|$. Thus (14) is a system of coupled, nonlinear differential equations for the σ_{sr}-s. As an example we can examine a two-level system with energy separation $E_1 - E_2 = \hbar\omega_{12}$. In the nonrelativistic dipole approximation we can get an asymptotic form of (14) valid when one of the two states dominates, e.g. $\sigma_{22} \ll \sigma_{11} \cong 1$

$$i\hbar\partial_t a_{12} = i\frac{4e^2}{3c^3}\omega_{12}^3|\mathbf{x}_{12}|^2 a_{12}, \tag{16}$$

where a_{12} is defined by $\sigma_{12} = a_{12}\exp(-\omega_{12}t)$, $\mathbf{x}_{12}$ is the dipole matrix element and double frequency terms, derivatives of order higher than three and terms with more than one of their indices different from 1 are neglected. Equation (16) has a solution of

form $a_{12}=A_{12}\exp(\Lambda_{12}t)$ where $\Lambda_{12}=4e^2\omega_{12}^3|\mathbf{x}_{12}|^2/(3\hbar c^3)$ is the Einstein coefficient. If $E_1>E_2$, then $\Lambda_{12}>0$ and the amplitude of the oscillation of frequency ω_{12} increases exponentially until (16) loses its validity and the electron goes to the state 2. If $E_1<E_2$ then $\Lambda_{12}<0$ and a_{12} decreases, the small disturbance a_{12} vanishes exponentially and no transition takes place to the state lying higher than the initial one. The fact that Λ_{12} can have positive and negative values suggests that J_e^μ can only be interpreted as a real four-current.

There is, however, another choice of the Green function for b^μ. As the negative frequency Green function produces solutions corresponding to incoming waves, therefore it is also appropriate for the building up of b^μ. This choice also leads to a radiative type self-interaction which is determined by

$$C^\mu=\int(D_{\text{ret}}-\bar{D}_c)J^\mu(x)d^4x, \tag{17}$$

where $\bar{D}_c$ denotes the negative frequency Green function which has the form $\bar{D}_c=\delta(s^2)+i/(\pi s^2)$. Using the concrete forms of the Green functions we get $D_{\text{ret}}-\bar{D}_c=D_-$ where $D_-=D/2+iD_1/2$ with $D_1=-2i/(\pi s^2)$ which obeys the homogeneous equation $\Box D_1=0$ and so does D_-.

It is easy to show that the time-dependent solution influenced by this type of self-interaction is governed by an equation similar to (14) but containing an other perturbing operator H_I' instead of H_I. The new operator H_I' is defined by $H_I'=(H_I+h_I)/2$ where the s, r matrix element of h_I has the form

$$h_{I,sr}=-\frac{2ie^2}{\pi}\int\bar{\varphi}_s\gamma^\mu\varphi_r\sum_{l,k}\sigma_{lk}(x_0)\bar{\varphi}_k\gamma_\mu\varphi_l\frac{dx_0d^3xd^3y}{\tau^2-r^2} \tag{18}$$

with the notation $\tau=ct-x_0$.

In the case of a two-level system Eq. (14) modified in this manner also leads to an asymptotical equation valid when $\sigma_{11}\gg\sigma_{22}$. This equation is the same as (16) and has the same solution when $E_1>E_2$ but it has zero on its right side when $E_2>E_1$. Thus it produces nonzero solution in the case $E_1>E_2$ only.

In order to make clear the connection between this second type of self-interaction and that in quantum electrodynamics we examine it in more detail. One can obtain level shifts and decay rates of states using Feynman's method [5] to solve the propagation kernel of Eq. (11) to first order in the presence of perturbation (17)

$$\Delta E=e^2\int\varphi_0\gamma^\mu K_v(2,1)\gamma_\mu\varphi_0 D_-(2,1)d^4x_2d^3x_1, \tag{19}$$

where $K_v(2,1)$ denotes the propagation kernel in the presence of an external potential A_s^μ without any perturbation. It can be shown that in the nonrelativistic approximation Eq. (19) is identical with the quantum electrodynamical expression of ΔE. The latter differs from (19) in that it contains the positive frequency Green function D_c instead of D_-. In the nonrelativistic approximation $K_v(2,1)$ differs from zero when $t_2>t_1$. As $D_c=\Theta(t_2-t_1)D_- -\Theta(t_1-t_2)D_+$ we can see that $K_vD_c=K_vD_-$ and therefore Eq. (19) equals the nonrelativistic, non-renormalised quantum electrodynamical expression of

the level shift and decay rate. Therefore, the second type self-interaction seems to be more appropriate than the first one. Furthermore in this case the decay rate results, in spite of the classical field theoretical foundation, admit of the probability interpretation of the wave function.

Thanks are due to T. Keszthelyi for valuable remarks and discussions.

References

1. L. I. Schiff, Quantum Mechanics, McGraw-Hill Book Co., Inc. 1949, Chapter X. p. 240; H. A. Kramers, Quantum Mechanics, North-Holland Publ. Co. Amsterdam, 1957, Chapter VIII; E. T. Jaynes, Coherence and Quantum Optics, Ed. by L. Mandel and E. Wolf, Plenum Press, New York, 1973, pp. 35.
2. E. T. Jaynes and F. W. Cummings, Proc. IEEE, *51*, 89, 1963; M. D. Crisp and E. T. Jaynes, Phys. Rev., *179*, 1253, 1969; L. Allen and J. H. Eberly, Optical Resonance and Two-Level Atoms, John Wiley & Sons, New York, 1975.
3. J. A. Armstrong, N. Bloembergen, J. Ducuing and P. S. Pershan, Phys. Rev., *127*, 1918, 1962; N. Bloembergen, Nonlinear Optics, W. A. Benjamin Inc. New York, 1965.
4. Our notation corresponds to that of J. D. Bjorken and S. D. Drell, Relativistic Quantum Mechanics, McGraw-Hill Book Co., New York, 1964.
5. R. P. Feynman, Quantum Electrodynamics, W. A. Benjamin Inc., New York, 1961; Phys. Rev., *76*, 749, 1949; *ibid* 769.

Acta Physica Hungarica 57 (1—2), pp. 125—129 (1985)

MÖSSBAUER EFFECT STUDY OF LIBYAN DESERT SILICA GLASS

H. A. SALLAM

Physics Department, Faculty of Science, Islamic College for Women
Nasr-City, Cairo, Egypt

N. A. EISSA

Physics Department, Faculty of Science, Al-Azhar University
Nasr-City, Cairo, Egypt

I. DÉZSI and D. L. NAGY

Central Research Institute for Physics, Budapest, Hungary

(Received in revised form 25 October 1983)

The Mössbauer effect and X-ray fluorescence techniques were applied to study Libyan Desert Silica Glass. The X-ray fluorescence proved the existence of 13 elements ($Z>20$) in each sample of 5 glass types with differences in their relative abundance. The Mössbauer effect spectra showed a complex two spectral lines. Computer analysis gave Mössbauer effect parameters to those of iron silicate minerals, but no Fe^{3+} ions are detected. It is concluded that this Libyan Silica Glass is one type of tektites formed by the impact of a huge meteorite on the earth's surface in that area.

Introduction

Libyan Desert Silica Glass (L.D.S.G.) is a natural silica glass discovered [1] in 1932 by the Egyptian geological survey expedition in the Western Desert of Egypt (North East of Africa and about 480 miles South-West of Cairo, lat.: 25° 30′—longt.: 25°). It is dated to about 38 million years ago and it is considered by some authors [2—6] a tektite, but it is found in much larger pieces and in greater quantities than any tektite yet known. It showed a variety of types classified [7] according to their colour and appearance, as: 1. transparent; 2. translucent; 3. milky; 4. vuggy; 5. xenilithic; 6. bonded and 7. blackish. Chemical analysis [8] of the blackish glass showed the following chemical composition:

SiO_2	98.20%
TiO_2	0.32
Al_2O_3	0.70
FeO	0.53
Fe_2O_3	0.24
NiO	0.02

MgO	0.01
CaO	0.30
Na_2O	0.33
K_2O	0.02
H_2O	0.06
Total	100.73%

Studies [2, 7] were undertaken concerning its features, locality and chemical composition. The study of the ratio Fe^{3+}/total iron was made [9] using the electron paramagnetic resonance in order to identify its origin.

The origin of this glass is still unsettled, as to whether it is a true tektite formed by the fall of a larger meteorite or an earthy natural glass. Giengengack and Al-Far [10] think that it retains significant terrestrial origin.

The aim of the present work is to apply the Mössbauer effect and X-ray fluorescence techniques for the first time to study this glass in an attempt to identify its origin.

Experimental

For the measurement of the Mössbauer spectrum an automatic Mössbauer spectrometer in constant acceleration mode was used. The 14.4 keV γ-rays were detected by a Xe-filled proportional counter having a Be window with no iron impurity. The sample weight was 391.3 mg and the iron content was about 1.28 mg/cm^2. A ^{57}Co(Cr) single line source was used. The isomer shift values are given relative to natural iron.

The analysis of the spectra was made by using our least squares fitting program.

Results and discussion

The results of the X-ray fluorescence analysis for the five types of the L.D.S.G. mentioned above are summarized in Table I. It can be observed that all types contain the same elements with slight differences in concentration.

The Mössbauer effect measurements were carried out on the blackish type only since it contains the highest percentage of iron. Fig. 1 shows the Mössbauer spectrum of this glass sample at room temperature. Computer analysis of this broad and asymmetric two line spectrum indicates the existence of two different quadrupole split doublets arising from two different electric field gradients and indicating that there are at least two inequivalent sites for the iron in this glass. Analysis of this spectrum gave the Mössbauer effect parameters listed in Table II.

Table I

The X-ray fluorescence results for five L. D. S. G. samples

Sample No.	1 Vuggy	2 Milky	3 Transparent	4 Blackish	5 Translucent	Error
Ti	850	360	720	900	320	+400 −600
V	310	–	260	–	260	±200
Cr	560	360	390	620	460	+150 −200
Mn	780	520	470	390	500	+100 −200
Fe	2400	2960	2500	5200	3500	±150
Co	320	340	270	165	245	±70
Ni	127	113	108	65	120	±40
Cu	130	120	365	247	156	±40
Zn	115	48	329	274	186	±30
Pb	14	12	14	22	16	±8
Sr	16.8	24.4	28	24.6	30.5	±3
Zr	192	240	266	200	136	±3
Y	35	43.5	33	39.5	27	±6

1 Practically above $Z > 20$ atomic number.
2 The concentrations are given in ppm ($10^{-4}\%$).

The values of the Mössbauer effect parameters are characteristic of Fe^{2+} ions only. The chemical analysis indicated a concentration of 0.24% as Fe_2O_3 but the analysis of the Mössbauer effect spectrum did not indicate the presence of Fe^{3+} peaks. Hovewer, because of the very low iron content the statistical error is rather high and the

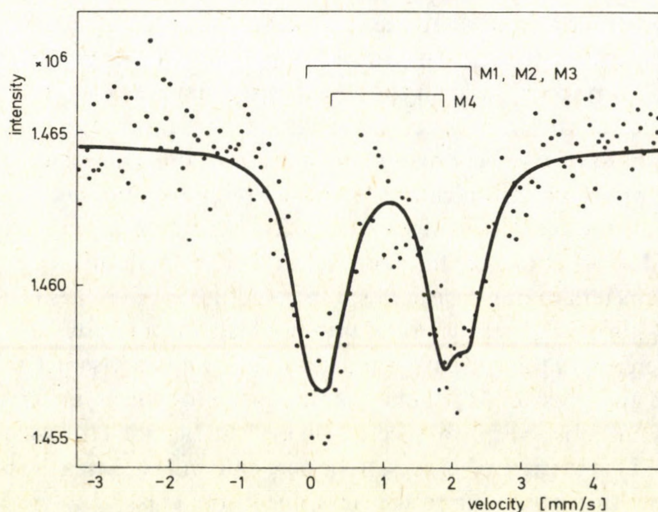


Fig. 1. Mössbauer spectrum of Libyan Desert Silica Glass sample measured at room temperature

Table II

ME parameters of the L. D. S. G. and some iron-bearing silicate minerals. δ: isomer shift; Δ: quadrupole splitting; Γ: line width, I: the relative Mössbauer intensity of the two components

Sample material	Lattice site	I. S. [mm/s]	Q. S. [mm/s]	Γ [mm/s]	I [%]	Reference
Libyan Desert Silica Glass	M_1, M_2, M_3	1.10 ± 0.02	2.30 ± 0.07	0.59 ± 0.05 assumed to be	45.3	Present work
	M_4	1.05 ± 0.02	1.56 ± 0.06	equal for each line	54.7	
Cummingtonite-grunerite $(Fe, Mg, Mn)_7Si_8O_{22}(OH)_2$	M_1 M_2 M_3	$1.14 - 1.18$	$2.76 - 290$		generally 50	11
	M_4	$1.05 - 1.11$	$1.50 - 1.68$		generally 50	
Anthophyllite $(Mg, F)_7Si_8O_{22}(OH)_2$	M_1 M_2	$1.12 - 1.13$	$2.58 - 2.61$		generally 50	11
	M_3	$1.09 - 1.11$	$1.80 - 1.81$		generally 50	
Fe		0	0		4.5	
Elmenite $(FeTiO_3)$		1.05	0.71		7.5	
Tektites		$0.80 - 0.92$	$1.84 - 2.08$	$0.61 - 0.74$ (for peak 1) $0.79 - 1.00$ (for peak 2)		12

presence of some Fe^{3+} ions certainly cannot be excluded, possibly as a small para- or superparamagnetic contribution to the low energy line. During the measurement the ± 11 mm/s velocity range has been scanned (Fig. 1 shows only the central part of the spectrum) and no trace of Zeeman split Fe^{3+} spectrum was observed.

The values of the Mössbauer effect parameters are in agreement with those obtained [11] for some iron bearing silicate minerals (Table II) such as cummingtonite, anthophyllite, pyroxene, in which iron ions occupy more than one site.

Moreover, the measured Mössbauer effect spectrum is very similar to those obtained by Marzolf et al [12] for some tektites. The small difference in the values evaluated parameters of our measured sample and those of tektites (Table II) can be due to the possibilities of resolving the two doublets in our case.

The difference in the values of quadrupole splitting and isomer shift of M_1, M_2 and M_3 group and those of M_4 doublet is the reason for the asymmetry in the width and height of the observed two spectral lines. This asymmetry was explained by Marzolf et al [12] in the case of tektites, to be due to the variations in the bond distances and angles from one iron atom to the next one in the glass.

Conclusion

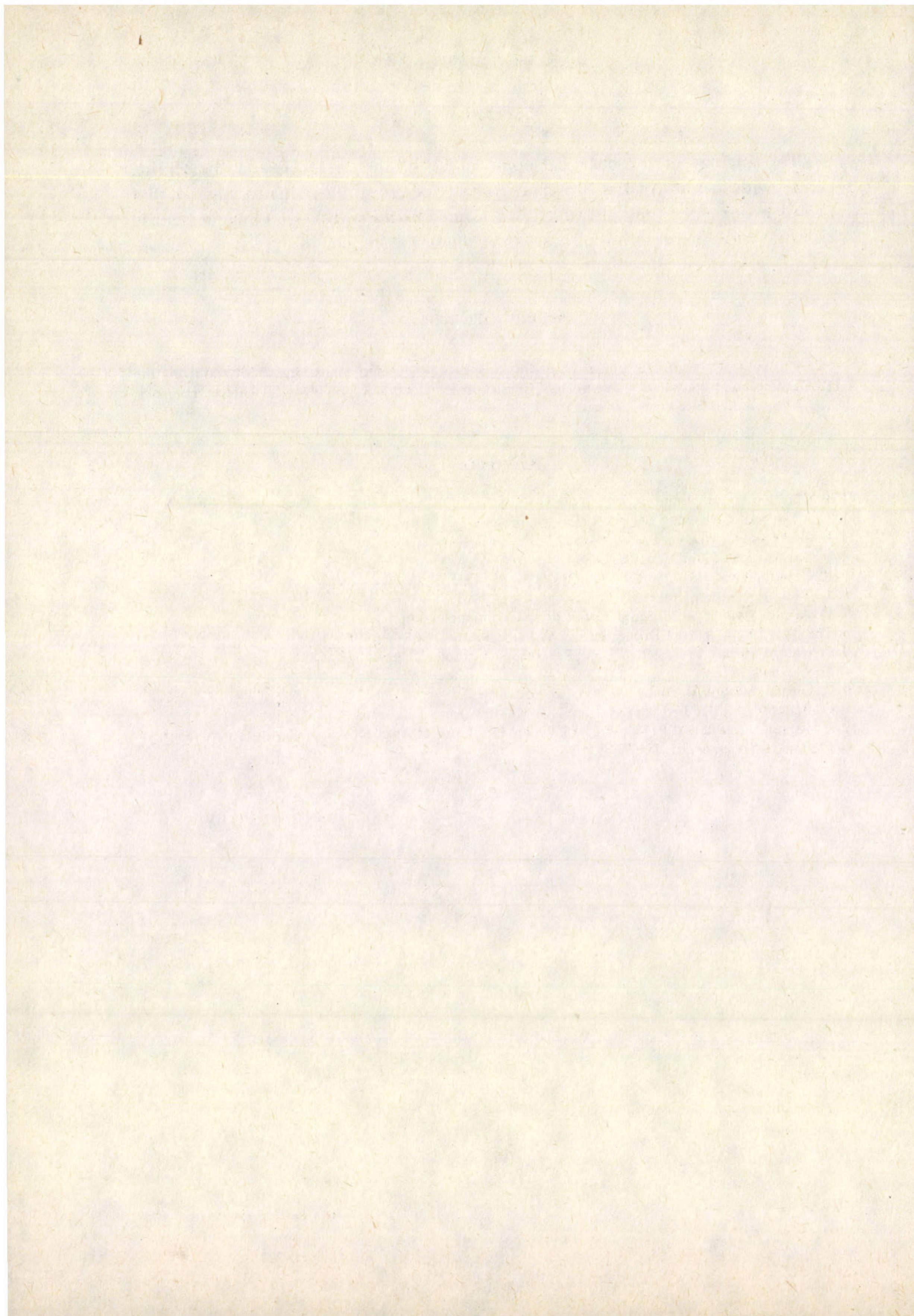
The analogy of the chemical composition of the L.D.S.G. and the tektite, together with identity of the Mössbauer effect spectrum allows us to suggest that L.D.S.G. is one type of tektites which was formed by the impact of a huge meteorite which may have fused with the rock at the site of the impact.

Acknowledgement

The authors are grateful to Dr. D. Al-Far, Geological Museum, Cairo, Egypt for supplying the L.D.S.G. samples. Thanks are also due to Dr. J. Bacsó and Mr. Gikanlinka, ATOMKI, Debrecen Hungary, for the X-ray fluorescence measurements.

References

1. P. Clayton, Nature, London, *132*, 22, 1933.
2. L. J. Spencer, Geol. Mag., *71*, 45, 1933.
3. B. Glass and G. Heezen, Scient. Am., *217*, 32, 1967.
4. V. Barnes, Scient. Am., *205*, 36, 1961.
5. R. Chapman and H. Larson, J. Geophys. Res., *68*, 4305, 1963.
6. J. O'Keefe, Tektites, University of Chicago Press, 1963.
7. D. Al-Far, Geological Museum, Cairo, private communication.
8. The New Encyclopaedia Britannica, *18*, William Benton Publisher, 1976.
9. S. Arafa, A. Bishay, R. Giegengack and R. Weeks, Proc. Amer. Geophys. Union, Ann. Meet., Washington D. C., June, 1975.
10. R. Giegengack and D. Al-Far, Bull. Geol. Soc. Amer., 1974.
11. G. Bancroft, A. Maddoch and R. Burns, Geochim. Cosmochim. Acta, *31*, 2219, 1967.
12. J. Marzolf, J. Dehn and F. Salmon, Mössbauer Effect and its Applications in Chemistry, Advances in Chemistry Series, *68*, 61, 1967.



Acta Physica Hungarica 57 (1—2), pp. 131—138 (1985)

DETERMINATION OF THE MEAN FREE PATH OF ELECTRONS IN SOLIDS FROM THE ELASTIC PEAK

I. A SIMPLIFIED THEORETICAL APPROACH

A. JABLONSKI, P. MROZEK

Institute of Physical Chemistry of the Polish Academy of Sciences
Warszawa, Poland

G. GERGELY, M. MENYHÁRD and A. SULYOK

Research Institute for Technical Physics of the Hungarian Academy of Sciences
1325 Budapest, Hungary

(Received 25 October 1983)

The inelastic mean free path λ of electrons is an important material parameter for quantitative AES and EELS. Instead of the usual methods, requiring a very delicate sample preparation procedure, λ can be determined by elastic peak electron spectroscopy. Beside geometrical parameters, the elastic peak is given by the product of λ and σ_{eff} backscattering cross section. A simple method is described for evaluating experimental elastic backscattering probability results using theoretical σ_{eff} data. σ_{eff} was calculated by integrating the differential electron scattering cross sections assuming a single scattering process and using three types of the atomic potential: the Thomas–Fermi, the Thomas–Fermi–Dirac and the Hartree–Fock models. Tabulated σ_{eff} data are presented for the $Z=2-50$ atomic number and $E_p=1$, 1.5, 2, 2.5 and 3 keV energy ranges. No significant differences were found with these potential models, as well as with those based on Fink's data obtained by computer solution of the Dirac equation. Our tabulated σ_{eff} data complete Fink's works restricted to $E_p \leq 1.5$ keV and to some elements. Our model is extended to compounds and alloys.

Introduction

The knowledge of inelastic mean free path (IMFP) λ of electrons is crucial for quantitative electron spectroscopy. The usual experimental methods for determining λ are rather troublesome since they require the preparation of perfect (hole-free and uniform) thin films (self-supporting or overlayers) with well known thickness. There is an obvious need for a method making routine measurements. This paper is presenting a simplified method for determining λ using elastic peak electron spectroscopy (EPES) described in [1]. The measurements can be made on semiinfinite samples or even on their different spots without requiring the delicate procedures for preparing perfect overlayers. The paper mainly deals with CMA measurements with normal incidence of primary electrons. It is, however, not confined to a CMA analyser. Formulae presented in the paper can be applied to any other electron analyser with considering the angles of excitation, electron reflection and the solid angle of acceptance.

1. Physical bases of the simplified method

Elastic peak electron spectroscopy is measuring the elastic peak $N(E_p)$ in absolute units determining the probability P_e of elastic backscattering of an electron from a solid surface. P_e is determined by the elastic backscattering of electrons within a surface layer of λ thickness. As shown in [1]

$$P_e = N_A \lambda \sigma_B \frac{\cos\alpha}{1+\cos\alpha}, \tag{1}$$

with the notations: N_A = density of atoms in the solid; σ_B = total cross section of elastic backscattering; α = detection angle of the CMA ($\alpha = 42°$).

P_e, σ_B and λ are determined by the energy of primary electrons E_p. In (1) P_e has been determined experimentally by comparing the integrated elastic peak $N(E_p)$ with the integrated spectrum $N(E)$ of secondary and backscattered electrons. The problems of experimental determination of P_e will be discussed in Part II. It is obvious from Eq. (1), with the knowledge of σ_B, that EPES offers a possibility for the experimental determination of λ. During the last years a great number of reliable experimental data have been published on various materials (e.g. [2, 3]). According to Seah, λ is mainly determined by E_p the electron energy but for different materials the IMFP data exhibit a considerable scattering over the "universal" $\lambda(E)$ curve. In the present work an analysis is based on the differential scattering cross section $\left(\frac{d\sigma}{d\Omega} \text{ or } \frac{d\sigma}{d\theta}, \theta = \text{angle of scattering}\right)$.

Abundant theoretical literature is available on the differential scattering cross sections of electrons using various quantum mechanical methods and models. Hereupon merely the fundamental works of Fink et al [4] and Ichimura et al [5] are mentioned. Our present simplified analysis is based on the assumption that electrons associated with the total elastic backscattering suffered only one elastic scattering. The small angle elastic scatterings and multiple elastic scattering events as well are neglected in the present considerations. These assumptions are justified by the small values of P_e($\sim 10^{-2}$ for large Z elements and decreasing drastically for low Z). Let us follow the path of an electron detected by the CMA within the elastic peak. Using a modified form of Eq. (4) in [1]

$$I_c(E_p) \sim I_p 2\pi N_A \sin\alpha \Delta\alpha \int_0^\infty \left(\frac{d\sigma}{d\Omega}\right) \exp\left[-\frac{x}{\lambda} - \frac{x}{\lambda\cos\alpha}\right] dx \tag{2}$$

with the notations: x = the distance below the surface; I_p = primary current; $I_c(E)$ = CMA current; Ω = solid angle; $\Delta\alpha$ = the angular opening of the CMA entrance slit.

Eq. (2) is not confined to the CMA and gives also the angular distribution of elastically backscattered electrons through α. Normalizing and integrating Eq. (2) over

$\alpha=0$ to $\pi/2$ with $\alpha=\pi-\Theta$, (since $\frac{d\sigma(\Theta)}{d\Theta}$, $d\sigma(\Theta)/d\Omega$ are conventionally available) is supplying P_e, the probability of total elastic backscattering:

$$d\sigma=|f(q)|^2 d\Omega=2\pi|f(q)|^2 \sin\Theta\, d\Theta\,; \tag{3}$$

$$P_e=N_A\lambda \int_{\pi/2}^{\pi} \frac{d\sigma}{d\Theta} \frac{d\Theta}{1-\sec\Theta}=\lambda N_A \sigma_{\text{eff}}\,. \tag{4}$$

Eq. (4) is very similar to Eq. (1) introducing σ_{eff} instead of σ_B In Eq. (3) $f(q)$ is the scattering amplitude presented in the Tables published by Fink et al [4] but available only for $E_p \leq 1.5$ keV. Our calculations have been extended up to $E_p=3.1$ keV and to $Z=2-50$. Let us consider an alloy or compound constituted of n components and containing $N_1, N_2, \ldots N_n$ atoms per cm^3 with N_i/N_A atomic compositions. The elastic backscattering produced by the components can be calculated in a similar way as given in Eqs (2) and (4)

$$P_e=\lambda \int_{\pi/2}^{\pi} \sum N_i \frac{d\sigma_i}{d\Theta} \frac{d\Theta}{1-\sec\Theta}, \tag{5}$$

$$P_e=\lambda N_A \frac{N_i}{N_A} \sigma_{i\,\text{eff}}\,. \tag{6}$$

2. Calculation of σ_{eff} using a simplified model

The effective backscattering cross section was calculated within the first Born approximation [5] for 3 types of atomic potentials. Further, Finks data have been inserted into Eqs (2) and (4) for $E_p=1$ and 1.5 keV, in order to compare them with our results. For the central field potential we have:

$$f(q)=-\frac{2\pi}{\hbar_0^2} \int_0^{\infty} V(r) \frac{\sin qr}{qr} r^2 dr\,, \tag{7}$$

with $q=2k \sin\frac{\Theta}{2}$. All symbols in Eq. (5) have their usual meaning. Three models of atomic potentials $V(r)$ are considered:

2.1. The Thomas–Fermi potential

$V(r)=(Ze^2/r)y(z)$ with $z=(Z^{1/3}/0.88534\, a_0)r$. The Rozental [6] approximation for the function $y(z)$ was used in the present calculation: $y(z)=\sum_{i=1}^{3} a_i \exp\left[-\frac{b_i}{x}\right]$, a_i and b_i are constants. In Table I the calculated σ_{eff} values are collected for $Z=2-50$.

Table I

Effective backscattering cross-section calculated for the Thomas—Fermi potential: σ_{eff} [10^{-18} cm^2]

Z	Kinetic energy of electrons in eV				
	1000 eV	1500 eV	2000 eV	2500 eV	3000 eV
2	0.01535	0.006947	0.003946	0.002540	0.001771
3	0.03400	0.01545	0.008798	0.005674	0.003960
4	0.05965	0.02721	0.01552	0.01002	0.007002
5	0.09214	0.04215	0.02409	0.01557	0.01089
6	0.1313	0.06022	0.03447	0.02230	0.01561
7	0.1771	0.08139	0.04664	0.03021	0.02116
8	0.2293	0.1056	0.06060	0.03928	0.02752
9	0.2880	0.1328	0.07631	0.04950	0.03471
10	0.3629	0.1630	0.09376	0.06087	0.04270
11	0.4240	0.1962	0.1129	0.07337	0.05149
12	0.5013	0.2323	0.1338	0.08700	0.06108
13	0.5847	0.2713	0.1564	0.1017	0.07147
14	0.6742	0.3131	0.1807	0.1176	0.08264
15	0.7696	0.3578	0.2067	0.1346	0.09460
16	0.8709	0.4054	0.2343	0.1526	0.1073
17	0.9781	0.4557	0.2635	0.1718	0.1209
18	1.091	0.5089	0.2945	0.1920	0.1351
19	1.210	0.5648	0.3270	0.2133	0.1502
20	1.335	0.6235	0.3612	0.2357	0.1660
21	1.465	0.6849	0.3970	0.2592	0.1826
22	1.601	0.7490	0.4344	0.2837	0.1999
23	1.742	0.8159	0.4734	0.3093	0.2180
24	1.889	0.8854	0.5140	0.3360	0.2369
25	2.042	0.9576	0.5562	0.3637	0.2565
26	2.200	1.033	0.5999	0.3924	0.2768
27	2.364	1.110	0.6453	0.4222	0.2979
28	2.533	1.190	0.6921	0.4530	0.3197
29	2.708	1.273	0.7406	0.4849	0.3423
30	2.888	1.358	0.7906	0.5178	0.3656
31	3.073	1.446	0.8422	0.5517	0.3897
32	3.264	1.537	0.8952	0.5866	0.4144
33	3.460	1.630	0.9499	0.6226	0.4399
34	3.662	1.726	1.006	0.6596	0.4661
35	3.868	1.824	1.064	0.6976	0.4931
36	4.080	1.925	1.123	0.7366	0.5208
37	4.297	2.029	1.184	0.7766	0.5492
38	4.520	2.135	1.246	0.8176	0.5783
39	4.747	2.243	1.310	0.8596	0.6081
40	4.980	2.354	1.375	0.9026	0.6386
41	5.218	2.468	1.441	0.9466	0.6699
42	5.461	2.584	1.510	0.9916	0.7019
43	5.709	2.702	1.579	1.038	0.7345
44	5.962	2.823	1.650	1.085	0.7679
45	6.220	2.947	1.723	1.133	0.8020
46	6.483	3.072	1.797	1.181	0.8368
47	6.752	3.201	1.873	1.231	0.8723
48	7.025	3.332	1.950	1.282	0.9084
49	7.303	3.465	2.028	1.334	0.9453
50	7.586	3.601	2.108	1.387	0.9829

Table II

Effective backscattering cross-section calculated for the Thomas—Fermi—Dirac potential: σ_{eff} [10^{-18} cm^2]

Z	Kinetic energy of electrons in eV				
	1000 eV	1500 eV	2000 eV	2500 eV	3000 eV
2	0.01509	0.006843	0.003891	0.002508	0.001750
3	0.03365	0.01529	0.008708	0.005616	0.003920
4	0.05919	0.02695	0.01536	0.009914	0.006926
5	0.09158	0.04177	0.02383	0.01539	0.01076
6	0.1307	0.05971	0.03411	0.02205	0.01542
7	0.1764	0.08074	0.04617	0.02987	0.02090
8	0.2287	0.1048	0.06000	0.03884	0.02720
9	0.2875	0.1319	0.07559	0.04897	0.03431
10	0.3526	0.1621	0.09292	0.06023	0.04222
11	0.4240	0.1951	0.1120	0.07263	0.05093
12	0.5016	0.2311	0.1327	0.08614	0.06043
13	0.5854	0.2700	0.1552	0.1008	0.07072
14	0.6752	0.3118	0.1793	0.1165	0.08180
15	0.7711	0.3565	0.2052	0.1333	0.09365
16	0.8729	0.4040	0.2326	0.1513	0.1063
17	0.9806	0.4543	0.2618	0.1703	0.1197
18	1.094	0.5074	0.2925	0.1904	0.1338
19	1.214	0.5632	0.3249	0.2116	0.1488
20	1.339	0.6219	0.3590	0.2338	0.1645
21	1.470	0.6833	0.3946	0.2571	0.1809
22	1.607	0.7474	0.4319	0.2815	0.1981
23	1.749	0.8143	0.4707	0.3069	0.2161
24	1.897	0.8838	0.5112	0.3334	0.2348
25	2.050	0.9561	0.5533	0.3610	0.2543
26	2.210	1.031	0.5969	0.3896	0.2745
27	2.374	1.109	0.6421	0.4192	0.2954
28	2.544	1.189	0.6889	0.4499	0.3171
29	2.720	1.272	0.7373	0.4816	0.3395
30	2.901	1.358	0.7872	0.5144	0.3627
31	3.088	1.446	0.8387	0.5482	0.3866
32	3.280	1.537	0.8917	0.5830	0.4113
33	3.477	1.630	0.9463	0.6189	0.4367
34	3.680	1.726	1.002	0.6557	0.4628
35	3.888	1.825	1.060	0.6936	0.4896
36	4.102	1.926	1.119	0.7326	0.5172
37	4.320	2.030	1.180	0.7725	0.5455
38	4.544	2.136	1.242	0.8135	0.5745
39	4.773	2.245	1.306	0.8554	0.6043
40	5.008	2.357	1.371	0.8984	0.6347
41	5.248	2.471	1.438	0.9424	0.6659
42	5.492	2.587	1.507	0.9874	0.6978
43	5.742	2.706	1.576	1.033	0.7305
44	5.998	2.828	1.648	1.080	0.7638
45	6.258	2.952	1.721	1.128	0.7978
46	6.523	3.079	1.795	1.177	0.8326
47	6.794	3.208	1.871	1.227	0.8681
48	7.069	3.340	1.948	1.278	0.9043
49	7.350	3.474	2.027	1.330	0.9412
50	7.635	3.610	2.107	1.383	0.9787

2.2 The Thomas–Fermi–Dirac potential

The scattering amplitudes $f(q)$ were calculated from the analytical formula published by Bonham and Strand [7],

$$f(q)=\frac{2Z}{a_0}\sum_{i=1}^{3}\frac{a_i}{b_i^2+q^2}, \tag{8}$$

where a_i and b_i are given by polynomials of ln Z. In Table II the calculated σ_{eff} values are presented for $Z=2-50$.

2.3. The Hartree–Fock potential

The analytical expression applying this potential for scattering amplitudes was published by Strand and Bonham [8]:

$$f(q)=\frac{2Z}{a_0}\left[\sum_i a_i[b_i^2+q^2]^{-1}+Z\sum_j c_j d_j(d_j^2+q^2)^{-2}\right], \tag{9}$$

with a_i, b_i, c_j and d_j constants listed for each element. Results of our calculations are summarized in Table III for $Z=2-18$.

Table III

Effective backscattering cross-section calculated for the Hartree—Fock potential: σ_{eff} [10^{-18} cm^2]

Z	Kinetic energy of electrons in eV				
	1000 eV	1500 eV	2000 eV	2500 eV	3000 eV
2	–	–	–	–	–
3	0.03588	0.01612	0.009104	0.005839	0.004059
4	0.06236	0.02830	0.01606	0.01032	0.007188
5	0.09485	0.04351	0.02483	0.01601	0.01117
6	0.1329	0.06153	0.03530	0.02284	0.01597
7	0.1766	0.08223	0.04738	0.03076	0.02155
8	0.2261	0.1056	0.06105	0.03973	0.02789
9	0.2821	0.1317	0.07629	0.04974	0.03498
10	0.3452	0.1608	0.09317	0.06081	0.04281
11	0.4150	0.1927	0.1116	0.07287	0.05135
12	0.4931	0.2268	0.1309	0.08539	0.06019
13	0.5814	0.2656	0.1525	0.09926	0.06989
14	0.6759	0.3083	0.1767	0.1148	0.08077
15	0.7769	0.3544	0.2028	0.1316	0.09247
16	0.8837	0.4039	0.2308	0.1496	0.1050
17	0.9912	0.4601	0.2665	0.1740	0.1227
18	1.109	0.5122	0.2962	0.1935	0.1364

2.4. Numerical solution of the Dirac equation by Fink et al [4]

The most exact model for calculating theoretical electron scattering factors has been developed by Fink et al [4] presenting a numerical solution of the Dirac equation. Their results have been tabulated for electron energies $E_p = 100$ to 1500 eV, for a number of elements between $Z = 2$ to 76. Tabulated data contain the differential cross

Table IV

Effective backscattering cross-sections based on Fink's results, solving the Dirac equation $\sigma_{eff} \times 10^{18}$ cm^2

Z	Element	Kinetic energy of electrons in eV	
		1000	1500
3	Li	0.04217	0.01811
4	Be	0.07696	0.03304
7	N	0.2468	0.1072
8	O	0.3283	0.1422
13	Al	0.9740	0.4124
15	P	1.294	0.5554
19	K	2.723	4.290
20	Ca	2.343	1.072
27	Co	4.005	1.912
29	Cu	4.575	2.219
31	Ga	5.086	2.521
33	As	5.558	2.822
42	Mo	7.236	4.074
47	Ag	8.076	4.688
50	Sn	8.607	5.027
51	Sb	8.799	5.140
73	Ta	11.70	7.766
74	W	11.69	7.848
76	Os	11.65	8.000

section $\frac{d\sigma(\Theta)}{d\Omega}$ through the scattering angle Θ, in units of 10^{-16} cm^2 per steradian. In order to compare Fink's results with our calculations, the tabulated data have been integrated using our Eqs (2) and (4), i.e. σ_{eff} was calculated on the basis of Fink's results. The results of these calculations are summarized in Table IV for a number of elements. It is interesting to compare them with Tables I—III. The σ_{eff} backscattering cross section values calculated with the Thomas–Fermi–Dirac model (supplying similar results to those in paragraphs 2.1 and 2.3) are quite close to the results deduced from Fink's data being larger by a factor of 1.1—1.75 in the $Z = 3 - 50$ atomic number range. Their ratio is almost constant between $Z = 20$—33 (Ca to As) making 1.6—1.7. This is valid for the integrated differential cross sections forming σ_{eff}. The angular distributions of $d\sigma(\Theta)/d\Omega$, however, exhibit larger deviations, especially above $Z > 30$ higher atomic numbers.

3. Conclusions

The probability P_e of electron elastic backscattering from a solid is determined by its σ_{eff} backscattering cross section and by λ IMFP. A simplified method was developed for determining λ from experimental P_e results. The method is based on theoretical values of σ_{eff} calculated by integrating the differential electron scattering cross sections and considering a single scattering process. A formula is deduced for σ_{eff} of compounds and alloys.

Tabulated data for σ_{eff} in the $Z=2-50$ atomic number and $E_p=1-3$ keV energy range have been calculated using three types of the atomic potential: the Thomas–Fermi, Thomas–Fermi–Dirac and Hartree–Fock models. No significant differences have been found with these models. Results compared with σ_{eff} calculated by integrating Fink's differential cross section data obtained by computer solution of the Dirac equation were similar within a constant factor 1.2—1.7. Our data recollected in three Tables complete Fink's results restricted to $E_p \leq 1.5$ keV and given only for some elements.

References

1. G. Gergely, Surf. Interface Anal., *3*, 201, 1981; Vacuum, *3*, 89, 1983.
2. M. P. Seah and W. A. Dench, Surf. Interface Anal., *1*, 2, 1979.
3. J. C. Ashley and C. J. Tung, Surf. Interface Anal., *4*, 52, 1982.
4. M. Fink et al, Atomic Data, *1*, 385, 1970; *4*, 129, 1972 and *14*, 39, 1974.
5. S. Ichimura and R. Shimizu, J. Appl. Phys., *51*, 2853, 1980; Surface Sci., *112*, 386, 1981.
6. S. Rozental, Z. Physik, *98*, 742, 1935.
7. R. A. Bonham and T. G. Strand, J. Chem. Phys., *39*, 2200, 1963.
8. T. G. Strand and R. A. Bonham, J. Chem. Phys., *40*, 1686, 1964.

Acta Physica Hungarica 57 (1—2), pp. 139—147 (1985)

DETERMINATION OF THE MEAN FREE PATH OF ELECTRONS IN SOLIDS FROM THE ELASTIC PEAK

II. EXPERIMENTAL RESULTS

G. GERGELY, M. MENYHÁRD, A. SULYOK

Research Institute for Technical Physics of the Hungarian Academy of Sciences
1325 Budapest, Hungary

A. JABLONSKI and P. MROZEK

Institute of Physical Chemistry of the Polish Academy of Sciences
Warszawa, Poland

(Received 22 November 1983)

The simple theoretical model described in Part I [1] is applied to evaluate experimental results determined by elastic peak electron spectroscopy and published previously. The inelastic mean free path λ determined from the elastic backscattering probability P_e of 2.2 and 3.1 keV electrons on C, Si, Ge and Mo are in reasonable agreement with results of Seah and Penn. In the evaluation σ_{eff} elastic backscattering cross sections calculated with the Thomas—Fermi—Dirac potential model have been applied. This model, however, is not suitable for W and Au high atomic number elements. A detailed discussion is given on the problems associated with the measurement of P_e taking into consideration the CMA response, deconvolution and angular distribution of electron spectra. Another method for determining λ is given by comparing the elastic peaks of two elements and their σ_{eff} backscattering cross sections. Experimental results are presented for Al, Si, Ge, GaAs, GaP, InSb, GaSb, Sb, InP, SiO_2 and Si_3N_4, atomic clean surfaces measured with a CMA at $E_p = 1$, 1.5, 2 and 3 keV at normal incidence. Good agreement was found with literature data collected by Ashley for Al—Si, Ge—GaAs. The IMPF λ of GaP and GaSb proved to be 10% larger than those of GaAs, results not available in the literature. New results as well are published for λ of Sb, InSb and InP. The good agreement with literature data justifies the application of the simplified model described in Part I.

Introduction

In Part I [1] of this work a simplified model has been described for calculating σ_{eff} electron backscattering cross sections of solid surfaces using various quantum mechanical models. In this paper the theoretical σ_{eff} values will be applied to determine the inelastic mean free path (IMFP λ) of electrons using elastic peak electron spectroscopy [2]. In the paper the same notations will be used as in [1]. As shown in [1] λ can be determined with the simple relationship

$$\lambda = \frac{P_e}{N_A \sigma_{eff}} \tag{1}$$

if experimental data are available on the elastic backscattering probability P_e. Comparing the elastic peak heights of two samples with the same spectrometer:

$$\frac{\lambda_i}{\lambda_1}=\frac{P_{ei}N_{A1}\sigma_{\mathrm{eff}1}}{P_{e1}N_{Ai}\sigma_{\mathrm{eff}i}}. \tag{2}$$

For many elements and compounds reliable experimental values of λ are available (e.g. [3] and [4]). Thus the variation of σ_{eff} with Z is needed only for determining λ_i. Such data are tabulated in [1]. In this paper problems of the determination of P_e are discussed and experimental results are presented.

1. Problems associated with the experimental determination of P_e

To determine λ reliable experimental values of P_e are needed. In [2] two simple experimental methods have been described based on the total electron spectrum $N(E)$ including low energy secondary and inelastically backscattered electrons or considering the spectrum only above $E>50$ eV. The latter is more practicable since the reflection coefficient of electrons r_e is a material parameter determined mainly by Z.

$$r_e\sim\int_{50}^{E_p}N(E)dE. \tag{3}$$

Abundant literature is available on r_e. Hereupon only the new experimental results of Hunger et al [5] are mentioned. Our experimental work has been carried out with a Riber OPC 103 CMA analyser operated in DC mode [6]. The principles described can be applied for any other type of analyser.

In [2] P_e was determined by comparing the integrated elastic peak area with the total integrated $N(E)$ electron spectrum area:

$$P_e=\frac{N(E_p)\Delta E_e}{r_e}. \tag{4}$$

ΔE_e is the FWHM of the elastic peak.

In this simple experimental procedure several problems occur and corrections are to be taken into consideration: The CMA DC signal $I_c(E)\sim\Delta E_c S_c(E)N(E)$ is determined by the resolution ΔE_c of the CMA and by $S_c(E)$ spectral response of the detector. $\Delta E_c/E=$const. (0.3% for our CMA). The electron spectrum $N(E)$ is continuous [6] below the plasmon peaks. In general, Auger peaks are negligible with respect to the background $N(E)$ except the giant carbon 270 eV peak.

Regarding $S_c(E)$ difficulties can occur for $E<50$ eV [7]. It was considered constant for $E>50$ eV. Local distortions of the CMA can affect $S_c(E)$. To insert the real distribution $N(E)$ into Eqs (3) and (4) a deconvolution of $I_c(E)$ CMA signals is needed. This is very simple for the continuous part of the spectrum considered as constant steps in the integration. The contribution of the elastic and of plasmon peaks to the integral

(3) is negligible. The distortion of these peaks cannot be avoided with the poor resolution of the CMA above $E > 1.5$ keV. Regarding the elastic peak $N(E_p)$, its physical FWHM ΔE_e is determined by the Boersch effect [8] of the electron gun making some eV. The phonon broadening (50 meV) can be neglected. ΔE_e can be estimated by the gradual decrease of E_p and reaching a nearly constant FWHM in the detected elastic peak. For our integrated electron gun it was $\simeq 1.3$ eV.

The CMA response was approximated with a simple model assuming Gaussian distributions with ΔE_c for the CMA energy window and ΔE_e for the elastic peak. Then the apparent distribution of the elastic peak is given by:

$$I_c(E-E_p) \sim N(E_p) \int_{-\infty}^{+\infty} \exp\left[-\frac{x^2}{\sigma_c^2} - \frac{(x-E)^2}{\sigma_e^2}\right] dx, \tag{5}$$

denoting now by E the energy distance from the peak maximum, and with $\sigma_c = \frac{\sqrt{\pi}}{2.36} \Delta E_c$, $\sigma_e = \frac{\sqrt{\pi}}{2.36} \Delta E_e$. Integrating Eq. (5) gives around the peak maximum:

$$I_c(E-E_p) = 1.06 \frac{\Delta E_e \Delta E_c}{\sqrt{\Delta E_e^2 + \Delta E_c^2}} \exp\left[-\frac{(E-E_p)^2}{0.36(\Delta E_e^2 + \Delta E_c^2)}\right] \times N(E_p). \tag{6}$$

The same model applies for a constant background $N_b \approx (N(E))$

$$I_c(E) = 1.06 \Delta E_c N_b = 0.0032\, E N(E). \tag{7}$$

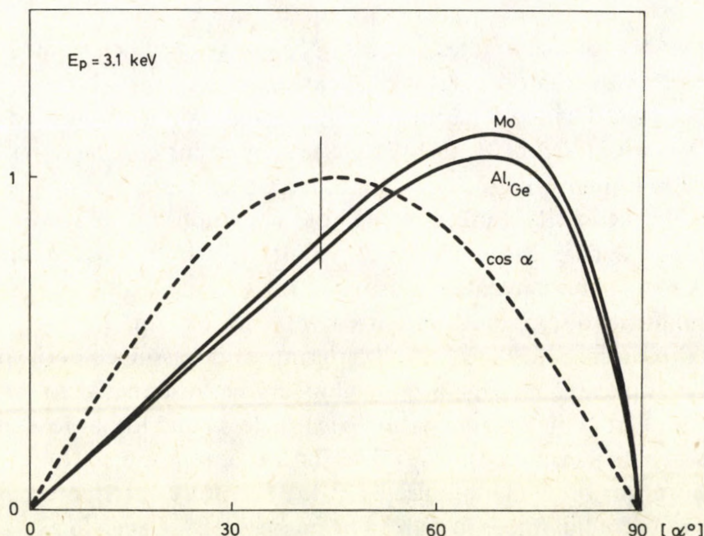


Fig. 1. Normalized angular distribution of elastically backscattered electrons calculated for the Thomas—Fermi—Dirac potential model

For deconvoluting experimental results [6] Eqs (5) and (7) are used. In the deconvoluted spectrum the FWHM of the elastic peak is the Boersch width.

The deconvolution of the plasmon loss peaks is more difficult. A procedure described by Blau [9] can be used, it must be emphasized however, that the FWHM of the plasmon peaks is a material parameter. It is more practicable to work with a spectrometer of high resolution [10]. Another problem in the determination of P_e occurs with the angular distribution of inelastically and elastically backscattered electrons. In our experiments normal incidence of the primary electrons was used and the analysis described in [1] is valid for this case. Inelastic backscattering of electrons is producing approximately Lambertian distribution [6, 11]. The same approach has been used for elastic backscattering on solid surfaces in the absence of accurate experimental data. The angular distribution of the elastic peak using Eq. (2) in [1] was calculated with the Thomas—Fermi—Dirac potential model. Some characteristic results are presented in Fig. 1. As shown in the Figure, for $\alpha = 42°$ CMA angle the deviations from the Lambertian distribution are slight but they should be taken into correction.

Our analysis refers to flat surfaces. Roughness of the sample surface (produced e.g. by ion bombardment) might be a source of error. Finally, uncertainty in the sample position (distance and angle) can be a source of error. The beam current and sensitivity level of the detector have been carefully kept constant and checked before and after each measurement. The reproducibility proved to be better than 10%.

2. Determination of IMFP λ from P_e experimental results

The IMPF λ of the samples studied in [2] have been determined by inserting experimental P_e values into Eq. (1). Results are summarized in Table I based on σ_{eff} calculated with the Thomas—Fermi—Dirac potential model. In the Table λ data of Seah and Dench [3] and of Penn [12] are presented for comparison. As shown in Tables I—III in [1] no significant differences have been noticed using the three potential models. Unfortunately, Fink et al [13] did not publish data above $E_p > 1.5$ keV. Comparing our λ data based on EPES results [2] with those of Seah [3] and Penn [12] a reasonable agreement was found for C, Fe, Ge and Mo as indicated in Table I. Significant discrepancies occurred with W and Au. Two reasons seem to account for this discrepancy: the elastic scattering model assumed in the present work might not be applicable to large atomic numbers because that the single scattering theory and the first Born approximation become less valid for large Z elements [14, 15]. The discrepancies between the λ values for low and medium atomic numbers are within the accuracy of the data collected in the literature and quoted in [3, 4]. The average deviation of the values in Table I from the Seah—Dench formula makes 19% (except Au and W).

Table I

Comparison of calculated and experimental values of λ. σ_{eff} was calculated for the Thomas–Fermi–Dirac potential.

Element	The inelastic mean free path λ[nm]		
	Based on [2]	Seah–Dench [3]	Penn [12]
	Primary electron energy $E_p = 2.2$ keV		
C	1.91	1.81	2.52
Si	4.96	2.72	3.7
Fe	1.61	2.09	2.21
Ge	3.13	2.89	2.85
Mo	2.23	2.40	2.53
W	0.95	2.42	2.63
	Primary electron energy $E_p = 3.1$ keV		
C	1.78	2.15	
Si	3.13	3.20	
Fe	1.64	2.48	
Ge	3.95	3.43	
Mo	2.65	2.85	
W	1.35	2.87	
Au	1.27	2.97	

3. Determination of IMFP λ by comparing the elastic peaks of two samples

Reliable experimental and theoretical λ data for a number of elements and compounds are presented by Ashley [4]. Those and others quoted can be used as standards for the determination of λ on other materials. The elastic peak spectra of Al, Si, SiO_2, Si_3N_4, Ge, GaAs, GaP and GaSb have been measured with our CMA for $E_p = 1$, 1.5, 2 and 3 keV in the DC mode [2] under normal incidence using the integrated CMA electron gun. The electron beam current (1 μA) was measured with a Faraday cup. The sensitivity level of the detector system was checked after each measurement recording the elastic peak of a graphite standard sample mounted on the same caroussel. The samples have been cleaned by Ar^+ ion bombardment and atomic clean surfaces have been tested by their Auger spectra (determined in the derivative mode). For evaluating our experimental results Eq. (2) was used with σ_{eff} data presented in Table II in [1] (Thomas—Fermi—Dirac model). In some cases (instead of σ_{eff} the differential cross sections $d\sigma(\Theta)/d\Theta = 138°$) were taken from Fink [13] according to Eq. (2) in [1]. For the experimental P_e values in Eq. (2) it is sufficient to compare the recorded elastic peak heights since their FWHM is constant (for a given E_p) and determined by the Boersch effect [8]. Some characteristic experimental results are presented in Fig. 2 (Si, SiO_2, Si_3N_4) for $E_p = 1$ keV and Fig. 3 (GaAs, GaP, GaSb) for $E_p = 1.5$ keV. The plasmon energy values are in good agreement with [16]. Results are summarized in Table II presenting the ratios λ_1/λ_2 for the samples compared. The

Table II

Determination of the IMFP λ by comparing two samples, using σ_{eff} values based on the Thomas–Fermi–Dirac (TFD) and Fink models, resp. For comparison literature data are presented.

Samples	E_p keV	P_{e1}/P_{e2}	N_{A2}/N_{A1}	λ_1/λ_2 TFD	λ_1/λ_2 Fink	λ_1/λ_2 Lit.	Ref.
Al/Si	1	1	0.83	0.957	0.966	1	4
	1.5	0.97		0.928	0.945		
	2	0.99		0.946			
	3	0.83		0.83			
GaAs/Ge	1	1	1	1			
	1.5	1		1	1	0.98	4
	2	0.95		0.95	1		
	3	1.01		1.02	1	0.98	
GaAs/Si	1	4.23		0.96	0.99	0.76	
	1.5	4.93		1.09	0.97	0.75	4
	2	4.5	1.1	0.99		0.75	
	3	4.23		0.92		0.75	
GaP/GaAs	1	0.71	0.902	1.09	1.09		
	1.5	0.72		1.11	1.17		
	2	0.75		1.13			
	3	0.86		1.23			
GaSb/GaAs	1	0.95	1.258	0.7	(1.19)		
	1.5	1.07		0.80	(1.1)		
	2	1.23		0.91	1.08		
	3	1.33		9.92	1.3		
Si_3N_4/Si	1	1.03	0.56	1.01	1.13		
	1.5	1.02		1	1.04		
	2	0.78		0.76			
	3	0.69		0.68			
SiO_2/Si	1	0.89		1	1.1	1.32	
	1.5	1		1.2	1.18	1.28	
	2	0.99		1.08		1.26	4
	3	0.8	0.63	0.91		1.25	
InP/GaAs	1	0.57	1.11	0.51	0.70	–	–
	1.5	0.77		0.69	1.05	–	–
	2	1.04		0.93	–		
	3	1.44		1.27	–		
InSb/Ge	1.5	0.97	1.49	0.62	0.81	–	–
	2	1.12		0.71	–		
	3	1.27		0.8	–		
Al_2O_3/Al	1.5	0.98	0.505	0.74	0.81	0.845	4
	2	0.93		0.75		0.83	
	3	0.93		0.70		0.84	
Al/Ge	1.5	0.22	0.728	0.91	1.09	1.28	4
	2	0.19		0.80	–	1.29	
	3	0.15		0.63		1.28	
Sb/Ge	1	1		0.56	0.82	–	
	1.5	1.04	1.343	0.57	0.72	–	
	2	1.32		0.72	–	–	
	3	1.31		0.71	–		

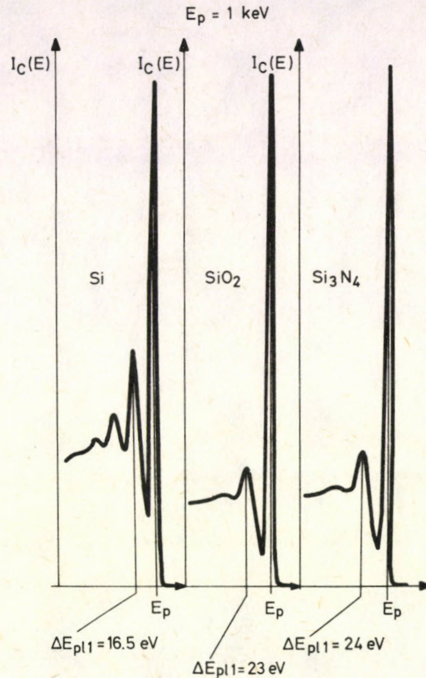


Fig. 2. Elastic peak spectra of Si, SiO_2 and Si_3N_4 surfaces for $E_p = 1$ keV. The CMA signal $I_c(E)$ is plotted in arbitrary units. Results are normalized in Table II. The first plasmon loss peaks are presented in the Figure

Table II contains the σ_{eff} ratios for the TFD and Fink models as well. As shown in the Table, no significant differences were found except the case of GaSb. The Table contains also literature data of λ_1/λ_2 indicating sources. Considering the experimental results presented in Table II and comparing them with the literature data, the following can be summarized: The $\lambda_{Al}/\lambda_{Si}$, $\lambda_{GaAs}/\lambda_{Ge}$ and $\lambda_{GaAs}/\lambda_{Si}$ are in good agreement with Ashley's results [4]. Unfortunately few experimental data are available on λ_{GaAs}, recollected in the references of [4]. A considerable scattering in the experimental data of λ_{Si} is met with in the literature quoted by Ashley [4].

No λ_{GaP} and λ_{GaSb} data are available in the literature. For the GaP—GaAs samples the ratio was approximately 1.1. Regarding GaSb—GaAs in the Table two corrections have been used: for $E_p = 3$ and 2 keV, experimental σ_B data published in [2] have been applied. For $E_p = 1$ and 1.5 keV the differential elastic scattering cross sections published by Fink have been used. In this way $\lambda_{GaSb}/\lambda_{GaAs} = 1.1$ was obtained. In a similar way new results are presented for λ of Sb, InP, InSb. They are similar to those on GaP and GaSb. The agreement of $\lambda_{SiO_2}/\lambda_{Si}$ and $\lambda_{Al_2O_3}/\lambda_{Al}$ with Ashley's results [4] is quite good.

These experimental results justify the application of our simplified method and model.

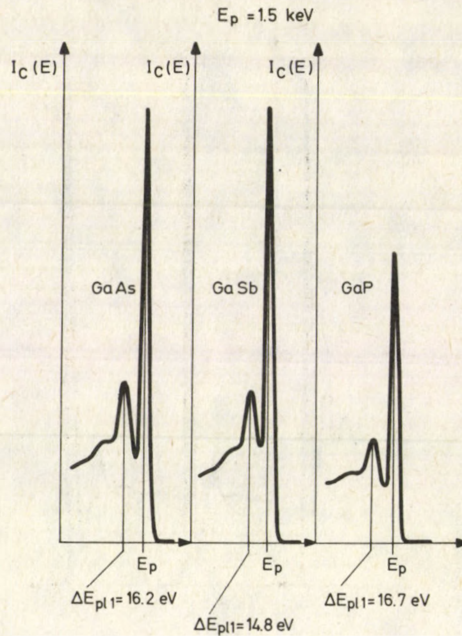


Fig. 3. Elastic peak spectra of GaAs, GaP and GaSb surfaces for $E_p = 1.5$ keV. The CMA signal $I_c(E)$ is plotted in arbitrary units. Results are normalized in Table II. The first plasmon loss peaks are presented in the Figure

Conclusions

The IMFP λ of electrons in solids can be determined by measuring the probability P_e of elastic reflection and using calculated values of the σ_{eff} effective elastic backscattering cross sections described in [1]. Evaluating the experimental results published in [2] a good agreement was found data published by Seah [3] below $Z = 42$ atomic number.

By comparing the elastic peak heights of two samples their IMFP ratio was determined with the σ_{eff} data of [1]. Using λ values of Ge and GaAs standard samples [3, 4] the IMFP λ was determined on GaP, GaSb, Sb, InSb and InP for $E_p = 1 - 3$ keV. Good agreement with literature data was found for Si, Al, Ge and GaAs published by Ashley [4], justifying the applicability of our simplified model described in [1]. The accuracy of the determination of λ can be improved by applying corrections to the angular distribution and deconvolution of elastic peak spectra.

References

1. A. Jablonski, P. Mrozek, G. Gergely, M. Menyhárd and A. Sulyok, Acta Phys. Hung., *57*, 131, 1985.
2. G. Gergely, Surf. Interface Anal., *3*, 201, 1981; Vacuum, *33*, 89, 1983.
3. M. P. Seah and W. A. Dench, Surf. Interface Anal., *1*, 2, 1979.
4. J. C. Ashley and C. J. Tung, Surf. Interface Anal., *4*, 52, 1982.
5. H.—J. Hunger and L. Küchler, Phys. Stat. Sol. (a) *56*, K45, 1979; Exp. Techn. Physik, *30*, 335, 1982.
6. G. Gergely, B. Gruzza and M. Menyhárd, Acta Phys. Hung., *48*, 337, 1980.
7. H. E. Bauer and H. Seiler, Optik, *62*, 107, 1982.
8. H. Boersch, Z. Physik, *139*, 115, 1954.
9. W. Blau and K. Kleinstück, Exp. Techn. Physik, *24*, 187 and 277, 1976.
10. G. Gergely, M. Menyhárd, C. Jardin, P. Michel, J. Electron Spectrosc. Relat. Phenom., *28*, 279, 1983.
11. L. Reimer, W. Pöpper and W. Bröcker, Scanning Electron Microscopy, *1*, 705, 1978.
12. D. R. Penn, J. Electron Spectrosc. Relat. Phenom., *9*, 29, 1976.
13. M. Fink et al, Atomic Data, *1*, 385, 1970; *4*, 52, 1972; *14*, 39, 1974.
14 A. Jablonski, Surface Sci., *74*, 621, 1978.
15. S. Ichimura and R. Shimizu, Surface Sci., *112*, 386, 1981.
16. H. Reather, Excitations of Plasmons and Interband Transitions by Electrons, Springer, Berlin, 1980.

Acta Physica Hungarica 57 (1—2), pp. 149—153 (1985)

ELECTRONIC STOPPING POWER FOR DIFFERENT ELEMENTS AND INELASTIC ENERGY LOSS IN COLLISION CASCADES

I. NAGY, J. LÁSZLÓ and J. GIBER

Physical Institute of the Technical University
1521 Budapest, Hungary

(Received 23 November 1983)

The expectation value of the electronic stopping power is obtained for different bombarding ions with different bombarding energies impinging in an Ag target. A universal form is fitted to the computer data of high reliability for this case. Energy losses are calculated in the cases of uniatomic bombardment of other elements, too. The dependence of the bombarding particle energy and atomic number on the distribution function of ε inelastic and E_L elastic energy loss are examined in simulated cascades in Al, Ag and Au targets.

Introduction

The determination of the inelastic energy loss due to atomic collisions is an important theoretical and experimental task. From the theoretical point of view it is important for the understanding of the collisions in the $v \ll e^2/\hbar$ velocity range, where the assumptions of the Born approximation are not fulfilled [1].

From the experimental viewpoint of the modern surface examinations it is necessary for the understanding of the explanation of the kinetic theory of the electron emission and secondary ionization [2, 3]. This would also serve as a useful background for the interpretation of the temperature for the thermal secondary ionization models and for the spike theory. The statistical model of Firsov [4] based on the statistical theory of the atom (Gombás [5]) is adequate for the determination of the inelastic energy loss due to binary collisions of keV energies between atoms with complicated electronic structures. One of the most important questions is the precise forming of the interaction potential. There are many known approximative forms [6]. The potential of Biersack et al [7] is used. The calculation of the expectation value of the electronic stopping power due to the binary inelastic loss (ε) is determined with the help of the probability density function of the impact parameter developed by Sanders et al [8] in amorphous targets. A more detailed analysis can be found in our earlier work [9].

Procedure

The INLOSS computer code was applied to calculate the expectation value of the electronic stopping power for binary interactions and for collisional sequences. The expectation value of the electronic stopping ($\langle S_e \rangle$) is derived applying the probability density function of the impact parameter ($f(b)db$) knowing the inelastic energy loss ($\varepsilon(E_1, Z_1, Z_2, b)$) determined in binary interactions at fixed bombarding energies (E_1):

$$\langle S_e \rangle = 2\pi \int_0^\infty \varepsilon(E_1, Z_1, Z_2, b) f(b) db, \quad (1)$$

where Z_1 and Z_2 are the atomic numbers of the bombarding and target particle, respectively, b is the impact parameter measured in units of the Thomas—Fermi screening radius of the target particle. $f(b)$ is a function of N, the atomic density of the target according to the definition of Sanders et al [8]. The electronic stopping power ($\langle dE/dx \rangle_e$) can be calculated from (1) as:

$$\langle dE/dx \rangle_e = N \langle S_e \rangle, \quad (2)$$

where N is in $\left(\frac{\text{nm}}{10}\right)^{-3}$, $\langle S_e \rangle$ in eV $\left(\frac{\text{nm}}{10}\right)^{2}$. Therefore the dimension of the stopping power is eV $\left(\frac{\text{nm}}{10}\right)^{-1}$. The calculation was carried out in a wide range of E_1, Z_1 and Z_2.

Results

A) Electronic stopping power

The calculated results of uniatomic bombardment are shown in Fig. 1 as the function of the atomic number. The bombarding energy (E_1) was changed between 0.5 and 100 keV. We have denoted characteristic results of the calculated $\langle dE/dx \rangle_e$ functions for some ion-target combinations: for $Al^+ \to Al$, $Cu^+ \to Cu$, $Ag^+ \to Ag$ and $Au^+ \to Au$. The curves show the variation in energy dependence.

Fig. 2 shows the important role of the target density. The bombarding energy is 100 keV and uniatomic collisions were calculated (Al^+, K^+, Ti^+, Cu^+, Zr^+, Ag^+, Te^+, Ba^+, Ta^+, Au^+).

The ratio of the losses is approximately determined by the ratio of the atomic densities $N = L\rho/M$, where L is the Avogadro number, ρ is the density and M is the atomic weight.

We have examined the features of the expectation value of the electronic stopping power in detail for an Ag target as the function of the bombarding energy and the atomic number of the bombarding ion. The energy was varied from 10 to 200 keV. The bombarding ions were Al^+, K^+, Ti^+, Cu^+, Se^+, Zr^+, Ag^+, Te^+, Ba^+, Ta^+ and Au^+.

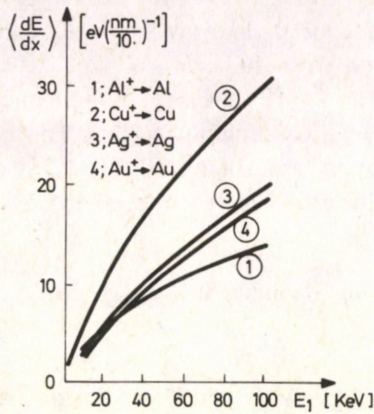


Fig. 1. The expectation value of the electronic stopping power vs the bombarding energy for 4 different cases of uniatomic bombardment

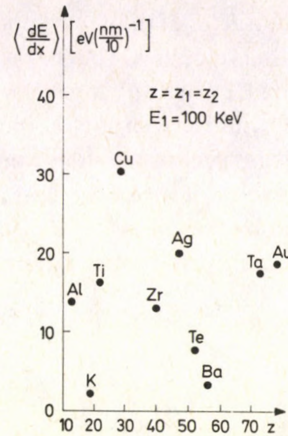


Fig. 2. The expectation value of the electronic stopping power vs the atomic number of the target for uniatomic bombardments with 100 keV bombarding energy

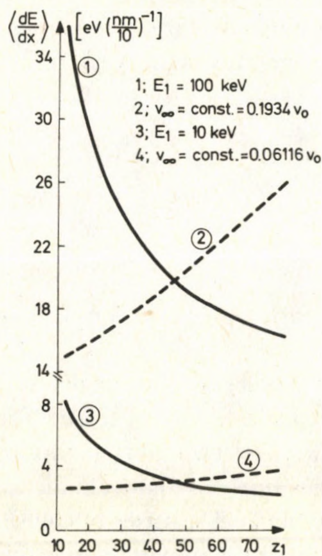


Fig. 3. The expectation value of the electronic stopping power vs the atomic number of the bombarding particle for 2 different constant energies (full lines) and for 2 different constant velocities of the bombarding particle (dashed lines)

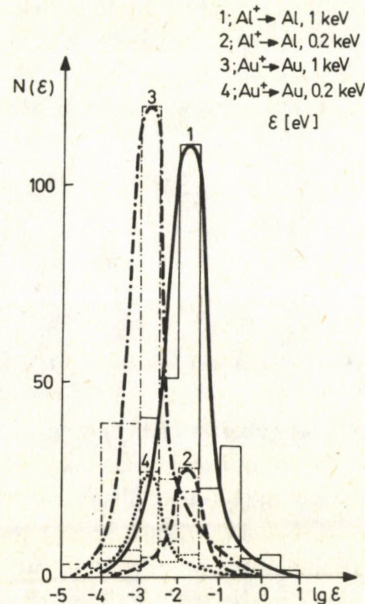


Fig. 4. The distribution function of the inelastic energy loss (N is in piece) in 4 different collision cascades

The character of the above functions was also investigated in the case of ions with constant initial velocities. The characteristic dependence on the atomic numbers can be seen in Fig. 3 drawn for $E_1=100$ keV and $E_1=10$ keV as well as for $v_\infty=0.1934\,v_0$ and $v_\infty=0.0611\,v_0$ (here $v_0=e^2/\hbar\cong 2.188\cdot 10^6$ m/s is the Bohr velocity).

A fit has succeeded to the computed electronic stopping power data. Thus, the expectation value of the energy loss ($\langle dE/dx\rangle_{ei}$) for an i type bombarding particle can be described by the following closed formula in the 10—200 keV energy region as a function of energy and type of bombarding ion:

$$\langle dE/dx\rangle_{ei}\cong 0.04875E_1^{\gamma}\times[Z_{\mathrm{Ag}}^{4/3}Z_i 2E_R\cdot 10^{-3}(m_{\mathrm{Ag}}+m_i)/(0.8853m_{\mathrm{Ag}})]^{-\gamma}\times$$
$$\times(Z_{\mathrm{Ag}}+Z_i)^2 m_i^{-1/2}, \qquad (3)$$

where $\gamma=E_1^{\alpha}$ (E_1 must be substituted in keV), $E_R=13.6$ eV, m_i and Z_i are the mass and atomic number of the bombarding particle, m_{Ag} and Z_{Ag} are those of the target. In the energy range in question $\alpha=-0.156$.

It can be seen from formula (3), that the exponent is decreasing from about 0.7 to 0.44 with increasing energy. About 10 keV γ becomes 0.5. Below 10 keV γ exceeds the value of 0.7. The deviation of the stopping power from the $\sim E_1^{1/2}$ behaviour (low energies — Lindhard [10]) is an experimentally supported fact [2, 11, 12]. For example a very good agreement was found for the electron yield in comparison to the results of experiment [2] applying formula (3) for the case of Ag bombarded by Se^+ ions [13].

The procedure should be repeated for different target materials, as the probability density function of the impact parameter contains the N_i atomic density in an exponent, namely:

$$f(b)db=\frac{32}{3}\pi N_i^2\int_0^\infty (x^2+b^2)^2\exp\left[-\frac{4}{3}\pi N_i(x^2+b^2)^{3/2}\right]dx\,db.$$

B) Distribution function of the inelastic energy loss in collision cascades

We have examined the energy deposition in a collision cascade developed in an amorphous target for the elastic as well as for the inelastic energy losses. The impact parameters of the binary collisions in the cascade were selected from a uniform random distribution between 0 and 0.3 nm. The energy deposition was slightly influenced by the choice of the lattice parameter within reasonable limits. The series of collisions were traced as far in the uniatomic target as the energy of the moving particle decreased below a certain threshold. This value was chosen to be equal to 10 eV.

The four curves of Fig. 4 are characteristic of the distribution of the inelastic energy loss (ε). In the case of different target materials the place of the maximum of the curves shifts. According to the simulated results, the maximum of the inelastic energy loss is lower if the target contains heavier atoms.

The deposition of the elastic energy loss (E_L) can be seen in Fig. 5. The drawn distributions show well the behaviour of the energy loss as a function of energy for different target materials.

Examining the mean collision frequency ($\langle n \rangle$) for two energies as a function of target material we have obtained the curves displayed in Fig. 6. It seems that in the case

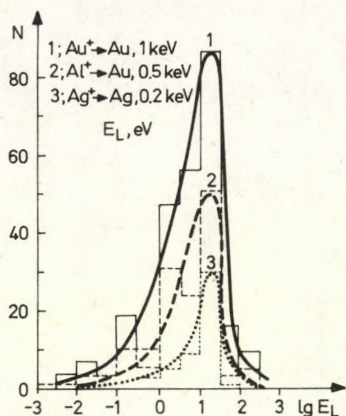


Fig. 5. The distribution function of the elastic energy loss (N is in piece) in 3 different collision cascades

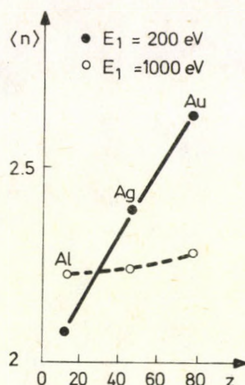


Fig. 6. The change of the mean collision number for 3 target materials and for 2 energies

of lower energy values the mean frequency increases with increasing atomic numbers. The relatively small values prove that the energy deposition is not determined primarily by the bombarding and first generation particles tracing long trajectories, but the particles pushed out with lower energies, i.e. particles from a domain of narrow geometrical limits. Therefore some thermodynamical considerations would take place basically.

References

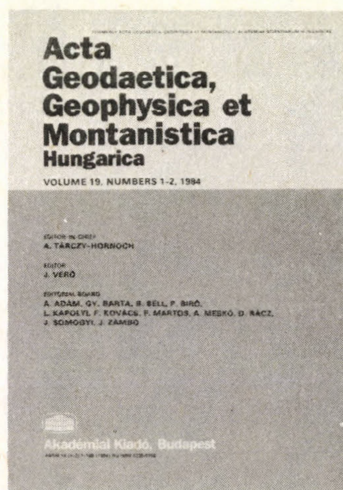
1. L. D. Landau and E. M. Lifshitz, Quantum mechanics, Pergamon, Oxford, 1958.
2. B. Svensson, G. Holmen and J. Linnors, Nucl. Instrum. Methods, *194*, 429, 1982.
3. R. Kelly, Phys. Rev. B, *25*, 700, 1982.
4. O. B. Firsov, Sov. Phys. JETP, *26*, 1076, 1959.
5. P. Gombás, Die statistische Theorie des Atoms und ihre Anwendungen, Springer, Wien, 1949.
6. I. M. Torrens, Interatomic Potentials, Academic Press, New York, 1972.
7. J. P. Biersack and J. F. Ziegler, Nucl. Instrum. Methods, *194*, 93, 1982.
8. J. B. Sanders and H. E. Roosendaal, Nucl. Instrum. Methods, *132*, 267, 1976.
9. I. Nagy, J. László and J. Giber, Acta Phys. Hung., *57*, 113, 1985.
10. J. Lindhard, M. Scharff and H. E. Schiott, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd., *33*, No 14, 1963.
11. R. Blume, W. Eckstein and H. Veerbeck, Nucl. Instrum. Methods., *194*, 67, 1982.
12. W. Heiland and E. Taglauer, The Physics of Ionized Gases, ed. by R. K. Janev, Beograd, 1978, p. 239.
13. I. Nagy, J. László and J. Giber: Appl. Phys. A., *31*, 153, 1983.

Acta Geodaetica, Geophysica et Montanistica Hungarica

(Formerly: Acta Geodaetica, Geophysica et Montanistica Academiae Scientiarum Hungaricae)

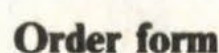
Editor in Chief:
A. Tárczy-Hornoch

Editor:
J. Verő



The journal publishes original papers in the field of geodesy, geophysics and mining engineering with special emphasis on certain topics as geodetic computation methods, instrumentation, electromagnetic induction, geomagnetic pulsations, rock mechanics, oil production, etc. It presents reviews as well.

Founded 1966
Papers in English
Publication: one volume of four issues annually
Price per volume: $44.00; DM 99,—
Size: 17 × 25 cm
ISSN 0374–1842



Order form
to be returned to
KULTURA
Hungarian Foreign Trading Company
P.O. Box 149, H-1389 Budapest, Hungary

☐ Please enter my/our subscription for ACTA GEODAETICA, GEOPHYSICA ET MONTANISTICA HUNGARICA for one year

☐ Please enter my/our standing order for ACTA GEODAETICA, GEOPHYSICA ET MONTANISTICA HUNGARICA starting with

Name: ______________________

Address: ______________________

Date and signature: ______________________

Contents of Volume 18. Numbers 1–2

Professor Lajos Homoródi 1911–1982
Izotov A. A.: Some problems of geodynamics and methods for their solution
Rinner K.: Entwicklungstendenzen in der Landesvermessung
Abdelhamid K. H.: On the adjustment of gyrotheodolite observations
Abdelhamid K. H.: Evaluation of the drift in the gyrotheodolite observations and its influence on gyro-azimuths of the amplitude method
Abdelhamid K. H., Szádeczky-Kardoss Gy.: Simple determination of the convergence of the meridians for the area of Egypt
Mentes Gy.: Problems of field data collection in geodetic measurements
Hazay I.: Quick zone-to-zone transformation in the Gauss–Krüger projection
Procházková D.: Source parameters of several earthquakes recorded at the seismic station Kašperské Hory
Le Minh Triet: On the hydromagnetic radial oscillations in the Earth's liquid core and the geomagnetic secular short-period variations
Bencze P.: Turbulence in the lower thermosphere during geomagnetic storms deduced from ionospheric sporadic *E* parameters
Zsiros T.: Faults and earthquakes in Hungarian territory east of the river Tisza
Zsiros T.: The Érmellék earthquake of 1834
Csernyák L., Steiner F.: Limit distribution of the most frequent values of samples from symmetrical distributions
Csernyák L., Steiner F.: Relation between two estimates of location: the most frequent value and the arithmetic mean
Märcz F.: Geomagnetic after-effect in atmospheric radio noise and its relation to Pc 1-type pulsations
Zhdanov M. S., Varentsov Iv. M., Bilinsky A.: Formalized 2D interpretation of the induction anomaly in the Soviet Carpathians
Kis K.: Derivation of coefficients reducing magnetic anomalies to the magnetic pole and to the magnetic equator
Meskó A.: Regional Bouguer gravity maps of Hungary
Agarwal A. K., Rastogi R. G., Singh B. P.: Equatorial electric fields generated by the quiet-day dynamo and by solar disturbances

BOOK REVIEWS

Akadémiai Kiadó

Publishing House
of the Hungarian Academy of Sciences
Budapest

Invitation for papers
Manuscripts should be sent to the editors
AGGM Editorial Office
P.O. Box 5
H-9401 Sopron
Hungary

NOTES TO CONTRIBUTORS

I. PAPERS will be considered for publication in Acta Physica Hungarica only if they have not previously been published or submitted for publication elsewhere. They may be written in English, French, German or Russian.

Papers should be submitted to

Prof. I. Kovács, Editor
Department of Atomic Physics, Technical University
1521 Budapest, Budafoki út 8, Hungary

Papers may be either articles with abstracts or short communications. Both should be as concise as possible, articles in general not exceeding 25 typed pages, short communications 8 typed pages.

II. MANUSCRIPTS

1. Papers should be submitted in three copies.
2. The text of papers must be of high stylistic standard, requiring minor corrections only.
3. Manuscripts should be typed in double spacing on good quality paper, with generous margins.
4. The name of the author(s) and of the institutes where the work was carried out should appear on the first page of the manuscript.
5. Particular care should be taken with mathematical expressions. The following should be clearly distinguished, e.g. by underlining in different colours: special founts (italics, script, bold type, Greek, Gothic, etc.); capital and small letters; subscripts and superscripts, e.g. x^2, x_3; small *l* and I; zero and capital O; in expressions written by hand: *e* and *l*, *n* and *u*, *v* and v, etc.

A List of Symbols on a separate sheet should be attached to each paper.

6. References should be numbered serially and listed at the end of the paper in the following form: J. Ise and W. D. Fretter, Phys. Rev., *76*, 933, 1949.

For books, please give the initials and family name of the author(s), title, name of publisher, place and year of publication, e.g.: J. C. Slater, Quantum Theory of Atomic Structures, I. McGraw-Hill Book Company, Inc., New York, 1960.

References should be given in the text in the following forms: Heisenberg [5] or [5].

7. Captions to illustrations should be listed on a separate sheet, not inserted in the text.
8. In papers submitted to Acta Physica all measures should be expressed in SI units.

III. ILLUSTRATIONS AND TABLES

1. Each paper should be accompanied by three sets of illustrations, one of which must be ready for the blockmaker. The other sets attached to the copies of the manuscript may be rough drawings in pencil or photocopies.
2. Illustrations must not be inserted in the text.
3. All illustrations should be identified in blue pencil by the author's name, abbreviated title of the paper and figure number.
4. Tables should be typed on separate pages and have captions describing their content. Clear wording of column heads is advisable. Tables should be numbered in Roman numerals (I, II, III, etc.).

IV. RETURN OF MATERIAL

Owing to high postage costs, the Editorial Office cannot undertake to return *all* material not accepted for any reason for publication. Of papers to be revised (for not being in conformity with the above Notes or other reasons) only *one* copy will be returned. Material rejected for lack of space or on account of the Referees' opinion will not be returned to authors outside Europe.

Periodicals of the Hungarian Academy of Sciences are obtainable at the following addresses:

AUSTRALIA
C.B.D. LIBRARY AND SUBSCRIPTION SERVICE
Box 4886, G.P.O., *Sydney N.S.W. 2001*
COSMOS BOOKSHOP, 145 Ackland Street
St. Kilda (Melbourne), Victoria 3182

AUSTRIA
GLOBUS, Höchstädtplatz 3, *1206 Wien XX*

BELGIUM
OFFICE INTERNATIONAL DE LIBRAIRIE
30 Avenue Marnix, *1050 Bruxelles*
LIBRAIRIE DU MONDE ENTIER
162 rue du Midi, *1000 Bruxelles*

BULGARIA
HEMUS, Bulvar Ruszki 6, *Sofia*

CANADA
PANNONIA BOOKS, P.O. Box 1017
Postal Station "**B**", *Toronto, Ontario M5T 2T8*

CHINA
CNPICOR, Periodical Department, P.O. Box 50
Peking

CZECHOSLOVAKIA
MAĎARSKÁ KULTURA, Národní třida 22
115 66 Praha
PNS DOVOZ TISKU, Vinohradská 46, *Praha 2*
PNS DOVOZ TLAČE, *Bratislava 2*

DENMARK
EJNAR MUNKSGAARD, Norregade 6
1165 Copenhagen K

FEDERAL REPUBLIC OF GERMANY
KUNST UND WISSEN ERICH BIEBER
Postfach 46, *7000 Stuttgart 1*

FINLAND
AKATEEMINEN KIRJAKAUPPA, P.O. Box 128 *SF- 00101 Helsinki 10*

FRANCE
DAWSON-FRANCE S. A., B. P. 40, *91121 Palaiseau*
EUROPÉRIODIQUES S. A., 31 Avenue de Versailles, *78170 La Celle St. Cloud*
OFFICE INTERNATIONAL DE DOCUMENTATION ET LIBRAIRIE, 48 rue Gay-Lussac
75240 Paris Cedex 05

GERMAN DEMOCRATIC REPUBLIC
HAUS DER UNGARISCHEN KULTUR
Karl Liebknecht-Straße 9, *DDR-102 Berlin*
DEUTSCHE POST ZEITUNGSVERTRIEBSAMT Straße der Pariser Kommüne 3–4, *DDR-104 Berlin*

GREAT BRITAIN
BLACKWELL'S PERIODICALS DIVISION
Hythe Bridge Street, *Oxford OXI 2ET*
BUMPUS, HALDANE AND MAXWELL LTD.
Cowper Works, *Olney, Bucks MK46 4BN*
COLLET'S HOLDINGS LTD., Denington Estate *Wellingborough, Northants NN8 2QT*
WM. DAWSON AND SONS LTD., Cannon House *Folkstone, Kent CT19 5EE*
H. K. LEWIS AND CO., 136 Gower Street
London WCIE 6BS

GREECE
KOSTARAKIS BROTHERS INTERNATIONAL BOOKSELLERS, 2 Hippokratous Street, *Athens-143*

HOLLAND
MEULENHOFF-BRUNA B. V., Beulingstraat 2, *Amsterdam*
MARTINUS NIJHOFF B.V.
Lange Voorhout 9–11, *Den Haag*
SWETS SUBSCRIPTION SERVICE
347b Heereweg, *Lisse*

INDIA
ALLIED PUBLISHING PRIVATE LTD., 13/14
Asaf Ali Road, *New Delhi 110001*
150 B-6 Mount Road, *Madras 600002*
INTERNATIONAL BOOK HOUSE PVT. LTD.
Madame Cama Road, *Bombay 400039*
THE STATE TRADING CORPORATION OF INDIA LTD., Books Import Division, Chandralok 36 Janpath, *New Delhi 110001*

ITALY
INTERSCIENTIA, Via Mazzé 28, *10149 Torino*
LIBRERIA COMMISSIONARIA SANSONI, Via Lamarmora 45, *50121 Firenze*
SANTO VANASIA, Via M. Macchi 58
20124 Milano
D. E. A., Via Lima 28, *00198 Roma*

JAPAN
KINOKUNIYA BOOK-STORE CO. LTD.
17-7 Shinjuku 3 chome, Shinjuku-ku, *Tokyo 160-91*
MARUZEN COMPANY LTD., Book Department, P.O. Box 5050 Tokyo International, *Tokyo 100-31*
NAUKA LTD. IMPORT DEPARTMENT
2-30-19 Minami Ikebukuro, Toshima-ku, *Tokyo 171*

KOREA
CHULPANMUL, *Phenjan*

NORWAY
TANUM-TIDSKRIFT-SENTRALEN A.S., Karl Johansgatan 41–43, *1000 Oslo*

POLAND
WĘGIERSKI INSTYTUT KULTURY, Marszalkowska 80, *00-517 Warszawa*
CKP-I W, ul. Towarowa 28, *00-958 Warszawa*

ROUMANIA
D. E. P., *Bucuresti*
ILEXIM, Calea Grivitei 64–66, *Bucuresti*

SOVIET UNION
SOJUZPECHAT — IMPORT, *Moscow*
and the post offices in each town
MEZHDUNARODNAYA KNIGA, *Moscow G-200*

SPAIN
DIAZ DE SANTOS, Lagasca 95, *Madrid 6*

SWEDEN
GUMPERTS UNIVERSITETSBOKHANDEL AB
Box 346, *401 25 Göteborg 1*

SWITZERLAND
KARGER LIBRI AG, Petersgraben 31, *4011 Basel*

USA
EBSCO SUBSCRIPTION SERVICES
P.O. Box 1943, *Birmingham, Alabama 35201*
F. W. FAXON COMPANY, INC.
15 Southwest Park, *Westwood Mass. 02090*
READ-MORE PUBLICATIONS, INC.
140 Cedar Street, *New York, N. Y. 10006*

YUGOSLAVIA
JUGOSLOVENSKA KNJIGA, Terazije 27, *Beograd*
FORUM, Vojvode Mišića 1, *21000 Novi Sad*

Index: 26.022

Acta Physica Hungarica

VOLUME 57, NUMBERS 3–4, 1985

EDITOR–IN–CHIEF

I. KOVÁCS

EDITORIAL BOARD

Z. BAY, R. GÁSPÁR, N. KÜRTI, K. NAGY,
L. PÁL, A. SZALAY, I. TARJÁN, B. TELEGDI,
L. TISZA, E. WIGNER

Akadémiai Kiadó, Budapest

ACTA PHYS HUNG. APAHAQ 57 (3–4) 155–358 (1985) HU ISSN 0231-4428

ACTA PHYSICA HUNGARICA

A JOURNAL OF THE HUNGARIAN ACADEMY OF SCIENCES

EDITED BY
I. KOVÁCS

Acta Physica publishes original papers on subjects in physics. Papers are accepted in English, French, German and Russian.

Acta Physica is published in two yearly volumes (4 issues each) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1054 Budapest, Alkotmány u. 21

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
1389 Budapest P.O. Box 149

or to its representatives abroad.

Acta Physica is indexed in Current Contents, in Physics Abstracts and in Current Papers in Physics.

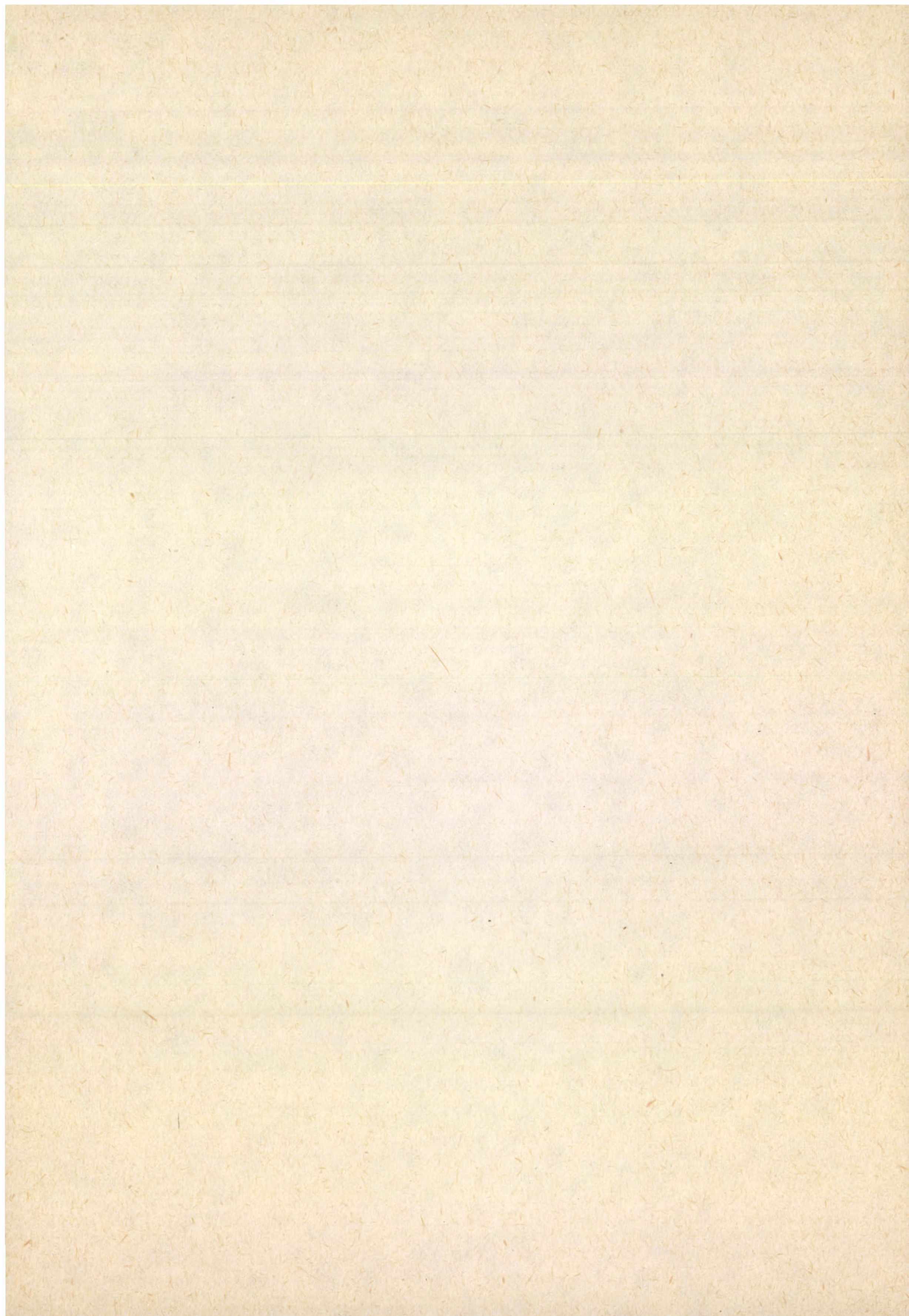
PROCEEDINGS
OF THE
THIRD HUNGARIAN CONFERENCE
ON CRYSTAL GROWTH
HCCG-3
WITH INTERNATIONAL PARTICIPATION

BUDAPEST, HUNGARY
19—21 SEPTEMBER, 1983

The Conference was organized by the Section of Crystal Physics
of the Roland Eötvös Physical Society, Budapest

Organizing Committee:

Prof. Dr. I. Tarján (Honorary Chairman)
Dr. E. Lendvay (Chairman)
Dr. E. Hartmann (Chairman of the Programme Committee)
Mr. T. Görög, Dr. E. Hidvégi, Mrs. S. Lándori-Keszthelyi,
Mr. J. Paitz, Mrs. M. Pardavi-Horváth, Dr. E. Sterk,
Mr. L. Vannay, Dr. R. Voszka



CONTENTS

Preface. *I. Tarján* 159

CONDENSED MATTER

The history of crystal growth. *J. Bohm* 161

Single crystal research — crystal quality control — industrial production. *R. Voszka* 179

Phantom controlled automatic Czochralski puller sensing the crystal weight. *B. Mező, L. Bálint* and *R. Voszka* 185

Finite element analysis of process control and operating limits in edge-defined filmfed growth (EFG) of silicon and sapphire ribbons: a review. *J. P. Kalejs, H. M. Ettouney* and *R. A. Brown* . 189

Preparation of aluminium–sapphire bicrystals. *H. G. Lindenkreuz, M. Jurisch* and *S. Sahling* 205

Preparation of ZnO single crystals by chemical transport reaction and determination of ZnO homogeneity range. *H. Oppermann, G. Stöver, A. Heinrich, K. Teske* and *E. Ziegler* 213

Preparation of GaAs structures for different devices by chloride transport. *T. Görög, I. Gyuró* and *K. Somogyi* 223

Growth of PbSe crystals from lead melt. *V. Rakovits* and *M. Hársy* 233

Selective epitaxial growth of $Al_xGa_yIn_{1-x-y}Sb$. *E. Lendvay, L. Petrás* and *V. A. Gevorkian* 237

Heat conduction vs crystal defects in GaSb ingots grown by the vertical Bridgman method. *M. Hársy, F. Koltai* and *T. Görög* 245

X-ray and SEM investigation of crystal defects in N-type, P-type and semi-insulating GaAs. *P. Franzosi* and *G. Salviati* 251

LEC-growth and the major electrical and structural characteristics of semi-insulating gallium arsenide. *R. Fornari, C. Paorici, L. Zanotti* and *G. Zuccalli* 263

Electrochemical etching. *Chr. Nanev* and *K. Dicheva* 271

Investigation of morphology and surface structure of small NaCl particles. *M. Rösler* 285

The growth of TeO_2 crystals from the boiling solution. *L. Ötsi, E. Hartmann* and *B. Vajna* 295

Growth and luminescence properties of GaSb single crystals. *W. Lewandowski, M. Gajewska, B. Pastuszka* and *M. Bugajski* 303

Evolution of impurity centres in NaCl:Ca single crystals under laser radiation. *M. V. Galustashvili* and *D. G. Driyaev* 309

KDP Q-switches and second harmonic generator for high-power solid-state lasers. *J. Bakos, T. Juhász, Cs. Kuti* and *L. Vannay* 313

On the variations in surface diffusion coefficient of LFM crystal. *M. Rak* 317

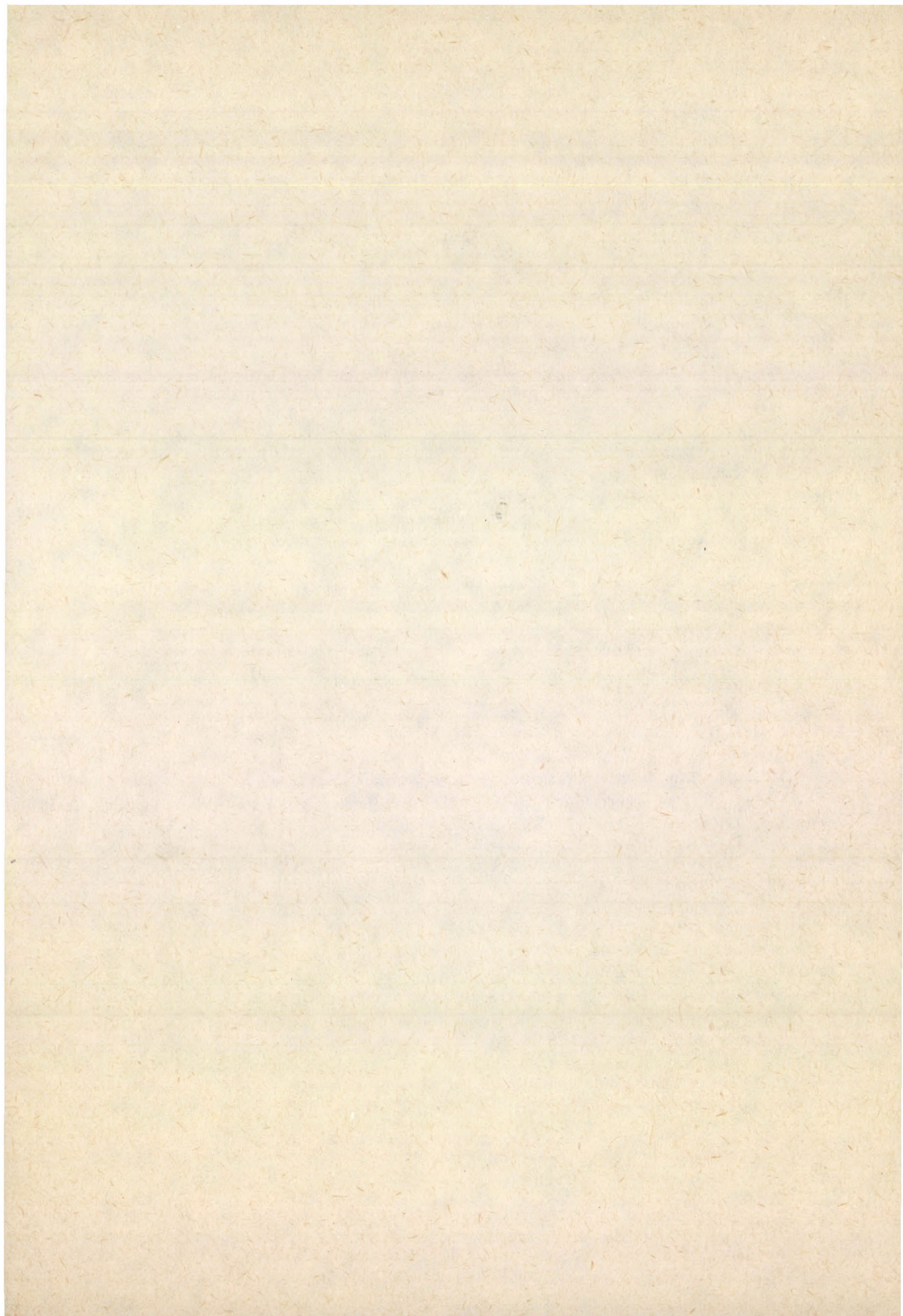
Preliminary studies on crystal growth of $Pb_{1-x}Cd_xS$ solid solutions. *P. Ştefiu* 323

Bunching of monomolecular steps during growth and evaporation of crystals. *L. Malicskó* 335

Macro- and microanalytical investigations of impurities during production of TeO_2 single crystals. *L. Malicskó, I. Cravero* and *L. Ötsi* 347

TITLES OF OTHER PAPERS PRESENTED AT HCCG-3 355

The final manuscript was received by the Editorial Office on 15th December 1983.



PREFACE

The *3rd Hungarian Conference on Crystal Growth* was held in Budapest in September 1983 with international participation.

One of the interesting presentations informed us that the history of crystal growth dates back about five hundred years. Glancing back over the past period one can state that this field represents a classical example of how human curiosity and playfulness become science and how science becomes technology.

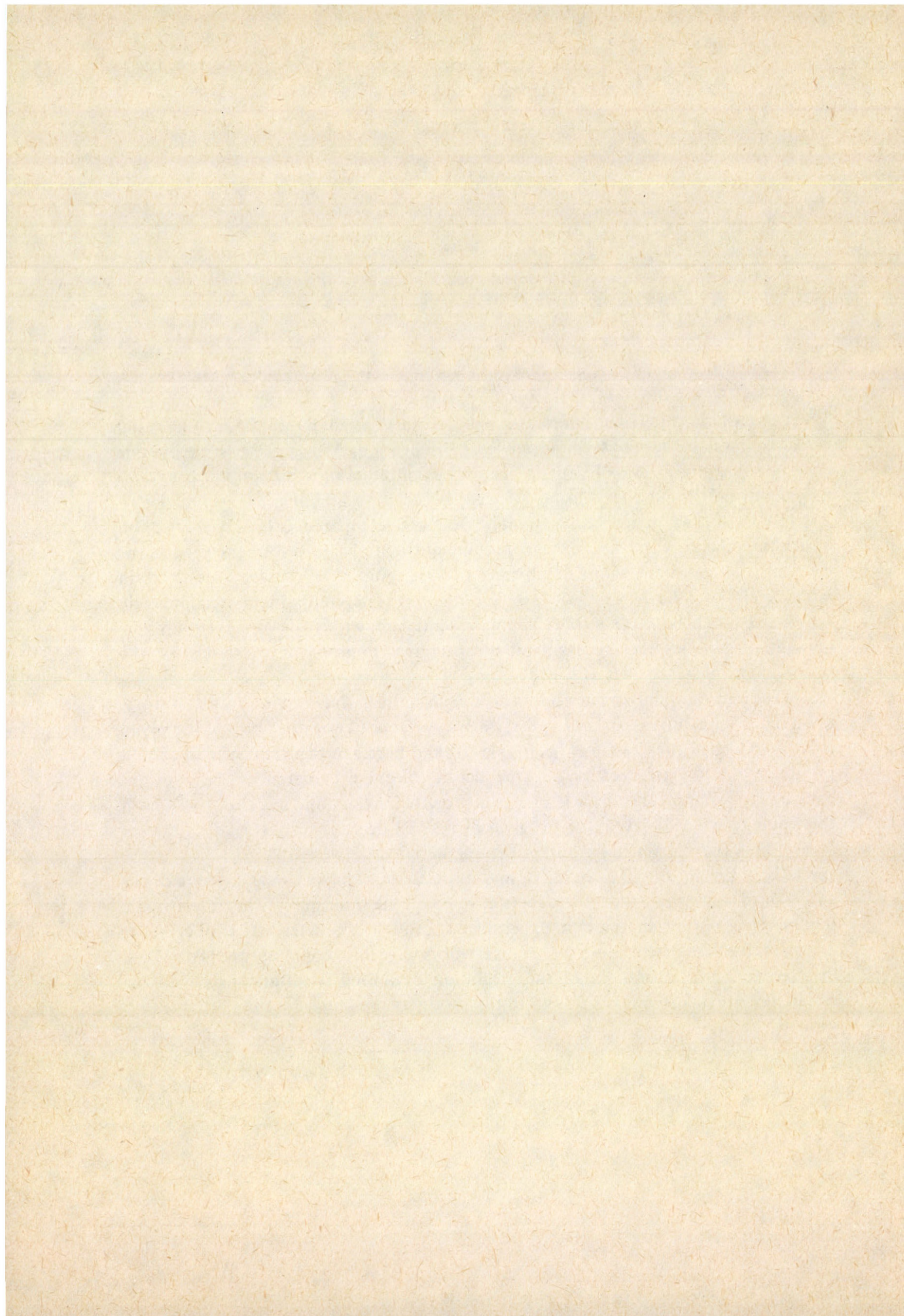
J. J. Gilman in his book twenty years ago compared crystal growth with agriculture: the growth takes place from some nutrient phase, to start the process the nutrient is often seeded with a small piece of the crystal to be grown, after a certain amount of time the harvest is ready for reaping. The comparison is completely reasonable since both operations supply essential materials for our everyday life: agriculture supplies the food, crystal growth gives appropriate materials for the scientific-technical facilities of life.

The international conferences on crystal growth also have an interesting history: not quite so long, but even so nearly forty years. As far as I know the first was in 1949 in Bristol and since then one has been able to take part almost every year in such a programme. The last decades, i.e. the period of these conferences, have meant enormous development in the field of crystal growth. On this point it is also true — as in the whole of science — that the progress of the last decades surpasses all of that of several hundreds of years. At our present conference we were able to welcome the specialists of more than 10 countries and more than 60 papers were presented. No special field was emphasized, the papers covered all the three main fields intertwined: theory of nucleation and growth, experimental growth, characterization of crystals.

It is particularly pleasing that several works gave an account of results obtained in international cooperation. At the exhibition organized simultaneously with the conference great interest was shown in the products presented by the participants.

Tarján Imre

Prof. Dr. I. Tarján



Acta Physica Hungarica, 57 (3—4) pp. 161—178 (1985)

THE HISTORY OF CRYSTAL GROWTH

J. Bohm

Central Institute of Optics and Spectroscopy, Academy of Sciences of the GDR
Berlin—Adlershof, GDR

A survey is given of the historical development of both theoretical and experimental knowledge from the early beginning to the fifties of our century. The survey is completed by a full bibliography of the most relevant papers, a timetable, and a list of conferences devoted to crystal growth.

The art of crystallization extends far back in the past and antedates considerably the written history of man. The crystallization of salt from sea water by evaporation was already practised at many places in prehistoric time and can be considered one of the oldest technical methods of transforming materials — perhaps as well as the burning of earthenware. Crystallization procedures were recorded in written documents well before the Christian era. The Roman Plinius in his "Naturalis historia" mentioned the crystallization of a number of salts, for instance of vitriols. The mediaeval alchemists, European as well as Arabian, had arrived at a stage of detailed knowledge about many crystallization processes and phenomena. The alchemist Geber, whose papers are dated in the 12th or 13th century (cf. Darmstaedter 1922), described the preparation and purification of various materials by recrystallization as well as by sublimation and distillation.

Towards the end of the Middle Ages, the general technical progress led to corresponding progress in the techniques of material production and transformation, too. In the middle of the 16th century, Birringuccio (1540) recorded in detail the leaching of saltpetre and its purification by recrystallization, and the Saxonian scientist Agricola (1556) in his famous, extensive work "De re metallica" gave instructions how to produce various salts, alums and vitriols. (Fig. 1)

In the following century, the word "crystal" came in use more and more in the modern, general sense. Originally, Homer had used the expression "crystallos" for ice crystals only, antiquity had extended it to quartz crystals (rock crystals). Also in the 17th century the denotation "crystallization" came in use, replacing earlier expressions like "condensation" or "coagulation".

On New Year's night of 1611, a snowflake landing on Johannes Kepler's sleeve was the point of departure for his charming essay: "A New Years's Gift, or On a Hexagonal Snowflake". Kepler (1611) concluded that snow crystals are built by closed packed spherical particles and posed, in such a way, the correct principle of crystallographic form and structure. Fifty years later Hooke (1665) claimed — in his

Die Pfanne A. Der Trog B. Die roſtartigen Stäbe C. Die Schnüre D. Kleine Steine E.

Fig. 1. Crystallization of vitriol. Woodcut from Agricola's "*De re metallica*" Basel: Froben 1556 (Strings are used for seeding)

"Micrographia" based on microscopic observations of many crystals — that every crystal form can be realized by arrangements of spherical particles. Looking at Kepler's drawings and remembering the work of Agricola and his precursors formerly cited, one can say that the very basic knowledge on both crystal growing and structure had already become available at this early time.

Indeed, a historical review reveals that the "modern" scientific development of crystallography started in the 17th century. In about 1600 Caesalpinus (1602) had already observed that crystals of specific materials, like sugar, saltpetre, alum, vitriols and so on, grown from solutions, exhibit typical forms, characteristic of each material. But it was not until as late as 1669 that Nicolaus Steno discovered the law of constancy

of crystal angles — the fundamental law underlying the growth of crystals. Steno's work was extended and generalized by Guglielmini (1668; 1705) and finally confirmed about 100 years later by Romé de l'Isle (1772; 1783). According to them, every chemical species has its specific crystal form.

In contrast to the extensive experimental experience and the accuracy in describing crystals, the generation and the growth process of crystals long remained subject to the speculations and mystifications deriving from the Middle Ages. The growth of crystals was often considered to be similar to that of plants or animals, viz. connected with mystic powers and virtues. At that time, of course, it seemed difficult to understand how crystals can form from fluid, transparent and even microscopically clear solutions. It was the great experimentalist Boyle (1666; 1672) who observed that the nature of solution-grown crystals depends both on impurities and the rate of deposition, i.e. the growth rate. He also concluded from crystal forms and inclusions, partly fluid, that gem stones and other minerals are generated from solutions, too. Nevertheless, Boyle believed the growing process to be driven by non-materialistic and imponderable powers, and he also gave credence to the medical virtues of crystals, especially of gem stones. But Steno (1669) had already concluded that crystals grow by the attachment of material from outside and not by any vegetative mode of growth. However, Steno's statement became accepted only gradually, for instance by Hottinger (1698) and by Homberg (1692), who wrote that dendrites also grow in a simple way from outside. But even in the 18th century some notable scientists, for instance Leeuwenhoek (1685; 1703; 1705), still gave credence to vegetative growth modes.

During the 18th century significant progress was achieved both in the systematic description of crystals, especially minerals (cf. Capeller 1723; Linnaeus 1768), as well as

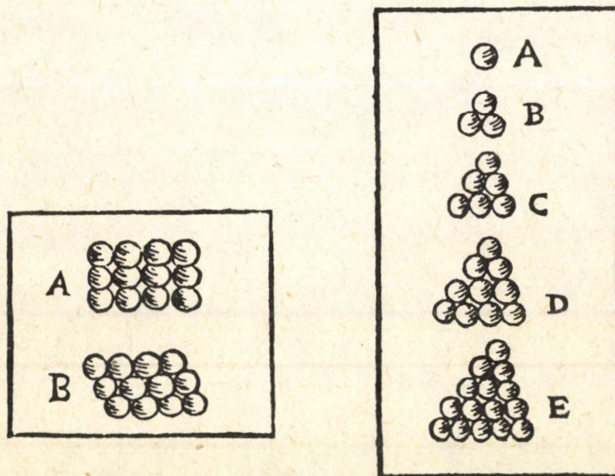


Fig. 2. Planar arrangements of close packed spheres. Reproduction from Johannes Kepler's "*Strena seu de Nive Sexangula*". Frankfurt: Tampach 1611

in crystal growth experiments. Fahrenheit (1724) discovered the supercooling of water and noted the release of heat when ice formation occurred. Towards the end of the century Lowitz (1795) in his extensive work reaffirmed the earlier implied requirement of supersaturation or supercooling for the initiation of crystal growth and described the now well-known features of supersaturated solutions. The supersaturation of a solution can be achieved both by evaporation or supercooling; the degree of supersaturation that can be attained depends on the particular salt and on the pretreatment of the solution. He also used seeding and recognized a specificity of different nucleating agents. From a mixed supersaturated solution, the separate salt that is used for seeding will be deposited. The identity of the crystallizing salt and the nucleating agent is not required in all cases.

Lowitz's work as well as the extended investigations of Leblanc (1802), Beudant (1817; 1818), Gay-Lussac (1813; 1819), Fuchs (1815; 1816; 1817) and others prepared the way for Mitscherlich's (1819) general formulations regarding isomorphism and epitaxy. Somewhat later (1822) Mitscherlich also discovered the phenomenon of polymorphism (originally of dimorphism), that all together led to a revision of the basic crystallo-chemical principles. At the same time Schweigger (1813) made the significant observation that a seed or nucleus must be of a certain size in order to initiate crystallization — the point of departure for the subsequent concepts of critical-sized nuclei. In the meantime the outstanding work of Haüy (1782; 1784) had appeared; after some precursors (Guglielmini 1688; 1705, Westfeld 1767, Bergman 1773; 1779), he propounded the view that continued cleavage of a crystal should ultimately result in the smallest possible unit, a "*molécule intégrante*", by a repetition of which the whole crystal is built up. The concept of a periodic crystal structure was now well established and, furthermore, the idea of molecular growth units was introduced to the crystallographic community, too. Despite Haüy's work, Weiss (1804) considered crystals to be anisotropic continua; he derived the crystal systems (1815) and discovered the law of rational intercepts (1816) and the zone law (1820). On the other hand, Seeber (1824) discarded the concept of polyhedral cleavage nuclei: he concluded from the compressibility of crystals that they are built by a parallelepipedic arrangement of spheres — that means a lattice. In the middle of the century Bravais (1849) derived the 14 lattice types, setting the periodicity of crystals on a sound footing. He hypothesized a correspondence of lattice and morphology according to which crystal faces are planes with a high density of lattice points.

Referring to experimental investigations, mention must be made of the excellent work of Löwel (1857). Following many early papers on the Na_2SO_4—H_2O system (e.g. Ziz 1815), he determined accurately the solubilities not only of the stable anhydrous Na_2SO_4 and its decahydrate (Glauber's salt, 1658), but of the metastable heptahydrate, too. His observation that from a supersaturated solution a crystal of the metastable heptahydrate crystallizes rather than the stable decahydrate afterwards prompted Ostwald (1897) to formulate his Law of Stages. Furthermore he established the expression of metastability (1893) and of a metastable region of supercooling or

supersolubility (1897), nowadays called the Ostwald–Miers region. Ostwald was also concerned with the critical size of nuclei, he gave an interpretation of Liesegang's rings as a supersaturation phenomenon (1897) and derived the thermodynamic formula of the enhanced solubility of small particles (1900). Liesegang (1896) generated his rings by placing a droplet of a solution of silver nitrate onto a layer of gelatine containing potassium chromate, and we may credit this as the introduction of the technique of crystal growth in a gel. Intensive investigations and speculations on the aspects of nucleation were stimulated further by the work of de Coppet (1872; 1875). According to his observations supersaturated solutions or supercooled melts remain stable for a limited interval of time depending on the size of the sample. De Coppet extended his experiments over exceptionally long periods up to several years until crystallization took place; he did in fact maintain solutions of Glauber's salt in the supersaturated condition for nearly 35 years, and these were still intact at the time of his last report (1907). De Coppet explained his results by the formation of crystal embryos via ordinary collisions — a first theory of homogeneous nucleation. This found considerable criticism, and there was a lengthy controversy between the followers of homogeneous and of heterogeneous nucleation theories. The first quantitative measurements of the linear growth rates were performed by Gernez (1882): he crystallized sulphur and phosphorus from their respective melts, using long glass capillary tubes. Later, in the early 20th century, Tammann (1898; 1903) became the leading exponent for quantitative measurements, both of nucleation and crystal growth rates.

But foremost, there is the masterly theoretical work of Gibbs (1878) on heterogeneous equilibria, but the value of this work was generally recognized only with great delay. Gibbs determined the energy needed to generate a nucleus and derived the equilibrium form of a crystal that fulfils the condition of minimum total free surface energy. But in a footnote he pointed out that the equilibrium form may determine the nature of small crystals only whereas the larger ones will be confined finally by such faces onto which the attachment of material proceeds most slowly. Curie (1885) independently concluded in a short meaningful paper that the stable form, as he said, of a crystal is given by the minimum of the sum of the products of surface tension times surface area. This led to the well-known construction of a crystal form by Wulff (1895; 1901). While the kinematic theory of crystal growth, based on the velocities of advance of the individual crystal faces, was developed by Becke (1894), by Johnsen (1910), and by Gross (1918) another alternative approach was introduced by Noyes and Whitney (1897) and by Nernst (1904) and Brunner (1904), who treated crystal growth as a diffusion-limited phenomenon. Early in this century Laue's invention of X-ray diffraction and the determination of crystal structures based thereon (first performed by Bragg) gave definite knowledge of the inner construction of crystals.

In the meantime, significant progress was achieved in crystal growth technology. A great deal of effort was made in the field of experimental mineralogy, stimulated not least by the search for recipes to make synthetic gem-stones. As early as 1837 Gaudin,

and then Böttger (1839) and Elsner (1839) prepared small ruby crystals by melting a mixture of potassium alum and potassium chromate. Fremy (1891), cooperating with Verneuil in his attempts to grow gem-stones and other crystals from high temperature solutions, used large crucibles containing up to 50 litres of melt but they got mm-sized crystals only, mainly because of the poor temperature stability of their furnaces. Among the numerous efforts to synthesize diamonds, only those of Hannay (1880) and Moissan (1894) are mentioned here. Hannay accidentally found diamonds when he heated a mixture of lithium, paraffin and bone oil in thick-walled iron tubes. Moissan claimed to have obtained diamonds by quenching a melt of iron saturated with carbon. At the beginning of the present century Verneuil (1902; 1904) published his well-known flame fusion method by which he succeeded in growing large ruby crystals. Soon afterwards the industrial production of synthetic rubies was established; today this method is still followed throughout the world in some 20 factories with an estimated 1000 growing machines in nearly the same manner as invented by Verneuil. These figures prove that the Verneuil method is well in advance of all other growth methods.

The old method of sublimation was used by Durocher (1849; 1851) to prepare crystals of transition metal sulphides, passing hydrogen sulphide over the corresponding chlorides; even at that time he used the expression "transportation". Nowadays it is often Lorenz (1891) — who reproduced and discussed many of the older experiments — who was credited with having reinvented the sublimation technique for crystal growth technology. Concerning the growth methods from solution, G. Wulff (1895) was credited with the first construction of an apparatus with a rotating vessel, in this way breaking with the principle of avoiding any movement in crystallization experiments. But he had, in fact, already had a precursor in the less-known L. Wulff (1886). Johnsen (1915) invented a vertical setup with a rotating crystal. The two-tank technique was patented in 1910 by Krüger and Finke, but it had already been described accurately as early as 1852 by Payen. Following some precursors in experiments with high pressure hydrothermal solutions, Spezia (1905; 1906; 1909) succeeded first in the hydrothermal synthesis of larger sized quartz crystals.

Concerning the development of melt growth, Nacken (1915; 1916) grew single crystals from the melt on a cooled rod or seed crystal dipped into the melt. This method was modified by Kyropoulos (1926), who additionally slowly raised the growing crystal. Thus, his set-up became similar to Czochralski's (1918) method of pulling crystals from the melt. With the aim of growing metal crystals with a constant diameter, Gomperz (1921) put a platelet of mica onto the melt surface and pulled the crystal through a hole in the centre of this platelet: in this way the technique of pulling profiled crystals was born; in the fifties this technique was reinvented by Stepanov (1959) and his co-workers. Concerning the freezing in crucibles, in 1914 Tammann had performed the growth of metal single crystals in small vertically arranged tubes by directional solidification; this method was made more sophisticated by Obreimov and Schubnikov (1924), who cooled the tip of the crucible blowing air onto it. With the same purpose Bridgman (1923; 1925) introduced the technique of lowering the crucible. In the 30s his

method was applied by Stockbarger (1936) to grow large alkali halide crystals. The gradient freezing of large crystals in a resting crucible was performed by Stöber (1925). Somewhat later, Kapitza (1928) used the horizontal gradient technique in an open boat to grow bismuth crystals; in addition, he attempted to grow bismuth crystals by vertical zone melting too, the invention of which as a crystal growth technique must be credited to him. Finally we mention that in this period the preparation of aluminium crystals by the strain-anneal technique was performed by Sauveur (1912; cf. Carpenter 1922; 1926), the recrystallization of tungsten by the Pintsch technique (1916; cf. Böttger 1917) and the deposition of tungsten and other refractory metals from the vapour phase of halide compounds onto a hot wire by Koref (1922) and by van Arkel (1923; 1925).

Returning again to the theory of crystal growth, it was Volmer (1922) who introduced the adsorption of growth units onto the crystal surface, their diffusion along the surface and the generation of two-dimensional nuclei. Then, Volmer and Weber (1926) extracted from a thermodynamic treatment the basic expression of the rate of nucleation. This expression gives an exponential dependence of the nucleation rate on the work of nucleation ("*Keimbildungsarbeit*"). Due to its kinetic peculiarity, in this expression the preexponential factor still remains undetermined. The first kinetic approach to nucleation was given by Farkas (1927). At the same time the molecular kinetic theory of crystal growth was founded by Kossel (1927), introducing the half crystal position ("*Halbkristallage*") and by Stranski (1928), introducing the detachment energy and somewhat later (1931; 1932; 1934; 1935) in common with Kaischev the average detachment energy. After this, Becker and Döring (1935) published their kinetic theory of nucleation.

Concerning the relationship between crystal structure and habit Donnay and Harker (1937) extended the principle of Bravais: they considered the influence of screw axis and glide mirror planes on the density of lattice points at the particular crystal faces and derived a morphological aspect for each space group. In the fifties Hartmann and Perdok (1955), regarding the actual structure of a crystal, introduced the concept of PBC-vectors that denote the chains of strongest bonds in a crystal — a concept that has proven fruitful with respect both to crystal habit and growth. Somewhat before this, Burton (1949), Cabrera (1949) and Frank (1949) founded the well-known theory of spiral growth, the nowadays so-called BCF theory, solving a hitherto marked discrepancy between growth theories and measurements of actual growth rates. Also in the early fifties, Burton, Prim and Slichter (1953) derived their frequently cited expression for the effective distribution coefficient, and Rutter and Chalmers (1953) described the phenomenon of constitutional supercooling that is caused by a drop in melt temperature due to the enrichment of impurities adjacent to the surface of a growing crystal. Both papers became very important with regard to the practical aspects of crystal growth.

The onset of the modern development of crystal growth technology, dating from the Second World War was boosted mainly by the demand for crystals for electronics, optics, and scientific instrumentation. Starting mostly from long-known growth

methods, the growth technologies had to be raised to a very high and advanced level to fulfil the increasing demands in crystal size and quality. Frequently, for economic as well as for political reasons, similar developments were performed independently at several places, an example of which was the hydrothermal synthesis of quartz crystals in the forties. This also holds for the production of semiconductor crystals, starting from the fifties. The main progress in the latter field was marked by the adaptation of the Czochralski method to grow germanium crystals by Teal and Little (1950) and by Roth and Taylor (1952), by the zone melting invented by Pfann (1952; 1953) and subsequently by the floating zone technique for silicium, invented by Keck and Golay (1953; 1954) and by Emeis (1954). Finally, the old dream of crystal growers, the synthesis of man-made diamonds, perhaps the most spectacular event in the history of crystal growth, was published first by Bundy, Hall, Strong and Wentorf (1955). But it was realized in at least three places in the world independently and at the very same time, all using surprisingly similar apparatus, that means high pressure equipment of the belt type. Nowadays about a half of the diamonds used for industry are produced synthetically.

The increasing investigations and efforts in the field of crystal growth have shown up in the literature as well as at relevant conferences. Nowadays, most of the important papers are concentrated in the two leading journals "Journal of Crystal Growth" and "Crystal Research and Technology". The latter was founded in 1966 as "Kristall und Technik", the "Journal of Crystal Growth" was founded one year after in 1967. But now as before papers on crystal growth appear in many other periodicals, too. The first conference dedicated especially to crystal growth was held at Bristol by the Faraday Society in 1949; a further one was held at Cooperstown (N.Y.) in 1958. The Soviet Union Conferences started as early as in 1956. The International Conferences on Crystal Growth were founded in 1966 at Boston. Since 1976, in Zürich, there have been European Conferences, too. Besides other regional conferences not mentioned here, we have the well-known Hungarian Conferences on Crystal Growth to the third of which this paper was dedicated.

Appendix 1

The history of crystal growth — Chronology

Year	Name	Contribution
1540	Birringuccio	Recrystallization of saltpetre
1556	Agricola	Production of various salts
1602	Caesalpinus	Typical forms of solution grown from crystal species
1611	Kepler	Structure of snow crystals
1658	Glauber	Crystallization of Glauber's salt
1665	Hooke	Structure of crystals
1666	Boyle	Influence of impurities and growth rate on crystal forms
1669	Steno	Law of constancy of crystal angles; crystal growth via addition of material from outside
1669	Bartholinus	Birefringence of calcite
1685	Leeuwenhoek	Description of crystals, also by microscopic observation
1688	Guglielmini	Correspondence of crystal forms and chemical species
1690	Huygens	Structural interpretation of birefringence
1692	Homberg	Crystal growth via addition of material from outside
1698	Hottinger	
1724	Fahrenheit	Supercooling of water
1767	Westfeld	Building of crystals from small growth units (cleavage nuclei)
1773	Bergman	
1782	Haüy	
1783	Romé de l'Isle	Description of crystals; change in nature of rocksalt by means of urea
1795	Lowitz	Supersaturation and crystallization of salt solutions; seeding
1813	Schweigger	Minimum size of crystal nuclei
1815	Fuchs	"Vicariates"
1815	Weiss	Crystal systems
1816	Weiss	Law of rational intercepts
1819	Mitscherlich	Isomorphism; epitaxy
1822	Mitscherlich	Polymorphism (dimorphism)
1824	Seeber	Lattice structure of crystals
1830	Hessel	Crystal classes
1837	Gaudin	Ruby from high-temperature solution
1839	Böttger; Elsner	
1839	Miller	Miller's indices
1849	Bravais	Lattice types; correspondence of lattice type and crystal form
1851	Durocher	Vapour growth of sulphide crystals; "transportation"
1852	Payen	Solution growth by the two-tank technique
1857	Löwel	Solubilities in the Na_2SO_4—H_2O system; metastable solutions
1865	Gernez	Reciprocal pairs of salts
1865	Marangoni	Liquid surface phenomena
1872	de Coppet	Spontaneous nucleation
1876	Sohncke	Groups of motion
1878	Gibbs	Heterogeneous phase equilibria
1880	Hannay	Man-made diamonds
1882	Gernez	Measurements of growth rates
1883	Barlow	Sphere packings
1885	Curie	Minimum surface energy of growth forms
1886	L. Wulff	Solution growth in a rotating vessel
1889	Stefan	Stefan's problems
1891	Schoenflies	Space groups
1891	Fedorov	

1893	Ostwald	Metastable region of supersaturation; critical size of nuclei
1894	Becke	Kinematic development of forms
1895	G. Wulff	Wulff's rule of construction of crystal forms; solution growth in a rotating vessel
1896	Liesegang	Liesegang's rings (crystal growth in a gel)
1897	Ostwald	Step rule
1897	Noyes & Whitney	Dissolution of crystals controlled by diffusion
1898	Tammann	Measurements of nucleation and growth rates
1900	Ostwald	Dependence of solubility on particle size
1902	Verneuil	Flame fusion technique (ruby)
1904	Nernst & Brunner	Diffusion layer on a crystal surface
1905	Spezia	Hydrothermal synthesis of quartz
1910	Johnsen	Kinematic development of forms
1910	Krüger & Finke	Solution growth by the two-tank technique
1911	Artemiew	Growth experiments with crystal spheres
1912	Sauveur	Strain-anneal technique
1912	Laue, Friedrich & Knipping	X-ray diffraction by crystals
1913	Bragg	X-ray crystal structure determination
1914	Johnston	Diffusion technique to grow lowly soluble compounds
1914	Tammann	Directional solidification of metals
1915	Johnsen	Solution growth using a rotating seed
1915	Nacken	Melt growth using a cooled seed
1916	Schaller & Orbig	Recrystallization of a tungsten wire (Pintsch technique)
1918	Czochralski	Pulling of metal crystals from their melt
1921	Gomperz	Pulling of profiled metal crystals from the melt using a mica orifice
1922	Volmer	Adsorption and surface diffusion; two-dimensional nuclei
1922	Koref	Vapour phase deposition (hot wire technique)
1923	Van Arkel	Vapour phase deposition (hot wire technique)
1923	Bridgman	Melt growth by lowering the crucible
1925	Stöber	Melt growth by the gradient technique
1926	Kyropoulos	Melt growth by using a cooled seed
1926	Volmer & Weber	Thermodynamic theory of nucleation
1927	Farkas	Kinetic approach to nucleation
1927	Kossel	Half crystal position – kinetic theory of crystal growth
1928	Stranski	Detachment energy – kinetic theory of crystal growth
1928	Kapitza	Crystal growth by zone melting and by the horizontal gradient technique
1935	Becker & Döring	Kinetic theory of nucleation
1936	Stockbarger	Melt growth by lowering the crucible
1937	Donnay & Harker	Morphological aspect
1949	Burton, Cabrera & Frank	Spiral growth (BCF theory)
1950	Teal & Little	Czochralski growth of germanium
1952	Pfann	Zone melting
1953	Keck & Golay	Floating-zone technique
1953	Rutter & Chalmers	Constitutional supercooling
1953	Burton, Prim & Slichter	Effective distribution coefficient
1955	Hartmann & Perdok	"PBC-Vectors"
1955	Bundy, Hall, Strong & Wentorf	High pressure synthesis of diamond

Appendix 2

Conferences on Crystal Growth

Conference			Publication
Bristol	1949		"Crystal Growth", Disc. Faraday Soc. No. 5, Butterworth, London: 1959.
Cooperstown (N.Y.)	1958		Doremus, R. H.; Roberts, B. W.; Turnbull, D. (Eds), "Growth and Perfection of Crystals", John Wiley New York/London 1962.
Boston	1966	ICCG[1]—1	Peiser, H. S. (Ed.), J. Phys. Chem. Solids, Suppl. N1 Pergamon, Oxford/New York: 1967.
Birmingham	1968	ICCG—2	Frank, F. C.; Mullin, J. B.; Peiser, H. S. (Eds), J. Crystal Growth, *3*, *4*, 1968.
Marseille	1971	ICCG—3	Laudise, R. A.; Mullin, J. B.; Mutaftschiev, B. (Eds), J. Crystal Growth, *13*, *14*, 1972.
Tokyo	1974	ICCG—4	Jackson, K. A.; Kato, N.; Mullin, J. B. (Eds), J. Crystal Growth, *24*, *25*, 1974.
Cambridge (Mass.)	1977	ICCG—5	Parker, R. L.; Chernov, A. A.; Cullen, G. W.; Mullin, J. B. (Eds), J. Crystal Growth, *42*, 1977.
Moscow	1980	ICCG—6	Givargizov, E. I.; Elwell, D.; Ghez, R.; Kuznetsov, F. A.; Peiser, H. S. (Eds), J. Crystal Growth, *52*, 1981.
Stuttgart	1983	ICCG—7	Tolksdorf, W.; Mullin, J. B. (Eds), J. Crystal Growth, *65*, 1983.
Zürich	1976	ECCG[2]—1	Kaldis, E.; Scheel, H. J. (Eds), Current Topics in Materials Science Vol. 2 North-Holland, Amsterdam: 1977.
Lancaster	1979	ECCG—2	(Proceedings)
Prague	1982	EMCG'82[3]	(Proceedings)

[1] International Conference on Crystal Growth
[2] European Conference on Crystal Growth
[3] European Meeting on Crystal Growth

Appendix 2 *(continued)*

Conference			Publication			
			Vol.	Rost kristallov[4] Main editor	Publisher	
Moscow	1956	I. VSRK[6]	I	Shubnikov, A. V.; Sheftal, N. N.	Acad. Sc. USSR	Moscow 1957
			II[5]	Shubnikov, A. V.; Sheftal, N. N.	Acad. Sc. USSR	Moscow 1959
Moscow	1959	II. VSRK	III	Shubnikov, A. V.; Sheftal, N. N.	Acad. Sci USSR	Moscow 1961
			IV[5]	Shubnikov, A. V.; Sheftal, N. N.	Nauka	Moscow 1964
Moscow	1963	III. VSRK	V	Shubnikov, A. V.; Sheftal, N. N.	Nauka	Moscow 1965
			VI	Sheftal, N. N.	Nauka	Moscow 1965
Moscow	1966	MSRK[7]	VII	Sheftal, N. N.	Nauka	Moscow 1967
			VIII	Sheftal, N. N.	Nauka	Moscow 1968
Leningrad	1969	FSRK[8]	IX	Sheftal, N. N.; Givargizov, E. I.	Nauka	Moscow 1972
			X[5]	Sheftal, N. N.	Nauka	Moscow 1974
Cachkadzor	1972	IV. VSRK	XI XII	Chernov, A. A.	University Erevan	Erevan 1975
			XII	Chernov, A. A.	University Erevan	Erevan 1977
Tbilisi	1977	V. VSRK	XIII	Givargizov, E. I.	Nauka	Moscow 1980
Budapest	1976	HCCG[9]—1	Abstracts, Budapest 1976			
Mátrafüred	1979	HCCG—2	Acta Physica Hungarica *47*, (1979)			
Budapest	1983	HCCG—3	Acta Physica Hungarica, this volume			

[4] English translation: "Growth of Crystals", Consultants Bureau, New York and London from 1958
[5] Interconference Volumes of "Rost kristallov"
[6] Vsesojuznoe soveshchanie po rostu kristallov (All-Union Conference on Crystal Growth)
[7] Mezhdunarodni simpozium po rostu kristallov (International Symposium on Crystal Growth)
[8] Simpozium po rostu kristallov pamyati Fedorova (Symposium on Crystal Growth in memory of Fedorov)
[9] Hungarian Conference on Crystal Growth

Appendix 3

Sources

Adams, F. D., The Birth and Development of Geological Sciences, Dover, New York: 1938 and 1954.

Bohm, J., Die historische Entwicklung der Kristallzüchtung, Crystal Research & Technol, *16*, 275—292, 1981.

Buckley, H. E., Crystal Growth, John Wiley, New York: Chapman and Hall, London: 1951.

Burke, J. G., Origins of the Science of Crystals, University of California Press, Berkeley, Los Angeles: 1966.

Caldwell, H. B., Crystallization. Chem. and Met. Eng., *42*, 213—214, 1935.

Darmstaedter, E.: Die Alchemie des Geber, Springer Berlin, 1922 (Repr. Sändig, Wiesbaden, 1969).

Eitel, W., 1. Mineralsynthese, 2. Kristallzüchtung: in: Handb. d. Arbeitsmethoden in der anorg. Chemie (Hrsg. Tiede, E.; Richter, F.) Bd. 4 S. 391—447 u. 448—473 de Gruyter, Berlin, Leipzig, 1926.

Elwell, D., Scheel, H. J., Crystal Growth from High-Temperature Solutions: Academic Press London, New York, San Francisco 1975.

Groth, P., Entwicklungsgeschichte der mineralog. Wissenschaften, Springer, Berlin, 1926.

Hook, A. van, Crystallization. Reinhold, New York: Chapman and Hall London:, 1961.

Kobell, F. v., Geschichte der Mineralogie von 1650—1860. J. G. Cottasche Buchhandlung München: 1864.

Leicester, H. M., Klickstein, H. S., A Source Book in Chemistry: McGraw-Hill, New York, Toronto, London 1952.

Lenz, H. O., Mineralogie der alten Griechen und Römer. E. F. Thienemann Gotha, 1861 (Repr. Sändig Wiesbaden, 1966).

Marx, C. M., Geschichte der Crystallkunde, D. R. Marx'sche Buchhandlung Carlsruhe und Baden 1825.

Niggli, P., Die Krystallologia von Johann Heinrich Hottinger (1698). Sauerländer Aarau:, 1946.

Ray, P., History of Chemistry in Ancient and Medieval India. Indian Chem. Soc. Calcutta: 1956 (cited after Elwell u. Scheel 1975).

Schneer, C. J. (Ed.), Crystal Form and Structure, Dowden, Hutchinson and Ross Stroudsburg (Penns. USA) 1977.

Schoen, H. M., Grove, C. S., Palermo, J. A., The early history of crystallization. J. Chem. Education *33*, 373—375, 1956.

Shafranovskij, I. I., Istorija Kristallografii. I u. II, Nauka Leningrad, 1978 and 1980.

Sohncke, L., Entwickelung einer Theorie der Krystallstruktur, B. G. Teubner Leipzig 1879.

Spangenberg, K., Wachstum und Auflösung der Kristalle, in: Handwörterbuch der Naturwissenschaften. Gustav Fischer Jena *10*, 362—401, 1934.

Wilke, K.-Th., Kristallzüchtung. Dtsch. Verlag Wiss. Berlin, 1973.

Appendix 4

Bibliography of original papers

Agricola, G., De re metallica libri XII. Froben Basel 1556 (New German translation by Schiffner, VDI-Verlag, Berlin, 1928, Düsseldorf 1953; cited after Georgius Agricola — Ausgewählte Werke, Bd. I, Dtsch. Verl. Wiss. Berlin 1956).

Arkel, A. E. van, Unikristallijn Wolfraam, Physica *3*, 76—87, 1923.

Arkel, A. E. van; de Boer, J. H., Darstellung von reinem Titan-, Zirkonium-, Hafnium- und Thoriummetall. Z. anorg. allg. Chem. *148*, 345—350, 1925.

Artemiew, D. N., Die Krystallisation der Kugeln als eine besondere Methode der krystallographischen Forschung. Z. Krist, *48*, 417—433, 1911.

Barlow, W., Probable nature of the internal nature of crystals, Nature *29*, 186—188; 205—207; 404, 1883, 1884.

Bartholinus, E., Experimenta Crystalli Islandici disdiaclastici, quibus mira et insolita Refractio detegitur Hafniae (Kopenhagen) 1669.

Becke, F., Der Aufbau der Kristalle aus Anwachskegeln, Lotos N. F. 14, 1894.

Becker, R., Döring, W., Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. Ann. Phys., *24*, 719—572, 1935.

Bergman, T., Variae crystallorum formae a spatho ortae, Nova Acta Reg. Societ. Scient. Upsalensis (Uppsala) *1*, 150—155, 1773.

Bergman, T., Opuscula Physica et Chemica, Uppsala 1779; Bd. II, 1780.

Beudant, F. S., Recherches sur les causes qui determinent les variations des formes cristallines d'une même substance minerale, Paris 1817; Ann. Mines, 1818.

Birringuccio, V., De la Pirotechnia. Venedig 1540 (Transl. into Engl.: The Pirotechnia of Vanuccio Birringuccio. Amer. Inst. Mining and Metallurg. Eng. New York: 1942.)

Böttger, R., Chemische Notizen: 6. Über die Anfertigung der Rubine auf dem Wege der Kunst (p. 85). Ann. d. Pharm. (Heidelberg), *29*, 75—88, 1839.

Böttger, W., Herstellung fadenförmiger Kristalle (Kristallfäden für Glühlampen) nach dem Verfahren der Firma Julius Pintsch A.-G. Z. Elektrochem., *23*, 121—126, 1917.

Boodt, Anselmus Böetius de, Gemmarum et lapidarum historia, Hanau, 1609.

Boyle, R., The Origine of Formes and Qualities, Oxford, 1666.

Boyle, R., An Essay about the Origine and Virtues of Gems, 1672. (Repr. Hafner New York 1972).

Bravais, A., Etudes cristallographiques. I: Du Cristal considéré comme un simple assemblage de points. p. 101—194. II: Du Cristal considéré comme un assemblage de molécules polyatomiques. p. 194—247. Paris, 1849.

Bridgman, P. W., The compressibility of thirty metals as a function of pressure and temperature, Proc. Amer. Acad. Arts Sci. (Boston), *58*, 165—242, 1923.

Bridgman, P. W., Certain physical properties of single crystals of tungsten, antimony, bismuth, tellurium, cadmium, zinc, and tin, Proc. Amer. Acad. Arts Sci. (Boston), *60*, 303—383, 1925.

Brunner, E., Reaktionsgeschwindigkeit in heterogenen Systemen, Z. phys. Chem., *47*, 56—102, 1904.

Bundy, F. P., Hall, H. T., Strong, H. M., Wentorf, R. H., Man-made diamonds, Nature (London), *176*, 51—55, 1955.

Burton, W. K.; Cabrera, N.; Frank, F. C., Role of dislocations in crystal growth, Nature, *163*, 398—399, 1949.

Burton, W. K., Cabrera, N., Crystal growth and surface structure, I, Discussion Faraday Soc., *5*, 33—39, 1949.

Burton, J. A.; Prim, R. C.; Slichter, W. P., The distribution of solute in crystals grown from the melt, J. Chem. Phys., *21*, 1987—1991, 1953.

Cabrera, N., Burton, W. K., Crystal growth and surface structure, II, Discussion Faraday Soc., *5*, 40—48, 1949.

Caesalpinus (Caesalpino), Andrea, De metallicis, 1583 and Nürnberg 1602 (cited after van Hook 1961 and Adams 1938).

Cappeller, M. A., Prodromus crystallographiae, Luzern, 1723 (Transl. Karl Mieleitner, München, 1922).

Carpenter, H. C. H., Elam, C. F., The production of single crystals of aluminium and their tensile properties, Proc. Royal Soc. (London), *A 100*, 329—353, 1922.

Carpenter, H. C. H., The production of single crystals of metals and some of their properties, J. Inst. Metals (London), *35*, 409—438, 1926.

Coppet, L. C., de, Sur la température de cristallisation spontanée de la solution sursaturée de sulfate de soude, Bull. Soc. chim. (Paris), *17*, 146—149, 1872.

Coppet, L. C., de, Théorie de la surfusion et de la sursaturation, d'après les principes de la théorie mécanique de la chaleur, Ann. Chim. et Phys. (Paris), *6*, 275—288, 1875.

Coppet, L. C., de, Recherches sur la surfusion et la sursaturation, Ann. Chim. et Phys. (Paris), *10*, 457—527, 1907.

Curie, M. P., Sur la formation des cristaux et sur les constantes capillaires de leurs différentes faces, Bull. Soc. Franc. Mineral. (Paris), *8*, 145—150, 1885.

Czochralski, J., Ein neues Verfahren zur Messung der Kristallisationsgeschwindigkeit der Metalle, Z. phys. Chem., *92*, 219—221, 1918.

Donnay, J. D. H., Harker, D., A new law of crystal morphology extending the law of Bravais. Amer. Mineralogist, *22* (5), 446—447, 1937.

Durocher, J., Recherches sur les dépôts métallifères et sur leur mode de formation, Comptes Rendus (Paris), *28*, 607—609, 1849.

Durocher, J., Production artificielle, par voie sèche, des principaux minéraux contenus dans les gites métallifères, Comptes Rendus (Paris), *32*, 823—826, 1851.

Elsner, L., Über die Darstellung künstlicher Rubine, Journ. prakt. Chem. (Leipzig), *17*, 175—176, 1839.
Emeis, R., Tiegelfreies Ziehen von Silicium-Einkristallen, Z. Naturforsch., *9a*, 67, 1954.
Fahrenheit, D. G., Experimenta & observationes de congelatione aquae in vacuo, Phil. Trans. Royal Soc. (London), *33*, 78—84, 1724.
Farkas, L., Keimbildungsgeschwindigkeit in übersättigten Dämpfen, Z. Physik. Chem., *125*, 236—242, 1927.
Fedorow, E. S. von, Symmetrie der regelmäßigen Systeme von Figuren, (Russ.) Verh. kaiserl. Min. Ges. Petersburg, *27*, 448, 1891.
Frank, F. C., The influence of dislocations on crystal growth, Discussions Faraday Soc., *5*, 48—54, 1949.
Frémy, E., Synthèse du rubin, Dunod Paris: 1891.
Fuchs, J. N., Über den Gehlenit, ein neues Mineral aus Tirol, J. Chem. Phys. v. J. S. C. Schweigger (Schweiggers J.), Nürnberg *15*, 377—386, 1815.
Fuchs, J. N., Ueber einige phosphorsaure Verbindungen, J. f. Chem. u. Phys. v. J. S. C. Schweigger (Schweiggers J.) Nürnberg, *18*, 288—296, 1816.
Fuchs, J. N., Ueber den Arragonit und Strontianit. J. f. Chem. u. Phys. v. J. S. C. Schweigger (Schweiggers J.) Nürnberg, *19*, 113—137, 1817.
Gaudin, A., Note sur la formation artificielle du corindon, Comptes Rendus (Paris), *4*, 999—1000, 1837.
Gay-Lussac, J. L., De l'influence de la pression de l'air sur la cristallisation des sels, Ann. Chim. et Phys. (Paris), *87*, 225—236, 1813.
Gay-Lussac, J. L., Sur la dissolubilité des sels dans l'eau, Ann. Chim. et Phys. (Paris), *11*, 296—315, 1819.
Gernez, D., Sur la cristallisation des dissolutions salines sursaturées et sur la présence normale du sulfate de soude dans l'air, Compt. Rend. (Paris), *60*, 833—837, 1865. Nouvelles études sur les dissolutions sursaturées, C. R. *60*, 1027—1030, 1865. Sur les causes d'erreur que présente l'étude des dissolutions sursaturées, C. R. *61*, 71—73, 1865. Addition à des notes précédentes sur la cristallisation des dissolutions salines sursaturées, C. R, *61*, 289—290, 1865. Sur les solutions sursaturées, C. R. *61*, 847—850, 1865.
Gernez, D., Recherches sur la durée de la solidification des corps surfondus., Compt. Rend. (Paris), *95*, 1278—1282, 1882.
Gibbs, J. W., On the equilibrium of heterogeneous substances, Trans. Connecticut Acad., *3*, 108—248 and 353—524 (1875—1878) (Abstr. Amer. J. Sci. (3) *16*, 441—458, (1878); also: The Collected Works of J. Willard Gibbs, Longmans, Green and Co New Haven 1928; Repr. 1948.
Glauber, J. R., Opera Chymica. Frankfurt, 1658.
Gomperz, E. von, Untersuchungen an Einkristalldrähten. Z. Phys., *8*, 184—190, 1921/22.
Gross, R., Zur Theorie des Wachstums- und Lösungsvorganges kristalliner Materie, Abh. d. math.-phys. Kl. d. Sächs. Ges. d. Wiss., Leipzig, *35*, 137—202, 1918.
Guglielmini, D., Riflessioni filosofiche dedotte dalle figure de'sali, Bologna, 1688.
Guglielmini, D., De salibus dissertatio epistolaris physico-medico-mechanica, Venedig, 1705.
Hannay, J. B., On the artificial formation of the diamond, Proc. Roy. Soc. (London), *30*, 188—189, 1880.
Hartman, P., Perdok, W. G., On the relations between structure and morphology of crystals. I., II. and III. Acta Cryst., *8*, 49—52; 521—529, 1955.
Haüy, R.-J., Extrait d'une mémoire sur la structure des cristaux de grenat., J. Phys. (Paris), *19*, 366—370, 1782. Extrait d'une mémoire sur la structure des spaths calcaires, J. Phys. (Paris), *20*, 33—39, 1782.
Haüy, R.-J., Essai d'une théorie sur la structure des cristaux, Paris, 1784.
Hessel, J. F. Ch., Krystall, Gehlen's physikal. Wörterbuch, *5*, 1290, 1830.
Homberg, G., Mem. Acad. Royal Sci. (Paris), 145—152, 1692.
Hooke, R., Micrographia, London, 1665.
Hottinger, J. H., Krystallologia, dissertatio de crystallis, harum naturam, admentem veterum, Bodmer Zürich: 1698.
Huygens, C. C., Traité de la lumière, Leyden, 1690.
Isle, J. B. L. Romé de L', Essai de cristallographie, Paris, 1772.
Isle, J. B. L. Romé de L', Cristallographie, ou description des formes propres à tous les corps du règne minéral, Paris, 1783.
Johnsen, A., Wachstum und Auflösung der Kristalle, Engelmann, Leipzig: 1910.
Johnsen, A., Optisches Drehungsvermögen von Lithiumsulfat-Monohydrat, Zentralbl. Mineral. Geol. Paläont. (Stuttgart), 233—243, 1915.
Johnston, J., The utilization of diffusion processes in the preparation of pure substances, J. Amer. Chem. Soc., *36*, 16—19, 1914.

Kapitza, P., The study of the specific resistance of bismuth crystals and its change in strong magnetic fields and some allied problems. Proc. Roy. Soc. London, *A119*, 358—442, 1928.
Keck, P. H., Golay, M. J. E., Crystallization of silicon from a floating liquid zone, Phys. Rev. (2), *89*, 1297, 1953.
Keck, P. H., Floating zone crystallization of silicon, Physica (Utrecht), *20*, 1059, 1954.
Kepler, J., Strena seu de nive sexangula, Frankfurt a. M, 1611.
Koref, F., Versuche über das Weiterwachsen von Metallkristallen durch Abscheiden aus der Gasphase, Z. Elektrochem., *28*, 511—517, 1922.
Kossel, W., Zur Theorie des Kristallwachstums, Nachr. Ges. Wiss. Göttingen, Math.-phys. Klasse, 135—143, 1927.
Krüger, F., Finke, W., Kristallisationsvorrichtung, Deutsches Reichspatent DRP 228 246, 1910.
Kyropoulos, S., Ein Verfahren zur Herstellung großer Kristalle, Z. anorg. u. allg. Chem., *154*, 308—313, 1926.
Leblanc, N., La Cristallotechnie, Paris, 1802.
Leeuwenhoek, A. van, Concerning the various figures of the salts contained in the several substances, Phil. Trans. Royal Soc. (London), *15*, 1073, 1685.
Leeuwenhoek, A. van, On the dissolution of silver, Phil. Trans. Royal Soc. (London), *23*, 1403—1443, 1703.
Leeuwenhoek, A. van, Concerning the figures of the salts of crystals, Phil. Trans. Royal Soc. (London), *24*, 1906—1917, 1705.
Liesegang, R. E., "A-Linien". Liesegang's photographisches Archiv, Düsseldorf, *21*, 321—326, 1896.
Linnaeus, C., Systema naturae III. Holmiae, Stockholm, 1768.
Lorenz, R., Mineralsynthetische Versuche. Über die Einwirkung trockenen Schwefelwasserstoffs auf einige Metalle, Ber. Dtsch. chem. Ges. (Berlin), *24*, 1501—1510, 1891.
Löwel, M. H., Observations sur la sursaturation des dissolutions salines, Ann. Chim. et Phys. (Paris), (3) *39*, 32—58, 1857.
Lowitz, J. T., Bemerkungen über das Krystallisiren der Salze, und Anzeige eines sicheren Mittels, regelmäßige Krystallen zu erhalten, Chem. Ann. v. D. L. v. Crell (Crell's Chem. Ann.), Helmstädt, *1*, 3—11, 1795.
Marangoni, C., Sull' espansione delle goccie di un liquido galleggiante sulla superficie di altro liquido, Fusi August, Pavia, 1865.
Marangoni, C., Ueber die Ausbreitung der Tropfen einer Flüssigkeit auf der Oberfläche einer anderen, Ann. d. Physik u. Chemie (Poggendorfs Ann.), Leipzig, *143*, 337—354, 1871.
Miller, W. H., A Treatise on Crystallography, London, 1839.
Mitscherlich, E., Über die Kristallisation der Salze in denen das Metall der Basis mit zwei Proportionen Sauerstoff verbunden ist, Abh. Königl. Akad. Wiss. Berlin, 427 to 437 (1818/1819) (also: Gesammelte Schriften von E. Mitscherlich, Berlin, 1896).
Mitscherlich, E, Über das Verhältnis der Krystallform zu den chemischen Proportionen. III.: Über die künstliche Darstellung der Mineralien aus ihren Bestandteilen, Abh. Königl. Akad. Wiss. Berlin 25—41, 1822/1823 (also: Gesammelte Schriften von Eilhard Mitscherlich, Berlin 1896).
Mitscherlich, E., Über das Verhältnis der Krystallform zu den chemischen Proportionen. IV.: Über die Körper, welche in zwei verschiedenen Formen krystallisiren, Abh. Königl. Akad. Wiss. Berlin 43—48, 1822/1823 (also: Gesammelte Schriften von Eilhard Mitscherlich, Berlin 1896).
Moissan, H., Nouvelles expériences sur la reproduction du diamant, Comptes Rendus (Paris), *118*, 320—326, 1894.
Nacken, R., Über das Wachsen von Kristallpolyedern in ihrem Schmelzfluß, Neues Jahrb. Mineral., Geol., Paläont. (Stuttgart), *2*, 133—164, 1915.
Nacken, R., Kristallzüchtungsapparate, Z. Instrumentenkunde, *36*, 12—20, 1916.
Nernst, W., Theorie der Reaktionsgeschwindigkeit in heterogenen Systemen, Z. phys. Chem., *47*, 52—55, 1904.
Neuhaus, A., Messungen von geometrischen Verschiebungsgeschwindigkeiten am NaCl und deren Abhängigkeit von Begrenzungsart, Konzentration und Lösungsgenossen, Z. Krist, *68*, 15—81, 1928.
Noyes, A. A., Whitney, W. R., Über die Auflösungsgeschwindigkeit von festen Stoffen in ihren eigenen Lösungen, Z. phys. Chem. (Leipzig), *23*, 689—692, 1897.
Obreimov, J., Schubnikov, L. W., Eine Methode zur Herstellung von Metall-Einkristallen, Z. Physik, *25*, 31—36, 1924.
Ostwald, W., Lehrbuch der Allgemeinen Chemie, 2. Aufl., Bd. II; S. 516. Engelmann Leipzig: 1893 (Bd. I 1891).

Ostwald, W., Studien über die Bildung und Umwandlung fester Körper. 1. Übersättigung und Überkaltung, Z. phys. Chem., *22*, 289—330, 1897.
Ostwald, W., A-Linien von R. E. Liesegang (Referat), Z. phys. Chem. (Leipzig), *23*, 365, 1897.
Ostwald, W., Über die vermeintliche Isomerie des roten und gelben Quecksilberoxyds und die Oberflächenspannung fester Körper, Z. phys. Chem. (Leipzig), *34*, 495, 1900.
Páyen, Méthode e cristallisation à l'aide d'une circulation continue, Comptes Rendus (Paris), *34*, 578—579, 1852.
Pfann, W. G., Principles of zone-melting, J. Metals, Trans. AIME, *194*, 747—753, 1952.
Pfann, W. G., Olsen, K. M., Purification and prevention of segregation in single crystals of germanium, Phys. Rev. (2), *89*, 322—323, 1953.
Pintsch (Julius Pintsch A. G.): Metalldrähte, -fäden oder -bänder und Verfahren zu ihrer Herstellung, Deutsches Reichspatent DRP 291 994, 1916.
Roth, L., Taylor, W. E., Preparation of germanium single crystals, Proc. Inst. Radio Eng., *40*, 1338—1341, 1952.
Rutter, J. W., Chalmers, B., A prismatic substructure formed during solidification of metals, Canad. J. Phys., *31*, 15—39, 1953.
Sauveur, Note on the crystalline growth of ferrite below its thermal critical range. Proc. Internat. Assoc. Test. Mat, *2* (11) (1912), Paper No. II. 6. (cited after Carpenter 1926)
Schoenflies, A. M., Kristallsysteme und Kristallstruktur, Leipzig, 1891.
Schweigger, J. S. C., J. Chem. Phys. v. J. S. C. Schweigger (Schweiggers J.) Nürnberg, *2*, 79, 1813.
Seeber, L. A., Versuch einer Erklärung des inneren Baus fester Körper, Gilbert's Annalen der Physik, *76*, 229—248; 349—372, 1824.
Sohncke, L., Die unbegrenzten regelmässigen Punktsysteme als Grundlage einer Theorie der Krystallstruktur, Karlsruhe, 1876.
Spezia, G., La pressione è chimicamente inattiva nella solubilità e ricostituzione del quarzo, Atti R. Accad. Sci. Torino, *40*, 254—262, 1905.
Spezia, G., Contribuzioni sperimentali alla cristallogenesi del quarzo, Atti R. Accad. Sci. Torino, *41*, 158—165, 1906.
Spezia, G., Sull' accrescimento del quarzo, Atti R. Accad. Sci. Torino, *44*, 95—107, 1909.
Stefan, J, Über die Theorie der Eisbildung, insbesondere über die Eisbildung im Polarmeere, Sitzungsber. Kaiserl. Akad. Wiss. Wien Math.-Nat. Cl, *98*, Abt. IIa, 965, 1889.
Steno, Nicolaus, De solido intra solidum naturaliter contento dissertationis prodromus. Florenz 1669 (Repr. W. Junk, Berlin, 1904; Transl. K. Mieleitner: „Vorläufer einer Dissertation über feste Körper, die innerhalb anderer fester Körper von Natur aus eingeschlossen sind" in: Ostwald's Klassiker der exakten Naturwissenschaften, Leipzig, 1923.
Stepanov, A. V., Novyj spocob polučenija izdelij (listov, trub, prutkov, raznogo profilja i t. d.) neposredstvenno iz rasplava, Ž. techn. Fiz. *29*, 381—393, 1959.
Stockbarger, D. C., The production of large single crystals of lithium fluoride, Rev. Sci. Instruments, *7*, 133—136, 1936.
Stöber, F., Künstliche Darstellung großer, fehlerfreier Kristalle, Z. Krist., *61*, 299, 1925.
Stranski, I. N., Zur Theorie des Kristallwachstums, Z. phys. Chem., *136*, 259—278, 1928.
Stranski, I. N., Beitrag zum Wachstum und Auflösen nichtpolarer Kristalle, Z. phys. Chem., *B 11*, 342—349, 1931.
Stranski, I. N., Wachstum und Auflösen der Kristalle vom NaCl-Typ, Z. phys. Chem., *B17*, 127—154, 1932.
Stranski, I. N., Kaischew, R., Über den Mechanismus des Gleichgewichts kleiner Kriställchen. Z. phys. Chem., *B26*, 100—113; 114—116; 312—316, 1934.
Stranski, I. N., Kaischew, R., Kristallwachstum und Kristallkeimbildung, Phys. Z., *36*, 393—403, 1935.
Stranski, I. N., Kaischew, R., Gleichgewichtsform und Wachstumsform der Kristalle, Ann. Physik, *23*, 330—338, 1935.
Strong, J., A method for growing large crystals of the alkali halides, Phys. Rev., *36*, 1663—1666, 1930.
Tammann, G., Über die Abhängigkeit der Zahl der Kerne, welche sich in verschiedenen unterkühlten Flüssigkeiten bilden, von der Temperatur, Z. phys. Chem. (Leipzig), *25*, 441—479, 1898.
Tamman, G., Kristallisieren und Schmelzen, Joh. Ambr. Barth Leipzig, 1903.
Tamman, G., Lehrbuch der Metallographie (p. 15) L. Voss Leipzig und Hamburg, 1914.
Teal, G. K.; Little, J. B., Growth of germanium single crystals, Phys. Rev., *25*, 16, 1950 and Bull. Amer. Phys. Soc., *78*, 647, 1950.

Verneuil, A., Production artificielle du rubis par fusion, Comptes Rendus (Paris), *135*, 791—794, 1902.
Verneuil, A., Memoire sur la reproduction artificielle du rubis par fusion, Ann. Chim. et Phys. (Paris), *3*, 20—48, 1904.
Volmer, M., Zum Problem des Kristallwachstums, Z. phys. Chem., *102*, 267—275, 1922.
Volmer, M.; Weber, A., Keimbildung in übersättigten Gebilden, Z. Phys. Chem., *119*, 277—301, 1926.
Weiss, C. S., Dynamische Ansicht der Krystallisation, Anhang, p. 365—389, in: Lehrbuch der Mineralogie von Haüy, translated by D. L. G. Karsten, C. H. Reclam Paris und Leipzig, 1804.
Weiss, C. S., Übersichtliche Darstellung der natürlichen Abteilungen der Krystallisationssysteme, Abh. Akad. Wiss. Berlin, 1815.
Weiss, C. S., Krystallographische Fundamentalbestimmung des Feldspathes, Abh. Akad. Wiss. Berlin, 1816.
Weiss, C. S., Über mehrere neubeobachtete Krystallflächen des Feldspathes und die Theorie seines Krystallwachstum im Allgemeinen, Abh. Akad. Wiss. Berlin, 1820.
Westfeld, C. F., Mineralogische Abhandlungen, Göttingen, 1767.
Wulff, G., Über Wachsthums- und Auflösungsgeschwindigkeit der Krystalle, Traveaux d. l. Soc. d. Natur. d. Varsovie (Warschauer Naturforscher-Gesellschaft) Ann. *6*, Livr. *9*, 7—11 (1895) ref. in Z. Krystallogr. (Leipzig) *30*, 309—311, 1899.
Wulff, G., Zur Frage der Geschwindigkeit des Wachstums und der Auflösung der Krystallflächen, Z. Krystallogr. u. Mineral., Leipzig *34*, 449—530, 1901.
Wulff, L., Krystallisation in Bewegung, Z. Kristallographie, Leipzig, *11*, 120—132, 1886.
Ziz, N. N., Ueber die Einwirkung der atmosphärischen Luft auf die Erzeugung von Krystallen in Auflösungen der schwefelsauren Salze und über verschiedene Arten dieser Krystalle, (in: Ueber den Einfluß des Luftdruckes auf Krystallisation der Salze.) J. Chem. Phys. von J. S. C. Schweigger (Schweiggers J.) Nürnberg, *15*, 157—171, 1815.

Acta Physica Hungarica, 57 (3—4) pp. 179—184 (1985)

SINGLE CRYSTAL RESEARCH — CRYSTAL QUALITY CONTROL — INDUSTRIAL PRODUCTION

R. VOSZKA

Research Laboratory for Crystal Physics
Hungarian Academy of Sciences
1112 Budapest, Hungary

A summary is given of the tasks involved in improving crystal quality, their interrelations and the organizational framework required for completing them. The possibilities inherent in single crystal research for the development of industrial products are demonstrated.

1. Introduction

The basis of experimental solid state physics is the properly characterized solid state sample. If this sample can be obtained in single crystal form the available scientific information is especially abundant. The story of single crystal production is one of intensive and sophisticated research work carried out to obtain a steady improvement in crystal quality. This result required and still requires the exploration of the natural laws valid in the interdisciplinary field of crystal growth itself, however, the special needs of experimental solid state research, and crystal based device development have also to be met simultaneously. In the case of successful device development the problem of the mass production of crystals of proper size, and the introduction of the growth technology into the industrial production process must be solved.

The useful properties of crystals for industrial practice are due either to their lattice or their defect structures (for instance the dopant structure). In both cases it is extremely important as a first step to produce crystals, which are structurally nearly perfect and as pure as possible, since the knowledge of their properties is indispensable as a reference for the investigation of the physical properties of the defective structures. Depending on the intended application, the term "good quality crystal" may refer to low defect concentration as well as to a properly defined defect structure that should be produced already in the growth process; for instance, by appropriate doping. Quality improvement requires the coordination of a series of procedures. It is our aim to describe these activities for the case of the Czochralski growth technologies though the ideas may be adapted to other methods.

2. Discussion

Let us consider block 1 in Fig. 1 representing *single crystal research.* This block includes those four tasks which have to be completed for successful single crystal research work. Let us discuss these tasks in some detail.

a) Basic material production. The activities for the production of the basic material consist of the purification and the synthesis of the starting material, the

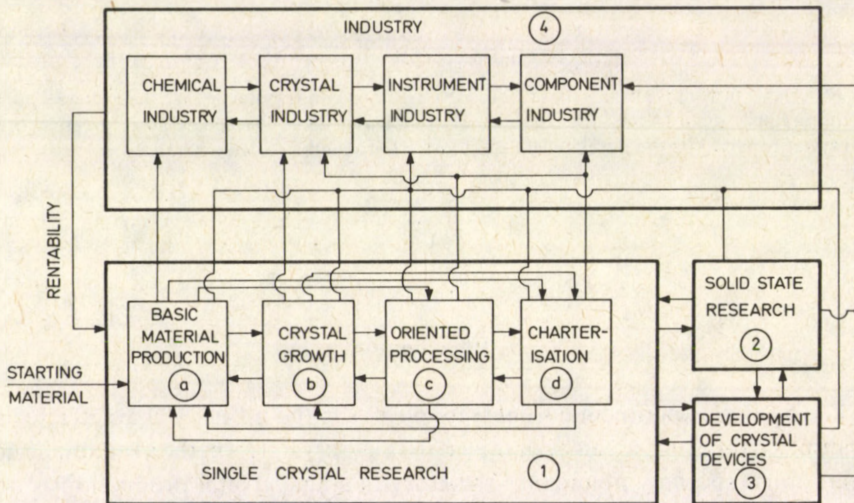


Fig. 1. Block diagram demonstrating the interrelations of single crystal research, solid state research, device development and industry.

investigation of the efficiency of the purifying and synthesizing technologies, as well as the optimization of these technologies.

Further on this work includes also the investigation of the interactions between the melt and the crucible, the study of the chemical interaction of the various impurities within the melt, the measurement of the distribution coefficients of these impurities, the study of stoichiometric problems, the chemical analysis of the as-grown crystals, the development of the various measuring devices required in these activities, etc.

b) The task of *crystal growth* includes the selection of the appropriate atmosphere and crucible, the shaping of a suitable heat space, e.g. by some well defined geometry, temperature isolation, afterheating system or temperature screening shields and also the optimization of the growth parameters such as the pulling rate, the number of revolutions per unit time, the crucible position, etc. The production of good seed crystal generations, and the investigation of the role of the crystallographic orientation

of the seed in crystal growth are also an important part of the work of the team engaged in growth research. A further task is presented by the study of the transport processes, the flow conditions, the growth kinetics, the shape of the phase boundary and the shoulder as well as the development of growth equipment and, especially, the automation of the equipment.

c) The activities of the team engaged in *crystal orientation and processing* consist of increasing and maintaining the exactness of orientation during the cutting processes, developing heat-shock-free fixation technique, optimizing the cutting parameters, developing the grinding and polishing technologies and the devices required for improved orientation and processing, etc.

d) Crystal characterization includes various research programmes.

A *basic characterization* is required in every case if single crystals of some substance have been grown for the first time. This type of characterization involves collecting basic data of the crystal such as, for instance, its chemical composition, molecular weight, crystallographic point and space groups, lattice constants, density, melting point, eventually the Curie point or temperatures of other phase transformations, the crystal hardness, water solubility, refractive index, coefficient of heat expansion, etc.

The most important aspect of crystal characterization refers *to the defect structure* which includes the concentration and distribution of impurity atoms, the structure of aggregations, the dislocation density and distribution, the stoichiometric structure, the mosaic and eventual domain structure as well as the stress structure.

In the case of *application oriented characterization* of crystals the determination of the physical properties directly involved in the given application is required. These are, for instance, the acousto-optical, electro-optical, piezo-electric, pyroelectric, dielectric, photoelectric, photoelastic, and magnetic properties, the electrical conductivity and the eventual luminescence parameters of the crystals.

It may occur that these parameters are influenced by some crystalline defect structure and faulty products are obtained. It is useful to investigate correlations of this type in detail and use the results for sorting out at an early stage the crystal samples which would result in faulty products, saving thus the expenses of processing.

Surface characterization. In most applications processed crystals are used. Consequently — especially for optical purposes — the smoothness of the surface, the thickness and structure of the eroded layer, its planeparallelity and, in the case of cleaved crystals, the surface structure are of very great importance.

Although it is not my intention to describe here the great variety of characterization methods, it should be stressed that the development of special characterizing devices may also be part of the activities of quality research.

It seems to be useful to distribute the discussed *four tasks in four organizational units.* We have done this in our laboratory and organized in fact four functional units, viz. chemical, crystal growth, orientation and processing, and characterizing subgroups which together constitute the crystal technology group.

Crystal quality can be improved by realizing close cooperation with a constant flow of information between all of these subgroups. Usually one has to intervene at a previous level, according to the results of the characterization mostly by a more thorough purification of the basic substances — though for a given purity the growth parameters and the quality of processing also have to be optimalized as a rule. In this chain-like team work the human nature of the individuals and the professional composition of the team play an important role since for success close cooperation of chemists, physicists, engineers and technicians is required.

There may not be complete agreement that the preparation of the basic substance should also be included in this single crystal technological chain. Of course one may start with substances produced by the chemical industry which today offers products of a purity as high as 1—10 ppm. However even this purity level, determined by factors outside the laboratory, has quite frequently proved unsatisfactory, especially in the field of solid state research. Sometimes substances are purified by industry to a very low concentration of certain impurities and on the other hand they still contain considerably high concentrations of others. Thus for instance the superpure alkali halides are extremely pure with respect to impurity cations, whereas for anion impurities they do not attain even the analytical purity level. The growth of extremely pure alkali halides [1, 2, 3] and also of TeO_2 [4, 5, 6, 7] has been successful in our laboratory only because we have developed a purifying technology of our own.

Let us now consider *experimental solid state* research (Fig. 1, block 2). One practical task of single crystal research consists of supplying this field with single crystals for the purpose of investigating new phenomena and new relations applying various methods. Thus the purpose in this case is not the characterization, though whatever research method is applied the results always contain information on the crystal quality as well. Consequently, there is a very close connection between the crystal and the solid state research and a bidirectional flow of information must be secured. We have a group in our laboratory in possession of certain methods for carrying out studies of solid state physics. A large number of additional research methods are offered in national and international cooperation. The possibilities given by such cooperation are fully used by our laboratory which has an extensive network of contracts enabling broad research schemes. In the course of solid state research the possibility is given to develop new devices working in accordance with the phenomena and natural laws detected. This kind of work has been quite successful in our laboratory. For instance we suggested — in cooperation with partners — the development of a device measuring the duration of laser pulses in the ps time range [8, 9]. This device has, in the meantime, actually been realized.

The next block (3) in Fig. 1 refers to the development of single crystal based devices. The limited size of our laboratory means that it has no such unit though we have a number of cooperation agreements with institutions where devices from crystals produced by our laboratory are being developed. Thus for instance TeO_2 based modulators and deflectors were developed by the Research Institute of Automation,

and $LiNbO_3$ based acoustic surface wave filters have been developed by the Institute of Technical Physics and by the Industrial Research Institute of Telecommunication. Similarly, $LiNbO_3$ based vibration detectors were developed in the Laboratory for Acoustics, and the Research Department of Medicor Works has produced $ZnWO_4$ based detectors for computer tomography.

When single crystal research and device development are institutionally separated one has to take considerable care of adequate cooperation since the technological steps of device production may change the defect structure of the crystal. Therefore in the case of faulty production it must be jointly investigated whether the crystal was originally defective or the defect was introduced into the crystal only in the course of device production. Thus the flow of information between groups engaged in single crystal research and device development is of outstanding importance. It is equally important to find those characterization methods whose application ensures a low level of faulty device production. Instruments enabling the control of the quality of new, ready made devices ought to be developed as well. Device development and production leads us to the last block which is *industry* (Fig. 1 block 4).

Single crystal based devices are developed in order to introduce their *industrial production*. The single crystals required for this purpose must be produced by the crystal industry out of basic materials produced by the chemical industry using at every step the technologies worked out in the research laboratory. Similarly, the processing technologies should be industrialized by the crystal or component industry. At the same time the research and development may be carried out in every functional phase at such a level that the devices developed can be produced by the instrument industry.

When industry is expected to start producing a research result, the first question to arise is rentability, this is followed by marketing and other problems. The first question should always be fed back to the research. Production oriented research must be separated at an early stage from other research. I should like to point out here that a crystal required for a certain product need not necessarily be the purest possible substance. If research has already revealed which of the defects are harmful from the point of view of the given application the material must be purified or these particular defects removed, and the rest may remain incorporated in the crystal.

In this way cheaper basic materials for special purposes can be produced. The choice of crystal size has also to be made by rentability considerations, though the increase in crystal size without quality deterioration is in itself an important and intriguing problem for the scientist.

In this respect, our laboratory had already obtained some results in this field in 1960 by the development of the technology of industrial NaI(Tl) single crystal scintillators. The production has been taken over by Gamma Works, Budapest, where 2 tons per year are now produced [10].

An optimal processing technology has to be worked out by minimizing the quantities of processing materials and of waste; in addition, economically optimal quality control methods must be found.

The rentability of prospective production must be kept in mind during the development of instruments and devices as well. Our laboratory has developed a simple and cheap automatic Czochralski growth apparatus where the crystal shape is determined by a phantom [11].

3. Conclusion

From what has been said it can be concluded that single crystal research should be carried out preferably in a complete materials science chain including preparation of the basic substance, crystal growth, oriented processing and crystal characterization, whereby close interaction among the teams can be ensured. Organization of this sort enables the continuous improvement of crystal quality, which is of advantage for solid state research as well as for the development of crystal based devices. For any unit in the materials science chain there is the possibility to develop devices required for the given unit. The technologies and devices developed in the basic functional units may be taken over by industry. The rentability problems posed by industry must be considered already at research level. Our own laboratory has been organized according to these functional principles and I am of the view that in some fields we can already report good results.

The support of the State Office for Technical Development (OMFB) is acknowledged.

References

1. R. Voszka, I. Tarján, L. Berkes and I. Krajsovszky, Kristall u. Technik, *1*, 423, 1966.
2. K. Raksányi and R. Voszka, Kristall u. Technik, *4*, 227, 1969.
3. R. Voszka, K. Raksányi and L. Berkes, Kristall u. Technik, *5*, 409, 1969.
4. I. M. Silvestrova, Yu. V. Pisarevskii, I. Földvári, Á. Péter, R. Voszka and J. Janszky, Phys. Stat. Sol., (a) *66*, K55, 1981.
5. I. Földváry, K. Raksányi, R. Voszka, E. Hartmann and Á. Péter, J. Cryst. Growth, *52*, 561, 1981.
6. J. Janszky, Á. Péter, A. Mecseki, R. Voszka, I. Tarján, I. Földvári, I. M. Silvestrova and Yu. V. Pisarevskii, Kristallografia, *27*, 152, 1982.
7. I. Földvári, R. Voszka and Á. Péter, J. Cryst. Growth, *59*, 651, 1982.
8. J. Janszky, G. Corradi and R. N. Gyuzalian, Optics Commun., *23*, 293, 1977.
9. J. Janszky, G. Corradi and R. N. Gyuzalian, Optica Acta, *29*, 51, 1982.
10. R. Voszka, Acta Phys. Hung., *47*, 5, 1979.
11. F. Schmidt and R. Voszka, Crystal Research and Technology, *16*, k127, 1981.

Acta Physica Hungarica, 57 (3—4) pp. 185—188 (1985)

PHANTOM CONTROLLED AUTOMATIC CZOCHRALSKI PULLER SENSING THE CRYSTAL WEIGHT

B. MEZŐ, L. BÁLINT and R. VOSZKA

Research Laboratory for Crystal Physics
Hungarian Academy of Sciences
1112 Budapest, Hungary

Automatic Czochralski crystal growth equipment is described. The crystals are grown by a phantom controlled system measuring the weight of the growing crystal. The equipment has been used with good results to produce $LiNbO_3$ single crystals with diameters of 5.0 cm.

1. Introduction

Automatic Czochralski crystal growth equipment operating by weight measurement is usually realized by one of two possible constructions. Either the weight change of the melt or the weight change of the pulled solid phase is measured. In both cases automation consists of a regulated change of weight fed back into the system. The regulation is carried out by a program which determines jointly with the heat distribution the shape of the crystal finally obtained.

A phantom controlled automatic Czochralski system detecting the weight change of the melt has already been developed [1] and used with good results to produce $LiNbO_3$, TeO_2, $Bi_4Ge_3O_{12}$, $ZnWO_4$ and other single crystals; these can be grown from the melt in air in resistance heated furnaces.

Here we wish to report on a phantom controlled method by which the crystal growth is regulated by sensing the weight change of the growing crystal.

2. Design and construction of apparatus

Figure 1 presents the schematic design. The console (Co) of the pulling spindle (S) driven by motor M_1 holds a two-beam balance (B) developed by us. The weight program-generator consisting of a vessel (V) containing glycerine with the immersed phantom (Ph) and also the tare (Ta) are mounted on one balance beam; the other holds the pulling rod with the seed holder and the crystal (C) rotated by motor M_2. The balance is connected to a converter generating a voltage (U_G) proportional to the resultant weight (G). During pulling the phantom is more and more immersed in the liquid and pulls down the program-generating side of the balance. The resulting

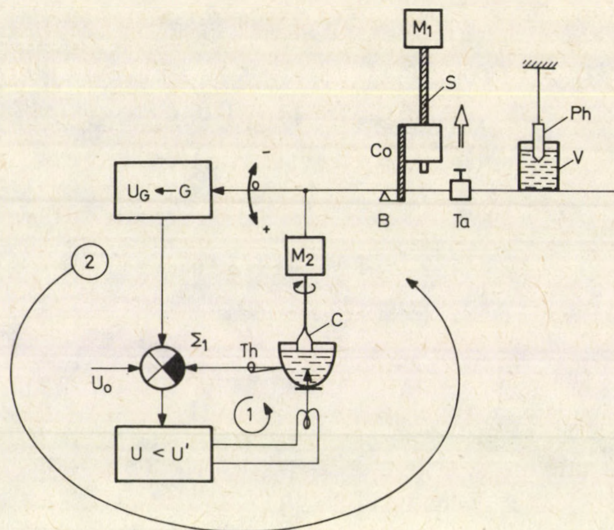


Fig. 1. Schematic diagram of the apparatus. M_1 pulling motor, S spindle, Co console, B balance, Ph phantom, V vessel, Ta tare, M_2 rotating motor, C crystal, $G \rightarrow U_G$ weight-voltage converter, Σ_1 summarizer, U_0 voltage level, Th thermocouple

regulating signal obtained from the converter relays a cooling program to the temperature regulator. This process results in an increase of the crystal diameter and the reestablishment of equilibrium. A continuous weight increase on the program-generator side continually increases the weight of the crystal while the balance remains near equilibrium during the growth process.

The control system is essentially a double-loop regulator (Fig. 2). The first loop represents a temperature program regulator whose program is generated by the second loop. Our system differs from the usual constructions by the summarizer Σ_2, i.e. the balance itself, which summarizes the weights. Figure 3 depicts the balance part of the automatic puller.

3. Experience

The system based on the measurement of the crystal weight compared with systems measuring the weight of the melt has some advantages, but also some disadvantages.

Disadvantages:

a) the system requires light weight and vibration-free motors;
b) the outlets of the motor and the eventual measuring circuits must be mounted momentum free;

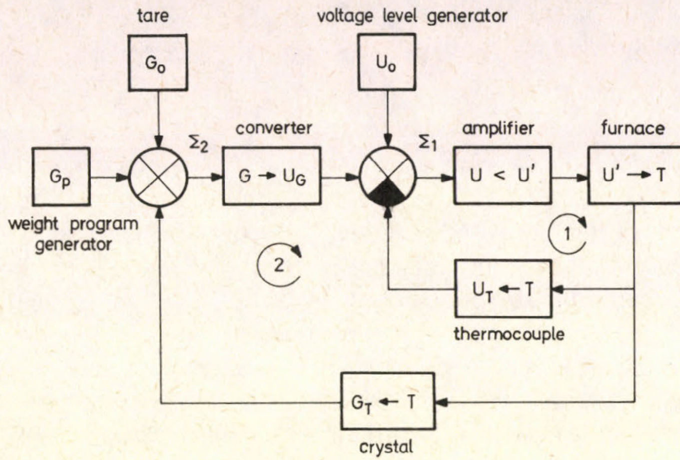


Fig. 2. Schematic diagram of operation during pulling. U_0 voltage level, U voltage, U′ power supplying voltage, T temperature, U_T thermovoltage, Σ_1 summarizer of the temperature programming regulator, G_T crystal weight, G_0 tare weight, G_p weight of the weight program generator, G resultant weight, U_G voltage proportional to the resultant weight, Σ_2 weight summarizer: in our case this is the balance itself

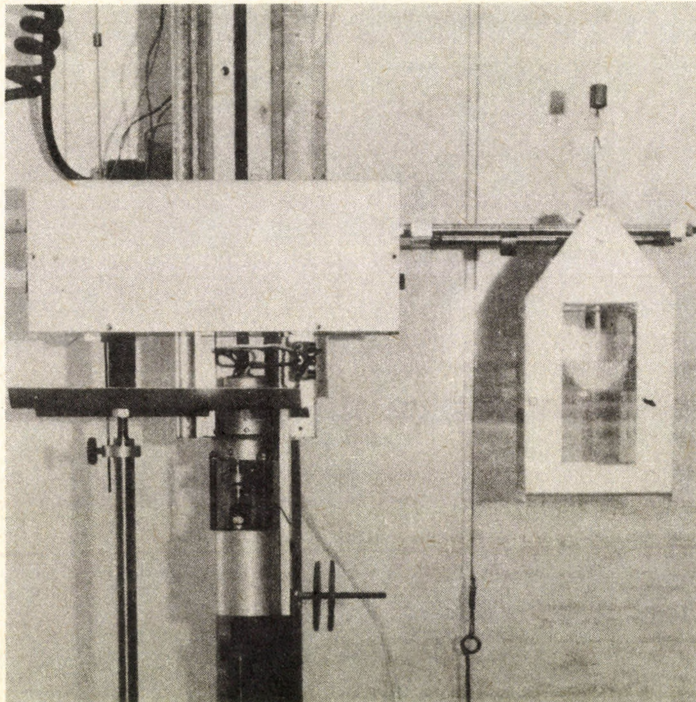


Fig. 3. Illustration of the balance part of the construction

c) the spindle loading must be eliminated by compensating the weight of the balance;
d) increased care must be taken to avoid touching the balance.

Advantages:

a) since the weight of the growing crystal is measured, the melt-evaporation does not interfere with the process;
b) the bottom of the furnace is closed, consequently a suitable temperature distribution and a better efficiency can be attained;
c) the crucible can easily be rotated;
d) the vertical position of the crucible is readily adjustable;
e) the crucible and the heating elements can be protected without technical difficulties;
f) with a single phantom, crystals of various diameters may be grown simply by displacing the weight program generator on its balance beam.

Utilizing this system we were able to grow whole series of $LiNbO_3$ crystals of a mass of 600 g and a diameter of 5.0 cm. The stability of the system, its insensitivity to fast changes and the operation of the simple electronics were found to be quite reliable.

Acknowledgement

This work was supported by the State Office for Technical Development (OMFB). The authors wish to thank the members of the crystal growth team of the Laboratory for valuable discussions on their experiences with the growth system.

Reference

1. F. Schmidt and R. Voszka, Crystal Research and Technology, *16*, K127, 1981.

Acta Physica Hungarica, 57 (3—4) pp. 189—203 (1985)

FINITE ELEMENT ANALYSIS OF PROCESS CONTROL AND OPERATING LIMITS IN EDGE-DEFINED FILM-FED GROWTH (EFG) OF SILICON AND SAPPHIRE RIBBONS: A REVIEW

J. P. KALEJS

Mobil Solar Energy Corporation
16 Hickory Drive
Waltham, Massachusetts 02254, U.S.A.

H. M. ETTOUNEY and R. A. BROWN

Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

Numerical solutions are described of the coupled heat transfer and capillary equations representing growth of silicon and Al_2O_3 ribbons by the EFG technique. The solutions allow study on a quantitative level of the relationship between process control parameters on the one hand, and operating limits and interface shapes on the other. Silicon and Al_2O_3 ribbon growth are contrasted and the role of thermophysical properties of these materials in setting operation regimes for achieving growth in practice is examined.

1. Introduction

The primary variables which must be controlled during the growth process for achievement of predetermined ribbon properties are its thickness and the height and shape of the solidification interface. These variables are in turn influenced by growth speed, by die geometry and by the temperature fields of the ribbon, meniscus and die environments, which are fixed by elements of the growth system design. A complete description of the process is obtained when the relationship between the environment temperature field and ribbon properties is known for a given growth speed and die geometry.

The most extensively studied Edge-defined Film-fed Growth systems, both experimentally and from a theoretical standpoint, have been those for silicon and Al_2O_3 ribbons [1]. Analytical solutions to the capillary equation for meniscus shape and for simplified heat transfer models have provided the information needed for understanding of system stability [2, 3] and for predicting its dynamic response [4]. On account of approximations made, which treat heat transport in one dimension, only qualitative trends are established in these analyses. Moreover, the relation of process variables to interface shape and impurity segregation, which control material properties, cannot be obtained. The approximations inherent in a one-dimensional

treatment of heat transfer in silicon ribbon have been examined using numerical solution of more detailed two-dimensional models [5]. Differences in the maximum growth rate and in location of the solid/liquid interface with respect to the die top are found between one- and two-dimensional models.

Initial attempts to calculate the interface shape from heat transfer considerations using numerical analysis did not model boundary conditions representative of real systems below the growth interface in meniscus-controlled ribbon growth [5, 6]. These analyses predict interfaces which are concave toward the melt for a wide range of growth variables. The influence of the die on impurity segregation and on heat and mass transport in silicon ribbon EFG was examined in detail in a later study [7]. Both symmetric and asymmetric growth (through the ribbon thickness) were modeled, and a full two-dimensional treatment of die top heat transfer and melt convection was used. The qualitative relationship between process variables and interface shape for symmetric growth was shown not to be changed with the added consideration of heat transfer in the die top and capillary and meniscus melts. However, die asymmetry caused significant interface inclination and changed impurity segregation behavior. The latter was attributed to convective enhancement of interface impurity transport that overwhelmed normal segregation arising from interface curvature. All of these studies failed to include the interaction between meniscus shape and crystal thickness and thus cannot be used to relate interface behavior to process and system variables in a self-consistent manner.

Only recently have numerical schemes been developed for solving simultaneously the heat transfer and capillary equations [8, 9]. These allow all parameters in meniscus-controlled growth to be included in relating process variables to operating conditions and interface shape. This provides the first opportunity to study at a quantitative level the influence of external temperature fields and process variables, such as growth speed and ambient temperature, on the ribbon thickness, on its temperature distribution and on the melt/solid interface shape. An additional aspect of system response is amenable to examination with comparison of silicon and Al_2O_3 ribbon growth [10]. Differences can be identified in the sensitivity of the growth process caused by changes in process parameters that arise from variations in the thermophysical properties of the two materials.

The model used for EFG ribbon growth is described in the next Section and the differences in the silicon and Al_2O_3 systems are highlighted. Typical results for operating diagrams for relevant system parameters are presented in Section 3. The influence of growth conditions and material properties on interface shape are examined in detail in Section 4, and model application is discussed in Section 5.

2. Model assumptions for ribbon EFG

The model of the EFG ribbon system studied here is shown schematically in Fig. 1. Relevant dimensions for the die and material constants for the silicon and Al_2O_3 systems are given in Table I. Other calculation domain parameters are: $l_1 = 0.005$ cm, $l_2 = 0.15$ cm, $l_3 = 0.0125$ cm, $l_4 = 0.25$ cm, and $l_5 = 5$ cm. All results are for these nominal values unless indicated otherwise.

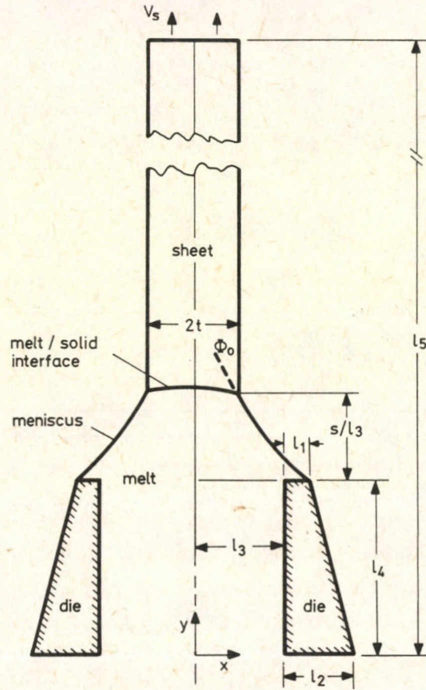


Fig. 1. Cross-sectional schematic of EFG system for ribbon growth

The basic two-dimensional heat transfer and Young—Laplace equations are written using dimensionless groups:

$$\nabla^2 T_i - Pe_i\,(\partial T_i/\partial y) = 0, \qquad Pe_i = V_i \rho_i c_{P_i} L^*/k_i\,, \qquad i = l, s, d\,, \tag{1}$$

$$2H = Bo\,(y + H_{\text{eff}})\,. \tag{2}$$

The subscript i denotes whether the heat balance is for the melt (l), the crystal (s), or the die (d). The mean curvature of the one-dimensional meniscus is $2H$. The physical constants and dimensionless groups used in the analysis are defined in Tables I and II. The Biot and Radiation numbers listed in Table II characterize the importance of

Table I

Thermophysical properties for silicon and Al_2O_3 systems

Property	Si	Al_2O_3
Melt		
Thermal conductivity, k_l [W/cm K]	0.64	0.1
Density, ρ_l [g/cm^3]	2.42	3.05
Specific heat, c_{p_l} [J/g K]	1.00	1.26
Emissivity, ε	0.64	0.9
Crystal		
Thermal conductivity, k_s [W/cm K]	0.22	0.1
Density, ρ_s [g/cm^3]	2.30	4.00
Specific heat, c_{p_s} [J/g K]	1.00	1.26
Emissivity, ε	0.64	0.9
Interfaces		
Melting temperature, $\tilde{T}_m$ [K]	1683	2316
Latent heat, ΔH_f [J/g]	1800	1046
Melt contact angle with die, φ_d	30°	26°
Melt/gas surface tension, γ [dynes/cm]	720	700
Melt/crystal growth angle (orientation)	11°	17° (⟨0001⟩) 35° (⟨10$\bar{1}$0⟩)

Property	Graphite	Molybdenum
Die		
Thermal conductivity, k_d [W/cm K]	0.43	0.84
Density, ρ_d [g/m^3]	2.1	10.2
Specific heat, c_{p_d} [J/g K]	1.7	0.4

convective and radiative heat transfer from the die/crystal/melt system to an environment at the dimensionless temperature T_∞ in relation to conduction across the ribbon. These appear in the heat transfer boundary condition

$$-(\mathbf{n}\cdot\nabla T_i)=Bi_i(T_i-T_\infty)+R_i(T_i^4-T_\infty^4),\qquad i=l,s,d\,, \tag{3}$$

where $\mathbf{n}$ is the unit vector normal to a surface separating the die, melt or crystal from the surroundings. All solutions presented here are for $Bi_i=0$. Additional details of the model are discussed elsewhere [9] and will not be repeated here.

A new finite element algorithm was developed to compute consistent solutions to the two-dimensional equations governing both heat transfer and capillarity in meniscus-controlled growth [8]. The ribbon thickness, the melt/solid and the melt/gas interface shapes are not specified *a priori*, but result from the calculation of the temperature fields in the three regions of interest in the growth system: the melt, the

Table II

Dimensionless groups for modeling Si and Al_2O_3 systems

Dimensionless group	Si	Al_2O_3
Biot numbers, $Bi_i \equiv hL^*/k_i$, $i = l, s, d$	0	0
Radiation number, $R_l \equiv \sigma\varepsilon_i L^* T^{*3}/k_l$	7.2×10^{-4}	7.2×10^{-3}
Melting temperature, $T_m = \tilde{T}_m/T^*$	0.95	1.30
Stefan number, $St \equiv \Delta H_f/c_{p_s} T^*$	1.0	0.47
Bond number, $Bo \equiv gL^{*2}\,\Delta\rho/\gamma$	2.1×10^{-3}	2.6×10^{-3}
Static head, $H_{eff} \equiv h_{eff}/L^*$	20	20
Surrounding temperature, $T_\infty \equiv \tilde{T}_\infty/T^*$	0.2	0.2
Reference temperature, T^*	1783 K	1783 K
Reference length, $L^* = 2l_3$	0.025 cm	0.025 cm

crystal and the EFG die. The finite-element formulation is described in detail in [8] and the strategy for iteration between interface shape and the temperature field calculations is presented in [9].

We model the growth of Si and Al_2O_3 sheets with thickness $2\tilde{t}$ that are continuously pulled and solidified from a die at the velocity V_s. The die geometry used is the same for both systems and is shown in Fig. 1. The die is made of graphite for Si growth and of molybdenum for the Al_2O_3 system.

The heat transfer models used here and in [9] assume that conduction is the dominant mode of heat transfer in both melt and crystal and account for convection simply by incorporating uniform velocities in the melt, V_l, and crystal, V_s, in the growth direction. Heat is transported through the die only by conduction. The transparency of the Si and Al_2O_3 is initially neglected. Radiation through solid Al_2O_3 is a more significant mechanism for heat transfer than for Si [11]. It is estimated that up to 80 per cent of the radiation emitted at the Al_2O_3 interface will be transmitted through the solid, as compared to about 20 per cent for silicon. While silicon transparency is confined to infrared frequencies, the transmission through Al_2O_3 is so pervasive that the growth interface can be visually observed during growth. Some uncertainty exists in the reported thermal conductivity of crystalline Al_2O_3, and the value used here may already partially include effects of crystal transparency. The effect of transparency on Al_2O_3 growth is examined separately below.

The higher melting temperature of Al_2O_3 results in a greater portion of heat being transmitted by radiation to the surroundings than for silicon. This is accounted for in our model by the larger dimensionless temperatures in the Al_2O_3 system; compare the melting temperatures in Table II. Radiative heat transfer to the surroundings also represents a larger portion of the heat lost from the melt in the Al_2O_3 system because of the much lower thermal conductivities in this system as compared to silicon; this effect is described quantitatively by the order-of-magnitude increase in R_l for Al_2O_3 over the radiation number for molten Si. The radiation numbers for the

crystal and die for both Si and Al_2O_3 are computed by scaling R_l with the appropriate thermal conductivity ratio. The molybdenum die used in the Al_2O_3 system is more conductive relative to the melt ($k_d/k_l = 8.4$) than the graphite die in the silicon system ($k_d/k_l = 0.67$) and this difference causes the shape of the isotherms near the die top to vary significantly between the two systems.

Although reported surface tensions of Si and Al_2O_3 melts are similar, the thickness of an Al_2O_3 crystal is much more sensitive to meniscus height because of the larger wetting angle φ_0 at the melt/gas/crystal junction. Values of φ_0 are listed in Table I for sapphire. These vary between 17° and 35°, depending on the crystal orientation. Crystal thicknesses $2\tilde{t}$ predicted by solution of the Young—Laplace equation are shown in Fig. 2 as a function of meniscus height, s, and the height of the die top above the level of the melt, h_{eff}, for wetting angles typical of the Al_2O_3 and Si systems.

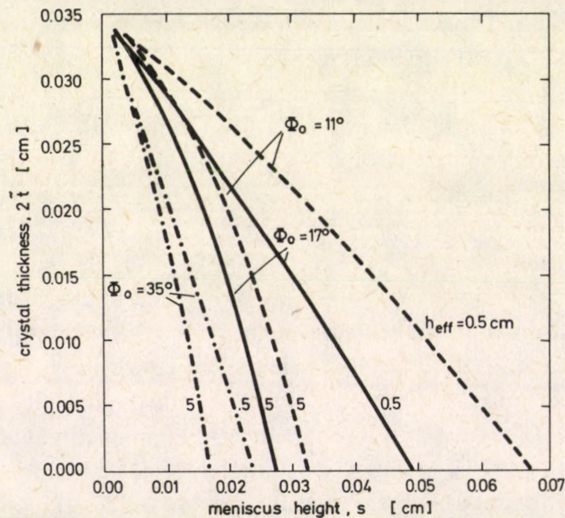


Fig. 2. Relationship between crystal thickness $2\tilde{t}$ and meniscus height s for a range of static heads h_{eff}. Growth angles φ_0 are for Si (11°) and for ⟨0001⟩ (17°) and ⟨10$\bar{1}$0⟩ (35°) grown Al_2O_3

The operating conditions for the EFG systems are characterized by h_{eff}, the pull rate of the sheet, V_s, the ambient temperature of the surroundings, $\tilde{T}_\infty$, and a setpoint temperature, $\tilde{T}_0$, for the melt entering at the bottom of the die. Calculations were performed to set the ranges of these parameters where steady-state operation is possible and to compare for Al_2O_3 and Si the sensitivity of crystal thickness changes to those in the parameters. In the calculations presented here, all other control parameters were held at the values given in Table II, unless changes are explicitly mentioned.

The two temperature variables chosen to parameterize the temperature field can, in principle, be related to system component temperatures. A single uniform ambient

temperature $\tilde{T}_\infty$ is used here for purposes of illustrating the main trends of the model. A variable $\tilde{T}_\infty$ has been incorporated into the calculation scheme, and results show that the basic trends established with uniform $\tilde{T}_\infty$ are modified only at a quantitative level when $\tilde{T}_\infty$ is varied [12]. Thus, the model has the potential for making predictions at a quantitative level with proper accounting for view factors associated with real system components. The die base temperature $\tilde{T}_0$ closely models the temperature setpoint associated with the control thermocouple for monitoring of the die top temperature. Control may be carried out via a feedback loop to the die heater used in the cartridge mode of EFG, or to a main zone (crucible) heater.

3. Operating variable relationships

Control aspects relating to ribbon thickness involving the main parameters of the coupled heat transfer and capillary solutions are illustrated in Figs 3 to 5. The results for silicon and Al_2O_3 contrast the relationship between ribbon thickness and selected operating variables: growth speed, die base temperature and die capillary dimension are parameters, respectively. Questions of stability are not directly addressed through such relationships, but a relative comparison between silicon and Al_2O_3 ribbon growth systems is possible on the basis of operating curve slope magnitudes.

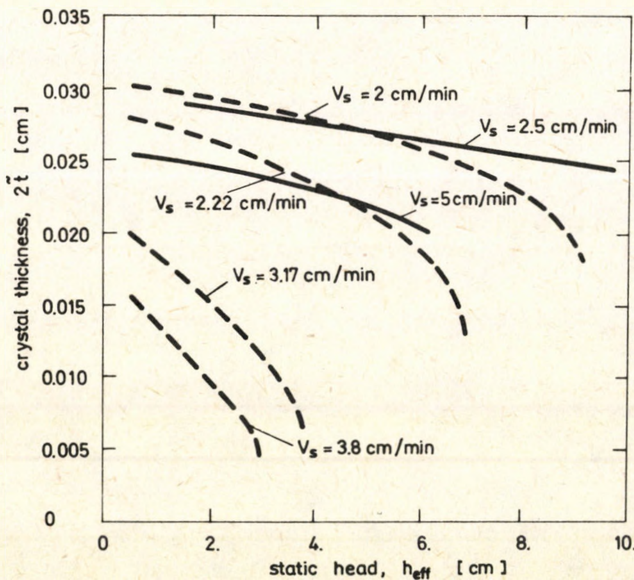


Fig. 3. Relationship between static head and crystal thickness at constant growth speed. $\tilde{T}_0 = 2550$ K and $\varphi_0 = 17°$ for the Al_2O_3 system (solid curves); $\tilde{T}_0 = 1753$ K for the Si system (dashed curves)

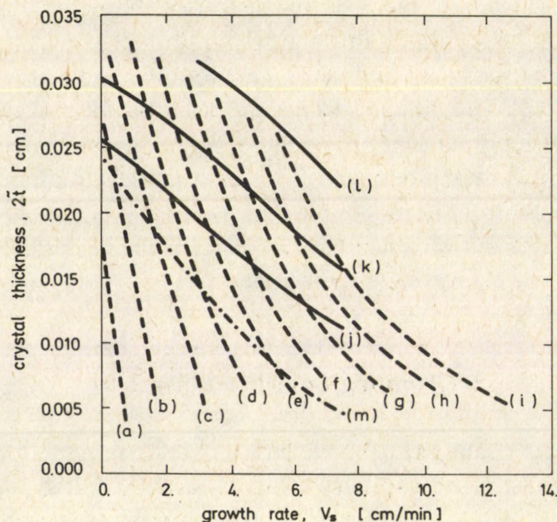


Fig. 4. Operating curves of crystal thickness as a function of growth rate for Si and Al_2O_3 ribbon grown in the systems described in Table II. Curves (a)—(i) are for Si with set-point temperatures $\tilde{T}_0$ varying from (a) 1793 to (i) 1713 K in ten degree increments. Curves (j)—(m) are for Al_2O_3 growth with different combinations of φ_0 and $\tilde{T}_0$; (j) 17°, 2675 K; (k) 17°, 2585 K; (l) 17°, 2496 K; (m) 35°, 2585 K

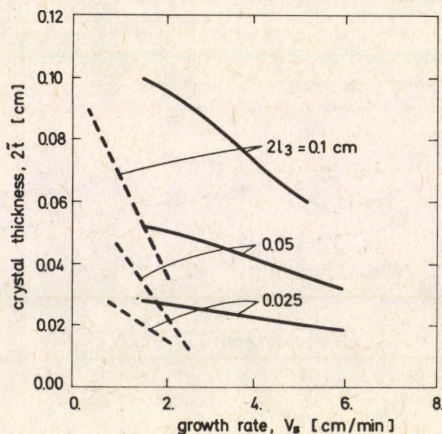


Fig. 5. Operating curves for ribbon grown at three die capillary gaps. Solid curves are for Al_2O_3 with $\tilde{T}_0 = 2585$ K and $\varphi_0 = 17°$; dashed curves are for Si with $\tilde{T}_0 = 1773$ K

The sensitivity of ribbon thickness to a given variable is shown to be markedly different in the two systems. The reduced sensitivity of the Al_2O_3 system, as measured by lower values of

$$\partial t/\partial h_{\mathrm{eff}}|_{V_s}, \quad \partial t/\partial V_s|_{T_0} \quad \text{and} \quad \partial t/\partial V_s|l_1,$$

has been attributed to the dominance of radiative heat transfer, coupled to a lower melt thermal conductivity, in this system [10]. In contrast, conduction acts as the dominant mode of heat transfer in silicon over this range of operating variables. This difference is further reflected in the dependence of the interface melt temperature gradient on growth speed, as is shown in Table III. Limits of stable growth with speed increases are

Table III

Melt temperature gradients at the center of the melt/crystal interface

Material	$\tilde{T}_0$ [K]	$\tilde{T}_\infty$ [K]	V_s [cm/min]	$d\tilde{T}/d\tilde{y}$ [K/cm]
Si	1763	357	1	-5.93×10^2
$2l_3=0.025$ cm	1763	357	2	-5.99×10^2
	1763	357	3	-6.69×10^2
	1763	357	4	-8.08×10^2
Al_2O_3	2585	357	0	-6.90×10^3
$2l_3=0.025$ cm	2585	357	1.5	-6.13×10^3
	2585	357	3.0	-5.35×10^3
Al_2O_3	2585	357	1.5	-2.54×10^3
$2l_3=0.1$ cm	2585	357	3	-2.03×10^3
	2585	357	4.5	-1.41×10^3
	2585	1783	0	-2.60×10^3
	2585	1783	1.5	-1.98×10^3
	2585	1783	3.0	-1.40×10^3
	2585	1783	4.5	-9.69×10^2

closely set by the interface melt temperature gradient, which must remain greater than zero if freezing of the ribbon to the die top is to be avoided. It is evident that the trend predicted for silicon results in growth failure by which the ribbon thickness goes to zero while the melt gradient increases with growth speed. This is contrary to the behavior predicted by one-dimensional heat transfer models [9]. The interface melt gradient dependence on speed for Al_2O_3 follows the reverse behavior.

The gradients in both systems are much higher than usually are realized in practice because thermal stresses caused by growing a crystal under these conditions can lead to cracking of the sheet. Temperature gradients of under 500 K/cm are approached, however, by increasing the ambient temperature to values nearer the melting point, as shown by the calculations for Al_2O_3 with $2l_3=0.1$ cm. Tailoring the ambient heat transfer conditions to an actual growth furnace and proper accounting of radiation through the crystal will lead to realistic temperature gradients but to growth conditions that are more sensitive to h_{eff} and V_s.

4. Interface shape behavior

The present model allows study of details of the dependence of melt/solid interface shape on operating variables not previously examined. In silicon ribbon, experimental data [7] show that impurity segregation exhibits characteristics not predicted by simple models which give interfaces uniformly concave toward the melt. Comparable segregation data are not available for Al_2O_3 ribbon, but this system is amenable to study on another account because the interface is visible due to crystal transparency. The model results for these two systems are examined later in view of the experimental features of interface behavior that need to be accounted for to explain segregation behavior.

Representative melt/solid interface shapes calculated for silicon are shown in Fig. 6. Figure 6(a) shows the effect of varying die setpoint temperature $\tilde{T}_0$; in Fig. 6(b),

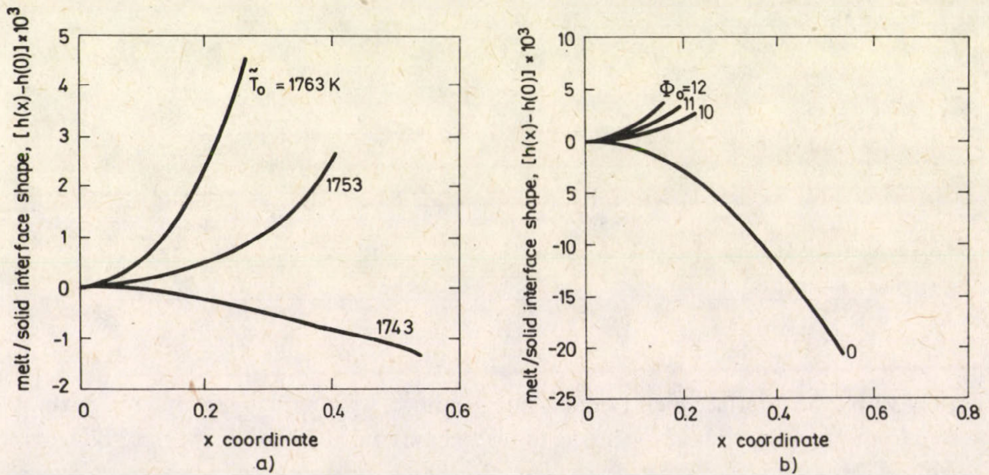


Fig. 6. Variation of melt/solid interface shape for silicon ribbon with (a) set-point temperature at $V=3$ cm/min, and (b) contact angle at $V_s=4$ cm/min and $\tilde{T}_0=1763$ K. The x-coordinate is scaled with $L^*=0.025$ cm

the influence of the contact angle change for the melt/solid/gas interface is shown. The location of each interface at the centerline of the ribbon has been translated to zero; the actual height of the meniscus can be calculated by reading the crystal thickness from this figure and using Fig. 2 for $h_{eff}=0.5$ cm to obtain s.

Interfaces convex toward the melt, obtained when the ribbon thickness decreases by increasing either $\tilde{T}_0$ or φ_0, are not predicted from earlier models [5—7]. This is the case because capillarity (meniscus shape) effects are not included and hence heat transfer in the meniscus region is not accounted for properly [9]. The contact angle for silicon is measured to be 11° in steady-state growth. However, fluctuations about the

steady-state value, such as during crystal thickness transients thus may affect interface shape on account of changes in heat transfer near the solidification front.

The interface shapes predicted for Al_2O_3 ribbon growth have generally more curvature than those found for silicon under comparable environment and growth conditions. Examples of this are shown in Fig. 7(a) and (b). This occurs both because of

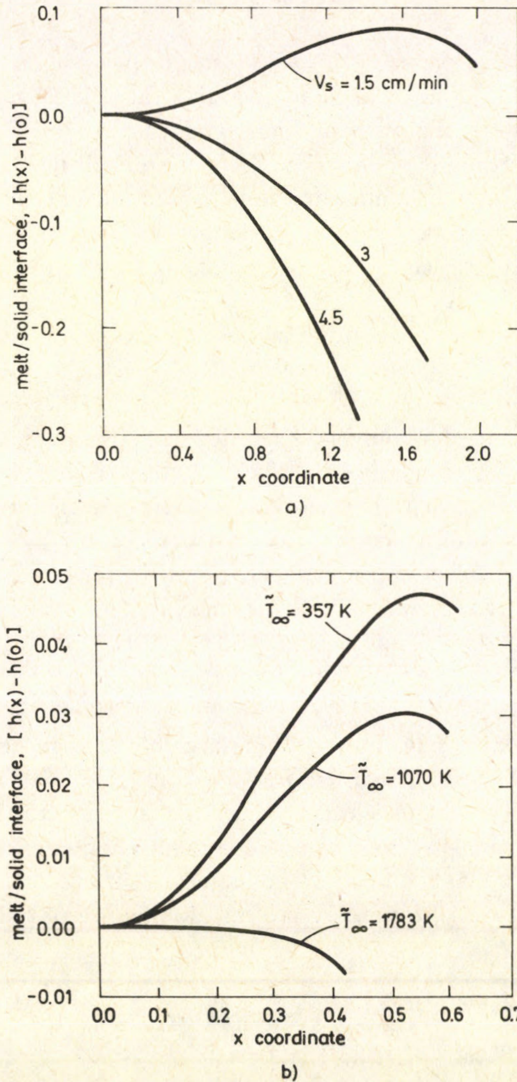


Fig. 7. Variation of melt/solid interface shape for Al_2O_3 ribbon with (a) growth speed at $\tilde{T}_\infty = 357$ K, and (b) environment temperature for $2l_3 = 0.025$ cm. $\tilde{T}_0 = 2585$ K and $\varphi_0 = 17°$ for both sets of results. The x-coordinate is scaled with $L^* = 0.025$ cm

the above mentioned enhancement of radiative heat transfer and of the reduction of conductive transfer in Al_2O_3 in comparison to Si, and also because of differences in die conductivity for these systems. The higher conductivity Mo die in the Al_2O_3 system distorts the shape of the isotherms near the capillary opening more than for Si. This has a particularly strong impact on the interface shape for growth at low meniscus heights.

5. Discussion

The model results presented above illustrate the basic relationships between ribbon thickness and interface shape and operating variables. A number of additional factors associated with real growth systems complicate application of the model at a quantitative level. Among these are: the effects of ambient (gas) heat transfer at the meniscus and ribbon surfaces, approximations in treatment of heat transfer in the solid, direct radiative heat transfer from the interface by light, and modeling of melt convection that becomes inadequate in the limit of high growth speeds. These are examined next.

Ambient convective heat transfer

Radiative heat transport only has been modeled up to this point (see Eq. (3)). Gas convection at the meniscus and ribbon surfaces is known to affect growth conditions. It has been suggested [7] to be responsible for more complicated behavior observed in silicon ribbon segregation than expected on the basis of interface shapes of the form presented in Fig. 6. The effects of ambient convection are included in the present model by introducing non-zero Bi_i in Eq. (3). Significant interface shape changes from those shown arise for non-zero Bi for silicon [5]. It is likely that ambient convection contributes toward changing segregation behavior in the near-surface region of the meniscus in practice.

The Al_2O_3 interfaces exhibit near-surface interface inflections under many growth conditions (Fig. 7). These features are consistent with shapes required to account for formation of certain defect configurations such as sheet microvoids at these locations [10]. They are produced by the dominant cooling effect of radiative heat transfer at the surface. Gas cooling does not influence surface heat transfer to as great an extent in Al_2O_3 as in silicon because of the dominance of radiation cooling.

Solid heat transfer

The values for the solid material constants chosen for the calculations, such as the absorptivity, emissivity and thermal conductivity, significantly influence calculated temperature gradients and profiles in the ribbon. The present results are all for constant

solid parameters. Both silicon and Al_2O_3 crystals are known to transmit significant amounts of radiation through "light piping". In general, the absorptivity and emissivity are strong functions of wavelength (hence temperature), and the lattice thermal conductivity also depends on temperature.

In silicon, free carrier absorption is very high near the melting point making the solid opaque, but this decreases rapidly with increasing wavelength so that the solid is transparent at lower temperatures [13]. Calculations have been done to incorporate these features in modeling heat transport in silicon ribbon [14], and show that direct radiation effects reduce interface gradients below those obtained for the black body model.

Light piping in Al_2O_3

Radiation into the crystal is clearly a more significant mode of heat transfer for Al_2O_3 than for silicon. In absence of accurate data on the lattice conductivity and a model to include light piping effects, we have estimated the influence of light transmission by introducing a heat sink along the melt/crystal interface to simulate direct radiation heat loss. The sink is parameterized by a variable β, which is introduced as a multiplier of a radiative heat loss term in the interface heat balance according to:

$$\mathbf{n}\cdot\nabla T_l-(k_s/k_l)(\mathbf{n}\cdot\nabla T_s)=Pe_l St\,(\mathbf{n}\cdot\mathbf{e}_y)+\beta R_l(T_M^4-T_\infty^4)\,. \tag{4}$$

$\mathbf{n}$ is the interface unit normal projecting into the solid, and $\mathbf{e}_y$ is the unit vector along the growth (y) axis. β is negative here.

The impact of non-zero β is to produce significant changes in interface curvature with respect to the $\beta=0$ solutions, as is shown in Fig. 8. At $\beta=-1.0$, the radiative interface heat loss term of Eq. (4) for $\tilde{T}_\infty=357$ K is of the order of the latent heat term $Pe_l St$ at the higher growth speed (2.5 cm/min), but more than an order of magnitude greater at the lower (0.005 cm/min) speed modeled (Fig. 8). This demonstrates the extent to which direct radiative heat loss influences interface shape as growth speed is varied. The exaggerated interface curvatures produced at the low speed by including direct radiative heat loss from the interface appear to bring the calculated interface profiles qualitatively into closer agreement with those reported experimentally from direct visual observation [15]. However, growth conditions and model parameters are different, and experimental data are not available to warrant a closer comparison at this time.

Melt heat convection approximation

Convective heat transfer by melt flow to the growth interface becomes more important with increasing growth speed. The ratio of convected to conducted heat is measured by the Peclet Number $Pe_l=V_sL^*/\alpha_l$. For Si and Al_2O_3 systems operating at

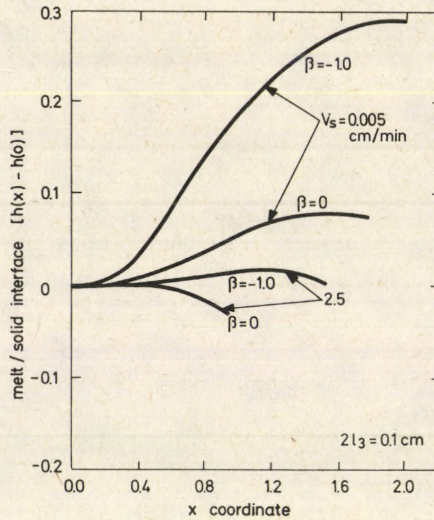


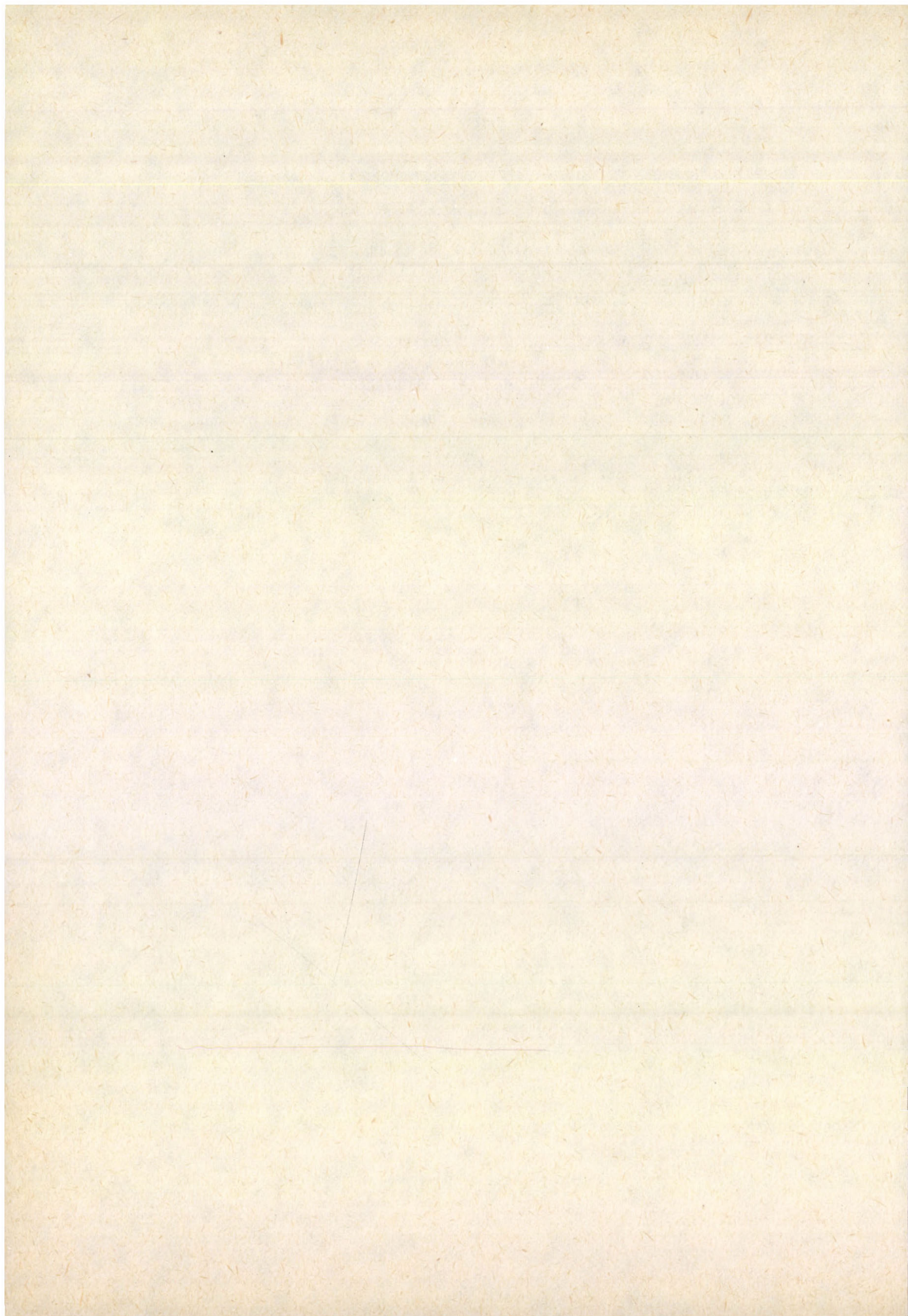
Fig. 8. Effect of direct interface radiation heat loss (see Eq. (4) in text) on interface shape of Al_2O_3 ribbon; $\tilde{T}_0 = 2671$ K and $\varphi_0 = 17°$. The x-coordinate is scaled with $L^* = 0.025$ cm

the same pull rate and die dimensions, comparison of Peclet numbers between Al_2O_3 and Si reduces to only a comparison of thermal diffusivities. The thermal diffusivity of Al_2O_3 melt ($\alpha_l \simeq 0.026$ cm^2/s) is ten times smaller than that of Si ($\alpha_l \simeq 0.264$ cm^2/s), and causes convection to be ten-fold more important in the Al_2O_3 system. As demonstrated in [12], incorporation of a full two-dimensional velocity field can lead to up to a 20% decrease in crystal thickness at $V_s = 5$ cm/min over the value of $2\tilde{t}$ predicted by the uniform flow model used here for Al_2O_3, but has a negligible impact in the case of silicon ribbon growth at this speed. Convective heat transfer will be of greater importance in any systems where fluid flows driven by density or surface-tension gradients are more intense than the growth velocity.

In summary, the model described here for meniscus controlled EFG of silicon and sapphire ribbons is capable of incorporating all relevant physical aspects of the growth process in obtaining solutions of coupled capillary and heat and mass transfer equations. Application of the analysis at a quantitative level requires more detailed knowledge of a number of material parameters of the solid phases of Si and Al_2O_3, and experimental data on system response to process variable changes. Preliminary experimental data have already been obtained in the study of 10 cm wide silicon ribbon growth to test certain aspects of the model [9, 12]. The basic relationships between process variables and ribbon properties discussed here are shown to be reproduced at this level of application.

References

1. A bibliography has been compiled by D. O. Bergin, J. Crystal Growth, *50*, 381, 1980.
2. T. Surek, S. R. Coriell and B. Chalmers, J. Crystal Growth, *50*, 21, 1980.
3. V. A. Tatarchenko and E. A. Brenner, J. Crystal Growth, *50*, 33, 1980.
4. E. M. Sachs and T. Surek, J. Crystal Growth, *50*, 114, 1980.
5. H. M. Ettouney and R. A. Brown, J. Crystal Growth, *58*, 313, 1982.
6. B. Chalmers et al., Mobil Tyco Solar Energy Corporation, Final Technical Report NSF/RANN/SE/GI—37067X/FR/75/1, August 1975.
7. J. P. Kalejs, L.-Y. Chin and F. M. Carlson, J. Crystal Growth, *61*, 473, 1983; J. P. Kalejs, J. Crystal Growth, *61*, 485, 1983.
8. H. M. Ettouney and R. A. Brown, J. Comput. Physics, *48*, 118, 1982.
9. H. M. Ettouney, R. A. Brown and J. P. Kalejs, J. Crystal Growth, *62*, 230, 1983.
10. J. P. Kalejs, H. M. Ettouney and R. A. Brown, J. Crystal Growth, *65*, 316, 1983.
11. P. I. Antonov, S. I. Bakholdin, E. A. Tropp and V. S. Yeferev, J. Crystal Growth, *50*, 62, 1980.
12. J. P. Kalejs et al., "Stress Studies in EFG", Monthly Report, DOE/JPL 956312/83 (August 1983), unpublished.
13. Yu. I. Ukhanov, Soviet Physics-Solid State, *3*, 1529, 1962.
14. J. P. Kalejs et al., "Stress Studies in EFG", Third Quarterly Report, DOE/JPL 956312/83/03 (April 15, 1983), unpublished.
15. R. E. Novak, R. Metzl, A. Dreeben and S. Berkman, J. Crystal Growth, *50*, 143, 1980.



Acta Physica Hungarica, 57 (3—4) pp. 205—211 (1985)

PREPARATION OF ALUMINIUM–SAPPHIRE BICRYSTALS

H.-G. LINDENKREUZ, M. JURISCH

Central Institute of Solid State Physics and Material Sciences
Academy of the Sciences of GDR, Dresden, GDR

and

S. SAHLING

Technical University, Department of Physics, Dresden, GDR

The growth of aluminium–sapphire bicrystals necessary for studies of the thermal boundary resistance between a metal and a dielectric at low temperatures is reported. Some results characterizing the perfection of the interface and the near interface region are discussed.

1. Introduction

The acoustic mismatch theory of the thermal boundary resistance (R_0) between two different materials developed by Khalatnikov [1a] and improved by Little [1b] predicts a theoretical temperature dependence for the low temperature region according to $R_{0\,th}=k_{th}\,T^{-3}$ (k_{th} = constant). The existence of electrons in the metal of a dielectric/metal contact should not contribute to the thermal boundary resistance if the sound velocity in the dielectric is greater than in the metal. Otherwise the electrons lead to a reduction of the thermal boundary resistance, i.e. the thermal boundary resistance of a contact with the metal in the superconducting state should be higher than in the normal state.

Measurements of the thermal boundary resistance of sapphire–indium [2] and sapphire–lead interfaces [3] show that the thermal boundary resistance in the normal state is lower than in the superconducting one. This qualitatively disagrees with the theory because the sound velocity in sapphire is larger than in the metals studied. Further $R_0 > R_{0\,th}$ and $R_0 \sim T^{-n}$ with $n > 3$ was observed for these interfaces. It was argued that the observed discrepancies between theory and experiment are caused by imperfections in the interface and in the near interface region. In order to elucidate the influence of imperfections high quality aluminium–sapphire bicrystals were grown. This communication gives a detailed description of the sample preparation and characterization. The results of the measurements of the thermal boundary resistance were published in detail in [4].

2. Crystal growth

In Fig. 1 the finally used graphite container for a vertical Bridgman arrangement under vacuum or hydrogen is shown. It consists of three parts: a lower one, the dimensions of which agree with the sapphire single crystal except of the part near the sapphire end face, where the diameter is increased from 5 to 5.8 mm, a middle part for

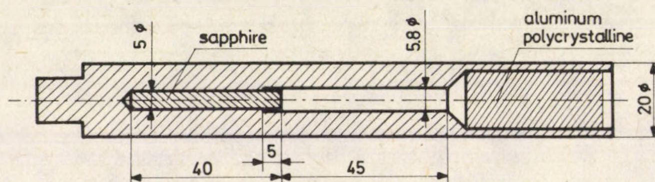


Fig. 1. Graphite container

the directional solidification of aluminium and an upper part for the initial polycrystalline aluminium. Using a two-zone resistance furnace or r.f. heating the whole container is heated slightly (app. 50 K) above melting temperature of aluminium. In this way, an oxide free melt is poured on the cleaned end face of the sapphire, filling the whole middle part of the container. Directional solidification was carried out with a displacement rate of the heater of $5 \cdot 10^{-3}$ mm/s. For growth under hydrogen a small slot parallel to the sapphire in the graphite in connection with a radial hole in the lower part of the container was necessary. Verneuil grown cylindrically shaped sapphire single crystals of 40 mm length with roughly ground surfaces and polished end faces of random orientation were used. The roughness of the end faces was varied: 0.5, 3.8, 8.4 µm. Two qualities of aluminium characterized by the data in Table I were used. An

Table I

Average content of the used aluminium (at. ppm)

	Sb	Ga	Cu	Fe	Mn	Cr	V	Ti	Ca	Si	Mg	r_0^{-1}
High purity	3.5	2.1	0.9	1.6	1.1	1.6	1.3	0.9	0.8	33	6.7	2500...1700
Pure	9.6	18	120	320	840	390	3.1	13	3.8	750	96	3

as grown bicrystal is shown in Fig. 2. Owing to the special pouring technique aluminium adhered tigthly to the graphite container. Usually the bicrystal could be removed by destruction of the container, only. This adhesion was more pronounced for growth under vacuum than under hydrogen. For measurements of the thermal boundary resistance the diameter of the aluminium crystal was reduced to 5 mm (for details see [4]).

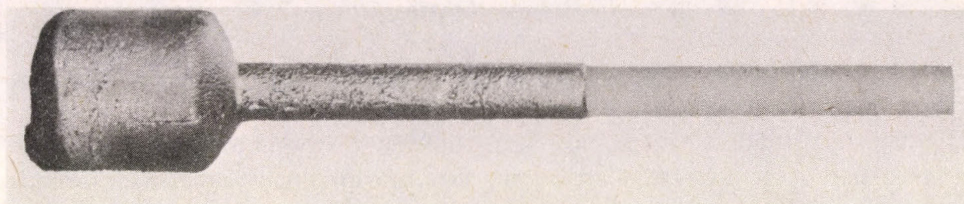


Fig. 2. As grown sapphire–aluminium bicrystal

3. Perfection of the grown bicrystals

3.1. Physical perfection

Looking through the sapphire single crystal an optical inspection of the interface was possible. After optimization of the melting and growth procedure homogeneous highly mirroring contact faces of aluminium without visible oxide particles, areas of insufficient wetting or cracks in the sapphire crystals independent of the roughness of the sapphire end face were found. The contact was extremely tight, attempts to separate the aluminium from the sapphire always resulted in the destruction of the sapphire.

The directionally solidified aluminium was single crystalline without favourable orientation for high purity initial material, but polycrystalline for lower grade material. No distinct orientation relation between the sapphire "seed" and the aluminium single crystals was found.

Cross sections in different distances from the interface and longitudinal sections containing the interface were cut from the crystals by a 50 μm tungsten wire saw, slightly electropolished using a standard procedure and investigated by Berg–Barrett Topography. A typical picture of a longitudinal section is shown in Fig. 3. Similar to Bridgman grown Al-single crystals a pronounced substructure with subgrains slightly elongated parallel to the growth direction was observed. There was no detectable difference between the substructures near and far from the interface.

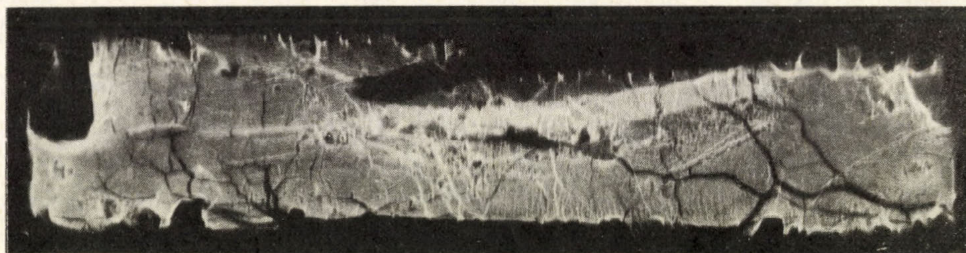


Fig. 3. X-ray topograph of a longitudinal section of the Al component single crystal (mag.: 8 : 1, {311})

3.2. Chemical perfection

The impurity distribution was studied by microautoradiography on ORWO K 106 stripping film after neutron irradiation of longitudinal sections containing the interface. Autoradiograms were taken after different decay times. Together with an investigation of the decay behaviour and γ-spectroscopic analysis the image forming isotopes could be identified.

In Fig. 4a the optical density which is approximately proportional to the average impurity content is shown as a function of the distance from the interface. The copy of

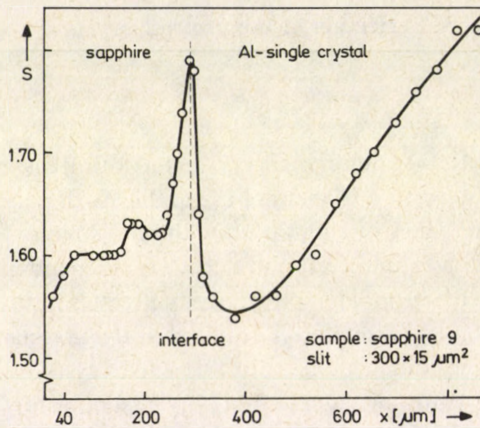


Fig. 4a. Optical density (proportional to average impurity concentration) as a function of the distance from the sapphire–aluminium interface

the corresponding autoradiograms is represented in Fig. 4b. Bright regions correspond to high impurity concentration. Three regions can be seen: the interface region with an enhanced impurity content followed by a minimum of concentration in the aluminium single crystal. After this region the concentration increases up to a nearly constant value. The decay behaviour of the interface region agrees with that of the bulk of the aluminium proving that the impurity enrichment in the interface region has a comparable composition. It is assumed that the impurity rich region is created if the sapphire temperature at the beginning of the pouring procedure is lower than the Al-melting temperature resulting in a high solidification rate of the first melt layer. After this the usual transient segregation takes place leading finally to a nearly constant concentration. Up to the optimization of the growth procedure the region of nearly constant concentration was characterized by an inhomogeneous cellular impurity distribution for both grades of aluminium. An example is shown in Fig. 5. Since cells are

oriented perpendicular to the growth front a curved solidification isotherm must be assumed for the crystal. This is due to a disturbance of the heat flux because of a crack in the sapphire single crystal. The region of cellular solidification could be completely avoided for high purity aluminium using the optimal growth procedure.

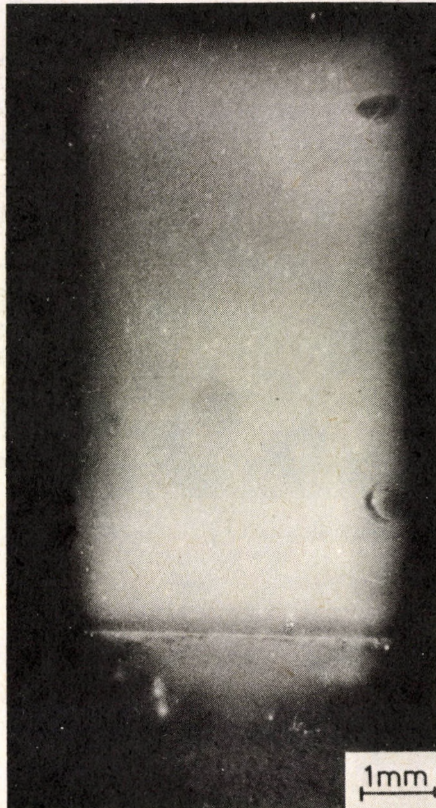


Fig. 4b. Corresponding autoradiograph

4. Thermal boundary resistance

In Figure 6 the thermal boundary resistance of an aluminium sapphire bicrystal is shown as a function of temperature (curve 1a, b) in comparison to the theoretical curve (2) and results from the literature (curves 3, 4). It follows that in the normal state of Al (curve 1a) the T^{-3} dependence is obeyed below 1.5 K in agreement with the theory. However, the measured thermal boundary resistance is smaller than the predicted (explanation see [4]). Comparing our results (curve 1a, b) with those of [2]

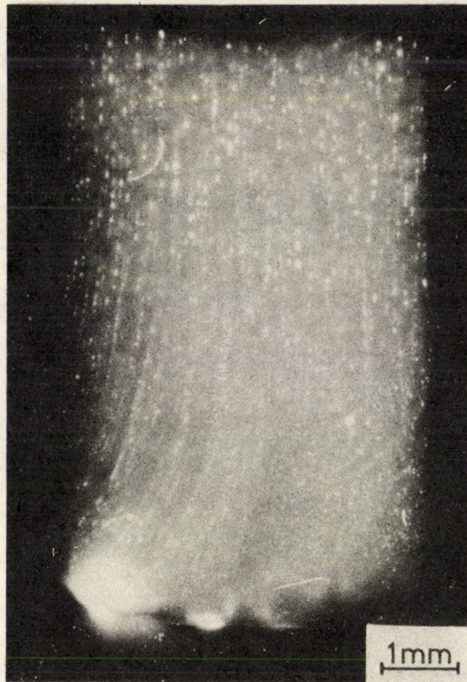


Fig. 5. Cellular impurity distribution near the interface of high purity Al single crystal (Background: Sc—46, impurities: Sb—122/124, Sn—113, Ce—141, Ir—192)

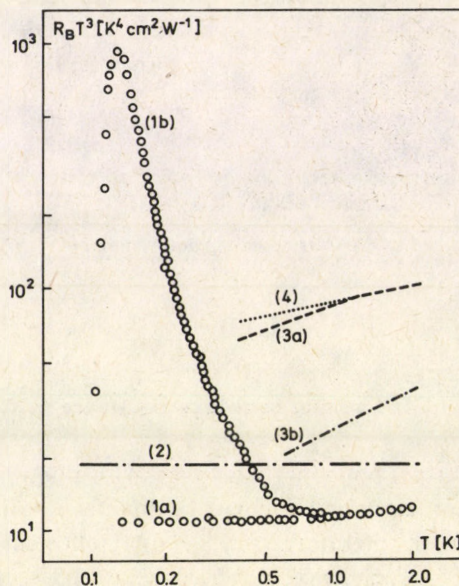


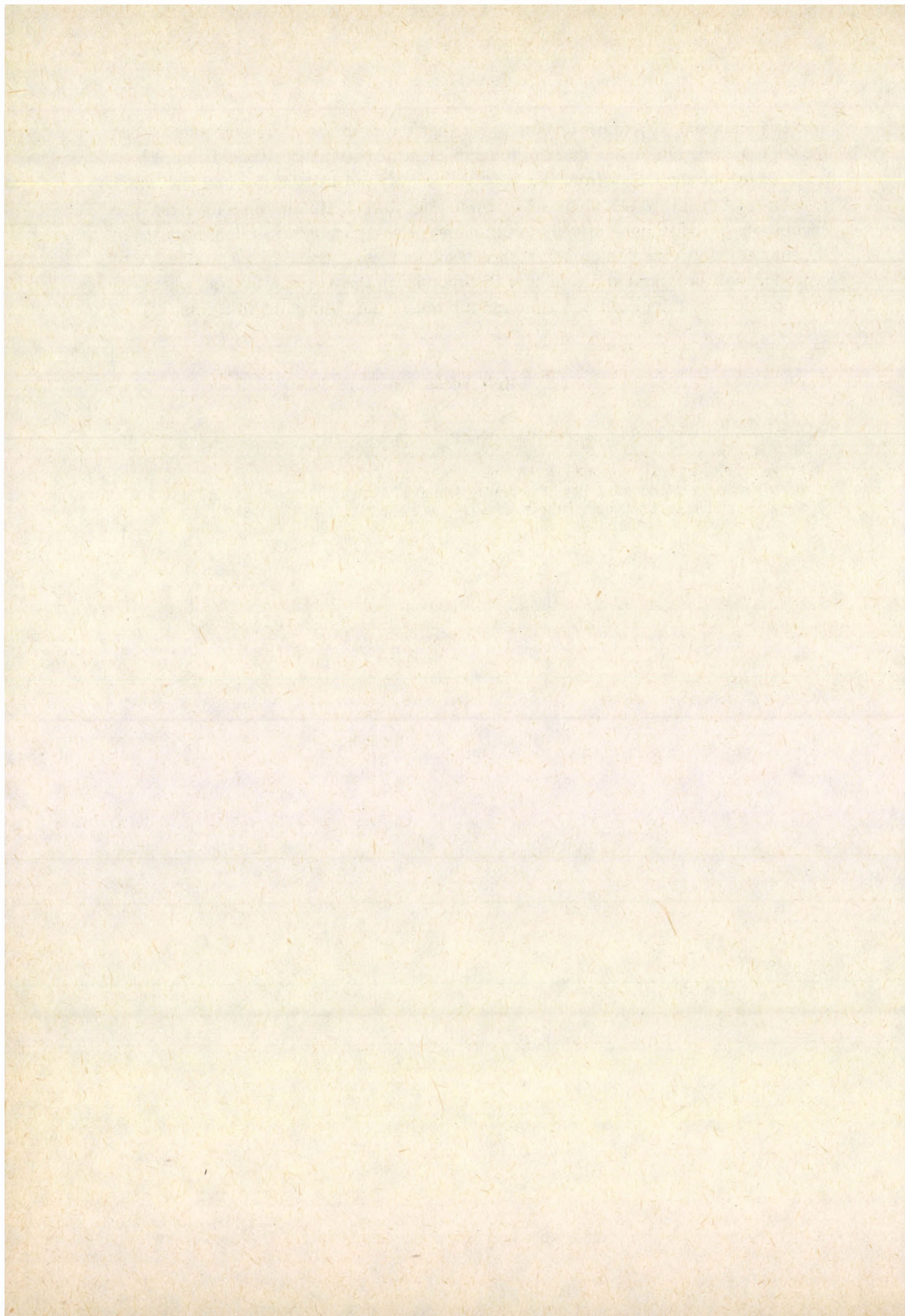
Fig. 6. Thermal boundary resistance for sapphire–metal samples as a function of temperature (1a) $Al_2O_3 - Al$ (normal state), roughness < 0.5 μm; (1b) $Al_2O_3 - Al$ (superconducting state), roughness < 0.5 μm; (2) calculated from the acoustic mismatch theory; (3a, b) $Al_2O_3 - In$, [2]; (4) $Al_2O_3 - Pb$, [3]

and [3] (curves 3, 4) it must be concluded that the reported disagreement between theory and experiment was due to the insufficient quality of the studied samples.

A drastic deviation from the T^{-3} law was found for the thermal boundary resistance of the bicrystals in the superconducting state of aluminium (curve 1a). The effect is observed for high purity aluminium only. A comparable result does not exist in the literature. An explanation for this behaviour was given in [4] taking into consideration an extended region near the interface in the Al-single crystal, where the quasiparticles of the electron system are still not in the thermal equilibrium.

References

1a. U. M. Khalatnikov, Zh. Eksp. Teor. Fiz., *22*, 687, 1952.
1b. W. A. Little, Can. J. Phys., *37*, 334, 1959.
2. M. W. Wolfmeyer, G. T. Fox and J. R. Dillinger, Phys. Lett., *31*, 401, 1970.
3. B. S. Park and Y. Narahara, J. Phys. Soc. Japan, *30*, 760, 1970.
4. S. Sahling, J. Engert, A. Gladun and R. Knöner, J. Low Temp. Phys., *45*, 457, 1981.



Acta Physica Hungarica, 57 (3—4) pp. 213—221 (1985)

PREPARATION OF ZnO SINGLE CRYSTALS BY CHEMICAL TRANSPORT REACTION AND DETERMINATION OF ZnO HOMOGENEITY RANGE

H. OPPERMANN, G. STÖVER and A. HEINRICH

Central Institute for Solid State Physics and Materials Research
Academy of Sciences of the GDR, Dresden, GDR

K. TESKE

Central Nuclear Research Institute of the Academy of Sciences
of the GDR, Rossendorf, GDR

E. ZIEGLER

Physics Section of the Mining Academy Freiberg, GDR

The homogeneity range of ZnO is very narrow and cannot be determined by conventional analytical methods. The determination was performed by oxidation of the substoichiometric samples in a gas stream with coulometric control of the gas-solid phase reaction by the aid of a solid electrolyte cell. ZnO single crystals were prepared by chemical transport. Transport of stoichiometric ZnO is possible with Cl_2, Br_2 and HCl going from the higher to the lower temperature, whereas HBr transports from the lower to the higher temperature. ZnO, which exhibits the lower phase boundary composition, can be transported only by HCl and HBr or their ammonium salts, respectively, from the higher to the lower temperature. The deviation of stoichiometry depends on the preparation temperature, i.e. the deposition temperature during the chemical transport. The value of the deviation amounts to 30—240 ppm in the range 800—1100 °C. Measurements of the electrical conductivity and the Hall mobility yield the same temperature dependence of the phase boundary composition.

Preparation of ZnO

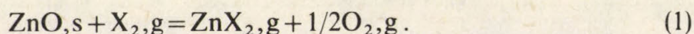
The application of the chemical deposition from the gas phase for preparing single crystals or layers of defined non-stoichiometric phases is one of the trends in chemical ιransport reactions (CTR). In this connection the question concerning the kind and composition of the deposited solid is a very important one. Generally, if a phase with homogeneity range exists, it is possible in CTR to deposit samples with the upper phase boundary and the lower phase boundary too, also in the case of a very small range of homogeneity.

Zinc oxide — as the only compound in the system Zn/O — shows a very small range of homogeneity. According to Hagemark and Toren [1] the phase width is between 1 and 100 ppm as a function of temperature. Within this range of homogeneity the physical properties of ZnO change very much.

The preparation of the investigated zinc oxide single crystals was carried out by chemical transport in a closed system.

In comparison with the deposition in an open system [2, 3] this preparation has the advantage of the possibility to adjust the coexistent pressure of oxygen more exactly, and therefore the properties of the crystals can be obtained as required.

In a closed quartz ampoule ZnO powder and the transport agent X_2 are brought to reaction:



$$ZnO,s + X_2,g = ZnX_2,g + 1/2O_2,g\,. \tag{1}$$

Using the dependence of the chemical reaction or reactions on the temperature of the equilibrium state (Eq. (1) — schematically) — (see Fig. 1) — the ZnO powder is

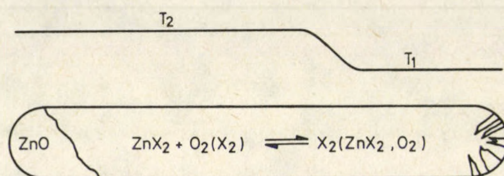


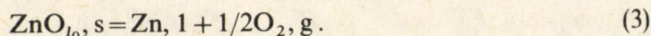
Fig. 1. Scheme of reaction and diffusion during crystal growth of ZnO

dissolved at T_2 and precipitated in crystals at T_1. The reaction products ZnX_2 and O_2 formed at T_2 diffuse to the location of deposition, where the backreaction takes place under formation of ZnO and X_2. The X_2 diffuses back to T_2, reaction takes place and so on.

Br_2, Cl_2, HBr, HCl, NH_4Cl and NH_4Br have been used as the transport agent. The transport behaviour depends on the transport agent, the total pressure and the temperature, or the difference between the temperatures of the location of dissolution and deposition. Moreover, to a high degree the transport behaviour is a function of the coexistent pressure of oxygen above ZnO, e.g. a function of the stoichiometric composition of the deposit at the dissolution part of the ampoule as shown generally as well as by some examples [4, 5, 6]. The coexistent pressure of O_2 above zinc oxide (ZnO_x):

$$[O]_{ZnO} = 1/2O_2,\, g \tag{2}$$

decreases with decreasing O_2-content (decreasing x) from the upper phase boundary ZnO_u to the lower phase boundary ZnO_{l_0}. Within the range of homogeneity ($l_0 < x < u$) the pressure of solubility of O_2 is therefore a function of x and T. At the lower phase boundary the pressure of O_2 is only a function of temperature because of the reaction of decomposition:



$$ZnO_{l_0},\, s = Zn,\, 1 + 1/2O_2,\, g\,. \tag{3}$$

The dependence of the pressure of decomposition of O_2 on the temperature at the phase boundary ZnO_{l_0}/Zn is shown in Fig. 2. This fact permits the unambiguous preparation of ZnO of the upper and the lower phase boundary by chemical transport [7].

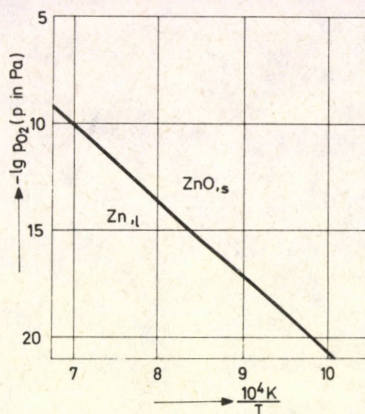
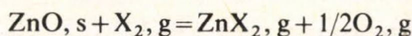


Fig. 2. Decomposition pressure versus temperature

Upper phase boundary

For unambiguous transport of zinc oxide of the upper phase boundary mainly Br_2 and Cl_2 are useful, giving an oxidizing atmosphere by the simple reaction (1):



$$ZnO, s + X_2, g = ZnX_2, g + 1/2O_2, g$$

under which deposition of ZnO rich in oxygen takes place. The transport behaviour can be described by one reaction because the pressure of O_2 for ZnO of the upper phase boundary is given by the ratio $p_{ZnX_2}/p_{O_2} = 2$; this means that ZnO is precipitated stoichiometrically.

The above-mentioned transport agents move from hot to cold (with the exception of HBr). According to thermodynamic expectations Br_2 shows the best transport effect, followed by the ammonium halides; the lowest effect is given by HCl. Optimum transport temperatures are at 1000—900 °C, where partly well-formed rod-shaped crystals with hexagonal cross-section are precipitated (Fig. 3) and partly crystals with a plaingrowing hexagonal area.

Concerning the quality of the crystals Br_2 is also the most suitable transport medium. The crystals of the upper phase boundary are predominantly colourless and clear — some are opaque.

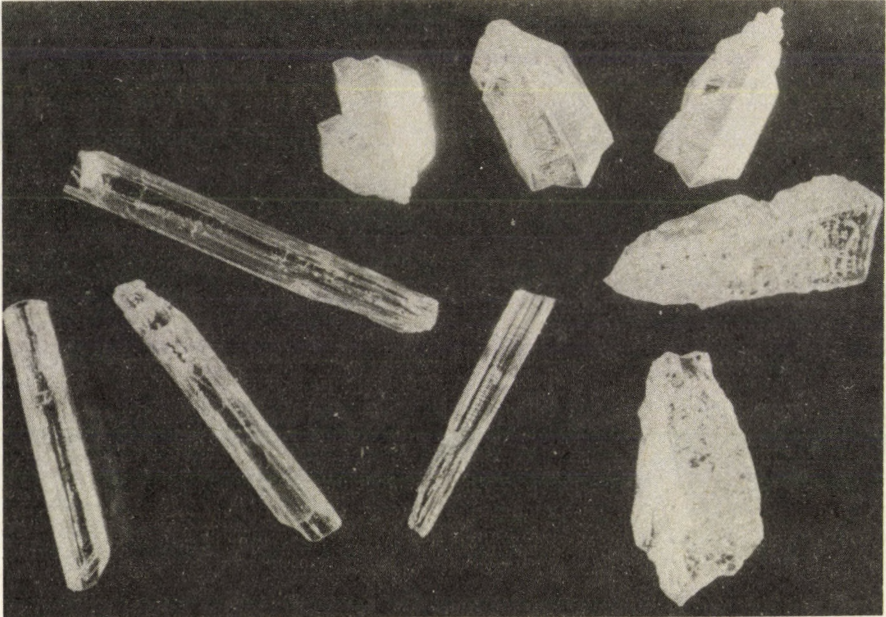
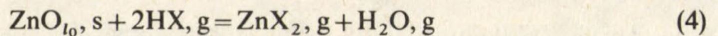


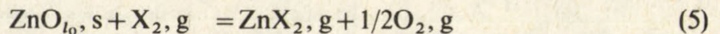
Fig. 3. ZnO crystals (upper phase boundary) by CTR with Br_2

Lower phase boundary

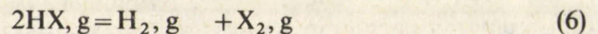
At the lower phase boundary p_{O_2} is fixed by reason of the decomposition equilibrium (3) therefore the transport of ZnO with the lower phase composition is only possible with the transport agents HX and NH_4X, where the mass transport of oxygen is carried out by H_2O. For all transport agents the transport at the lower phase boundary takes place from hot to cold because of the decreased p_{O_2}. Contrary to the upper phase boundary three independent reaction equations are necessary besides the consideration of the decomposition equilibrium (3) to describe the transport:



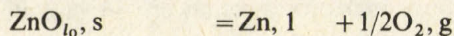
$$ZnO_{l_0}, s + 2HX, g = ZnX_2, g + H_2O, g \quad (4)$$



$$ZnO_{l_0}, s + X_2, g \quad = ZnX_2, g + 1/2O_2, g \quad (5)$$



$$2HX, g = H_2, g \quad + X_2, g \quad (6)$$



$$ZnO_{l_0}, s \qquad = Zn, 1 \quad + 1/2O_2, g$$

To realize the chemical transport for lower phase boundary a small amount of elemental Zn besides ZnO is given into the ampoules, reducing ZnO at the reaction temperature to the lower phase boundary and maintaining the coexistent partial

pressure of oxygen of ZnO_{l_0} at the location of deposition during the transport. If the deposition temperature T_1 is changed from 700 °C to 1100 °C, crystals of the corresponding lower phase boundary ZnO_{l_0} can be obtained even if the phase width is small and analysis of the phase boundary difficult. According to Hagemark and Toren [1], with increasing temperature the phase boundary line is shifted to the side of Zn-excess. Crystals of zinc oxides of the lower phase boundary are orange to brown; the intensity of the colour increases with increasing deposition temperature and therefore increasing deficit of oxygen or excess of zinc. Crystals of the lower phase boundary transported by HBr at 1000—900 °C are shown in Fig. 4. Often the Zn-particles adhere to the crystals, complicating the formation of crystals without defects.

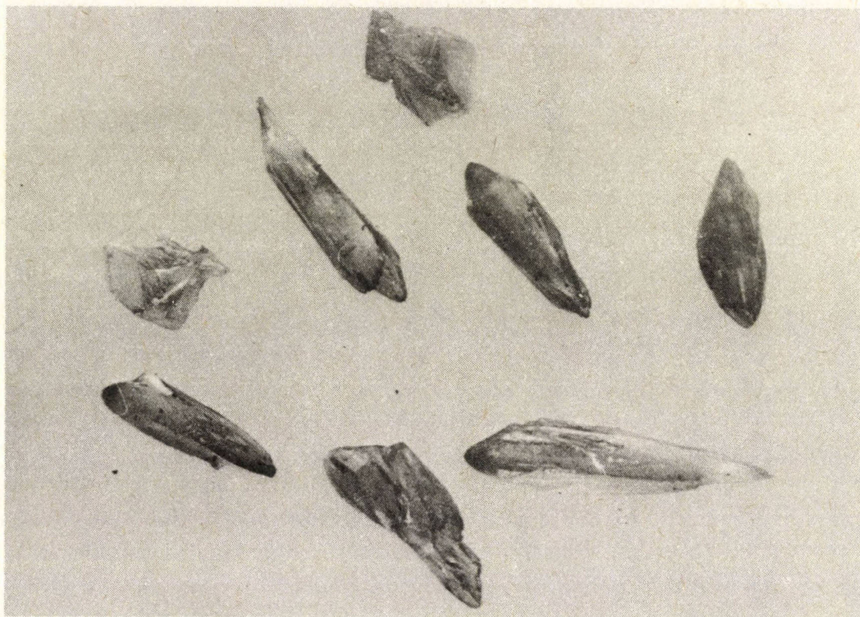


Fig. 4. ZnO crystals (lower phase boundary) by CTR with HBr

Determination of homogeneity range

The homogeneity range of ZnO is very narrow and cannot be determined by conventional analytical methods [8]. The determination was performed by oxidation of the substoichiometric samples in a gas stream with coulometric control of the gas–solid phase reaction using a solid electrolytic cell [9, 10], Fig. 5. An inert gas containing only a few vol.-ppm oxygen was used as the gas phase for the sensitive detection of the deviation from stoichiometry.

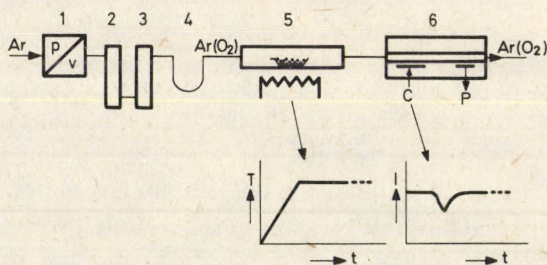


Fig. 5. Scheme of coulometric determination of the deviation from stoichiometry; 1 inductance coil, 2 Ni catalyst, 3 molecular sieve (O_2 elimination), 4 diffusion membrane, 5 quartz boat, 6 solid electrolytic cell

If the ZnO_{1-x}-sample reacts with the gas, we obtain the mass of oxygen via the potentiostatic measured stream I over the time

$$q_{O_2}=k\int_{t_1}^{t_2}(I(t)-I_0)\,dt \qquad k=82.8\ \mu g\ O_2/A\cdot s \tag{7}$$

and from this the deviation from the stoichiometry

$$\Delta x=x=\frac{q_{O_2}\cdot M_1}{(E+q_{O_2})\cdot M_2}, \tag{8}$$

where E denotes the mass of the sample, M_1 the mol mass of ZnO, and M_2 the mol mass of O_2.

If the composition of the sample is equal with the lower phase boundary and the end of the oxidation equal with the stoichiometric oxide the difference is equal with the homogeneity range of ZnO.

Figure 6 demonstrates a typical curve for oxidation, the endpoint of which is indicated by the stream $I(t)$ going back to I_0. This Figure shows two types of reactions: only the second type can be assigned to nonstoichiometry; the first one is due to impurities, especially from the transport agent.

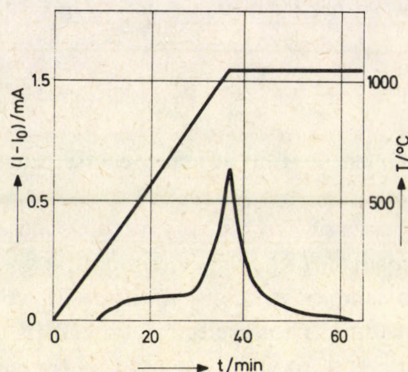


Fig. 6. Typical curve for oxidation

The deviation from stoichiometry depends on the preparation temperature, i.e. on the deposition temperature during the chemical transport reaction. The value of the deviation amounts to 30—240 ppm in the range 800—1100 °C (see Fig. 7).

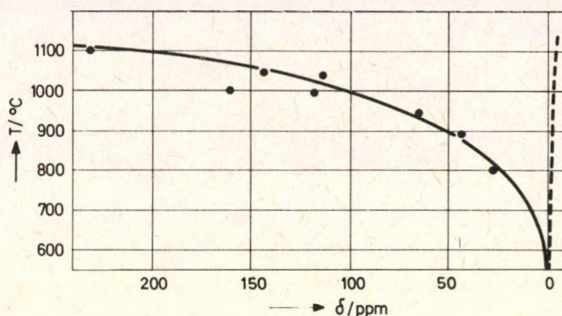


Fig. 7. Phase diagram of the Zn/O system. The lower phase boundary results from our coulometric measurements, the upper phase boundary is drawn schematically [1]

Measurements of the electrical conductivity and the Hall mobility also show the temperature dependence of the phase boundary composition [7]. In both types of crystals the electrical conductivity σ and the Hall coefficient R_H were measured in the temperature range from 77 to 1000 K by the van der Pauw method. We can summarize these measurements as follows: The crystals from the lower phase boundary show a clear dependence of the carrier concentration n, the electrical conductivity σ, the Hall mobility μ and the donor concentration N_D on the growth temperature T_1 (Fig. 8), whereas for crystals from the upper phase boundary no correlations were found. The correlation between growth temperature T_1 and carrier concentration n and donor concentration N_D in crystals from the lower phase boundary can be understood with the help of the phase diagram assuming that native donors play the decisive role for the electrical properties. On the other hand, this correlation can be used to construct the phase diagram as carried out by us (see Fig. 9). The concentration of the donor given by N_D is, on average, larger by a factor of 2.2 than n (300 K) as is shown in Fig. 9. The line N_D represents the lower phase boundary composition in the system ZnO/Zn. The right-hand scale gives the concentration in ppm. The results are in good agreement with the data of Hagemark and Toren [1], but they are at lower temperatures by a factor 2 lower than our results from coulometric and electrical measurements — with another temperature dependence. The described preparation of ZnO and the determination of its very small homogeneity range provide a striking example of the use of chemical transport reactions for the defined preparation of non-stoichiometric compounds.

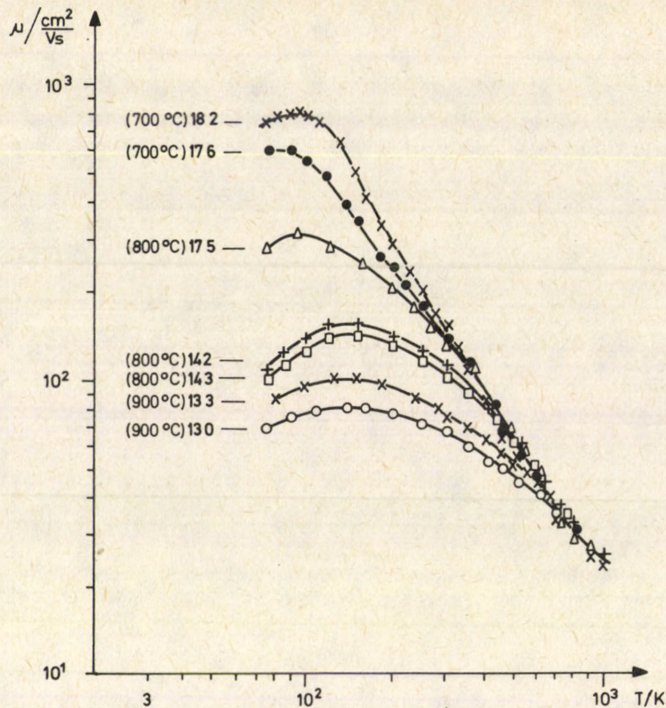


Fig. 8. Temperature dependence of the Hall mobility for samples from the lower phase boundary [7]

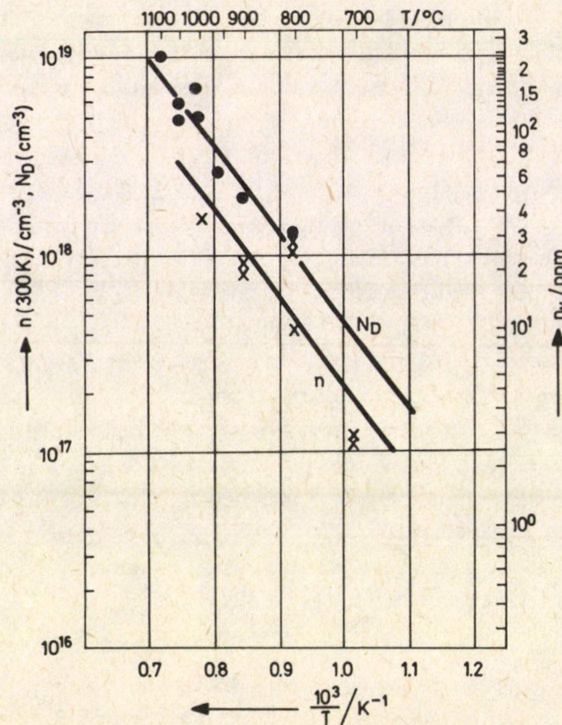
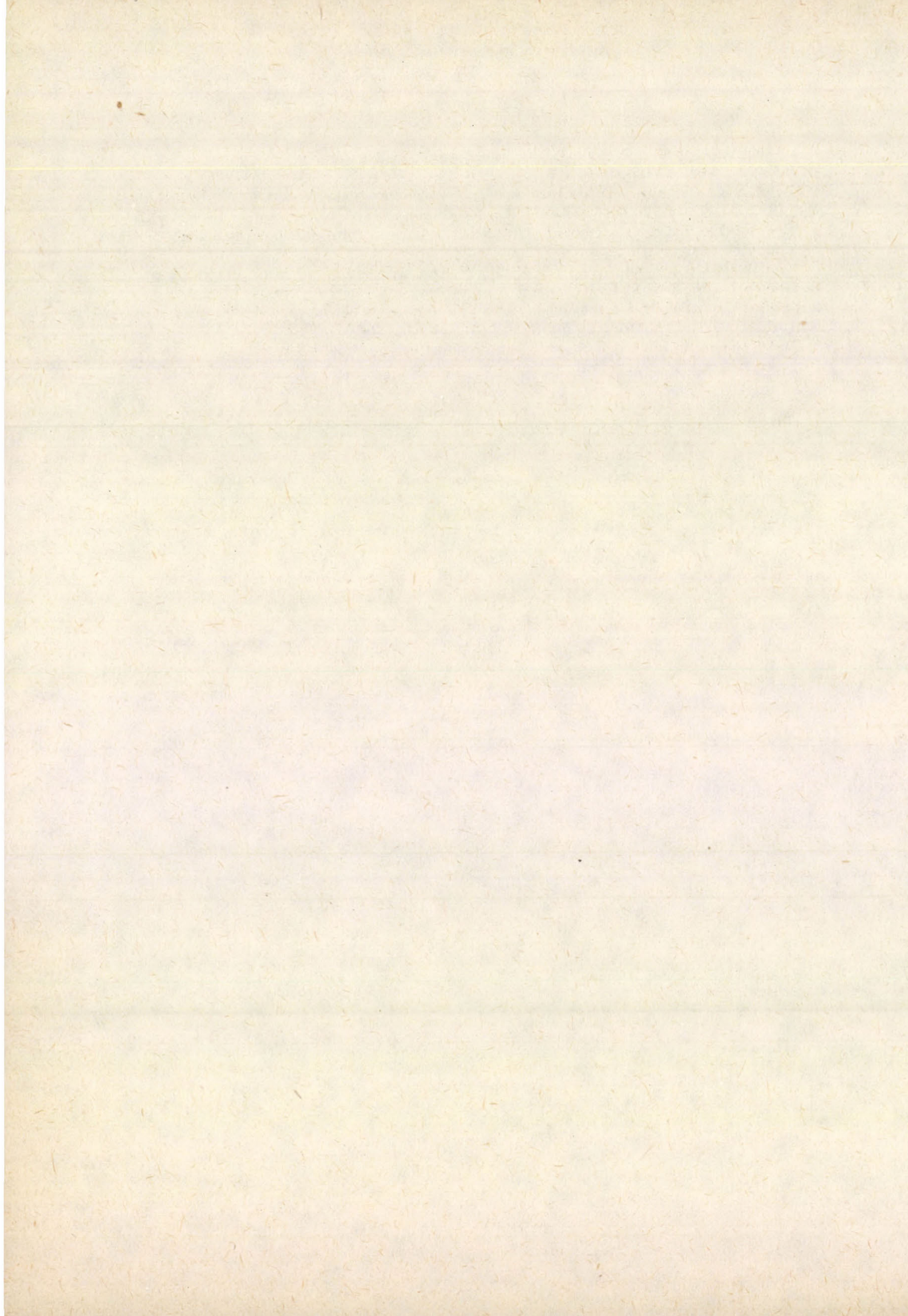


Fig. 9. Electron concentration *n* (*x*-experimental), the donor concentration N_D and the coulometric determined lower phase boundary (0) vs the reciprocal growth temperature

References

1. K. I. Hagemark and P. E. Toren, J. Electrochem. Soc., *122*, 992, 1975.
2. H. J. Fischer, J. Cryst. Growth, *34*, 139, 1976.
3. P. G. Pasko, B. I. Kidjarov, T. D. Levitzkaja and P. P. Neojanceva, Neorg. Mat., *8*, 2, 1972; Neorg. Mat., *8*, 3, 1972; Izd. Nauka Sib. Otd., 193, 1975; Izd. Nauka Sib. Otd., 46, 1975.
4. E. Wolf, H. Oppermann, G. Krabbes and W. Reichelt, Current Topics in Materials Science, North Holland, Amsterdam, *1*, 1978.
5. G. Krabbes, H. Oppermann and E. Wolf, Z. anorg. allg. Chem., *416*, 65, 1975.
6. H. Oppermann, W. Reichelt, G. Krabbes, E. Wolf, Kristall u. Technik, *12*, 717, 1977.
7. E. Ziegler, A. Heinrich, H. Oppermann and G. Stöver, Phys. Stat. Sol. (a) *66*, 635, 1981.
8. R. Kucharkowski and A. Drescher, Fresenius Z. Anal. Chem., *314*, 555, 1983.
9. K. Teske, H. Oppermann and G. Stöver, Z. anorg. allg. Chem., *511*, 72, 1984.
10. K. Teske, Kernenergie, *24*, 20, 1981.



Acta Physica Hungarica, 57 (3—4) pp. 223—232 (1985)

PREPARATION OF GaAs STRUCTURES FOR DIFFERENT DEVICES BY CHLORIDE TRANSPORT

T. Görög, I. Gyuró and K. Somogyi

Research Institute for Technical Physics
Hungarian Academy of Sciences
1325 Budapest, Hungary

GaAs, one of the basic materials of many electronic devices, has become widely used because of the improvements in epitaxial growth techniques. This paper presents the results obtained during the development of GaAs structures. Data are given on the optimum growth parameters, on the requirements against the raw materials, the flow velocity of gases, the growth rate of the layers, on the possibility of reducing the background impurities to $5 \cdot 10^{14}/cm^3$, and on the different electrical measurements performed on structures such as $n^{++}-n_b-n_s$, $n_k^{++}-n-n_b-n_s^{++}$, $n-n_b^{-}-n_s$ and submicron layers.

The wide application of GaAs for device production has been promoted by the intense development of the crystal growing and epitaxial technologies during the last decade [1—2]. At present numerous types of GaAs devices (Gunn and Schottky diodes, MESFET transistors, integrated circuits) are on the market.

The majority of GaAs devices contain different layer structures which can be fabricated by vapour phase epitaxy (VPE), liquid phase epitaxy (LPE), or molecular beam epitaxy (MBE). However, for mass production, VPE has some advantages over the other methods:

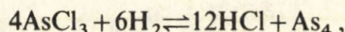
— it can be used to develop a great variety of structures consisting of different $A^{III}B^{V}$ compounds and differently doped layers,

— uniform doping and sharp transitions between the individual layers can be maintained,

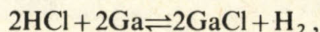
— besides the good reproducibility, the background impurity level is below $5 \cdot 10^{14}$ at/cm^3,

— the short technological cycle and the relatively large substrate area used in every run result in high yields.

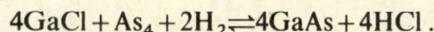
The VPE technology is based on the following transport reactions:



$$4AsCl_3 + 6H_2 \rightleftharpoons 12HCl + As_4 ,$$



$$2HCl + 2Ga \rightleftharpoons 2GaCl + H_2 ,$$



$$4GaCl + As_4 + 2H_2 \rightleftharpoons 4GaAs + 4HCl .$$

The development of GaAsP LED-s and GaAs homoepitaxial layers made in laboratory conditions was reported earlier [3]. The increased interest in our GaAs products required the installation of first-class VPE equipment. The aim of this paper is to present those technological experiments which have been carried out to reach the optimum growth conditions for producing GaAs devices of high quality. Among other aspects the effects of varying the growth rate, the background impurity level and the dopant concentration on the different electrical and other physical properties of layer structures are also discussed.

Experimental

Figure 1 illustrates the technical set-up of the VPE equipment. The system contains a Pd-diffusion cell to purify the hydrogen carrier gas and the four gas lines for transport, etching, by-pass, and doping.

Hydrogen saturated with $AsCl_3$ reaches the source and the deposition zone through the transport and the etching lines, respectively. The by-pass line prevents the diffusion of reaction products back to the source area. The transport line can carry simultaneously two different gas mixtures to the deposition zone.

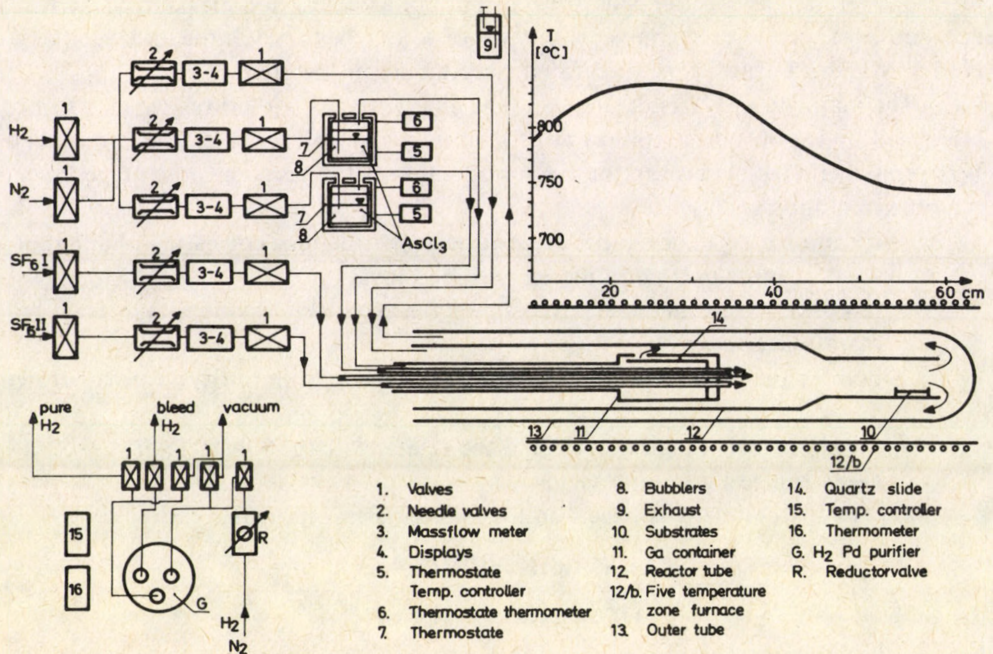


Fig. 1. Temperature profile and scheme of the VPE reactor

After proper saturation of the source material the growth process consists of the following steps:

— the GaAs wafers are placed in the substrate holder,
— the system is heated up to the required temperature,
— the surface of the substrates is cleaned by etching,
— the layers are deposited,
— the reactor is cooled to room temperature.

Materials used for the experiments:

Hydrogen (Tungsram, Hungary) purified in a Pd-cell.
dew point: 183 K
oxygen concentration less than 1 ppm
Gallium (Aluminium Works, Hungary) 6N purity
$AsCl_3$ of semiconductor grade (USSR)

The typical growth parameters of the epitaxial layers:

Source temperature: 1093—1103 K
Deposition zone: 1003—1113 K
Temperature gradient in the deposition zone: 0.8—1.5 K/cm
Temperature of $AsCl_3$: 290.5 K
Flow rates of hydrogen
— in the transport line (b): 400—1000 cm^3/min
— in the by-pass line (a): 2000—4000 cm^3/min
— in the etching line (c): 400—800 cm^3/min

The experiments were designed to determine the optimum parameters for producing $n-n_b-n_s^{++}$, $n_k^{++}-n-n_b-n_s^{++}$, $n-n_{is}$ and $n-n_b-n_i$ structures, where n = active layer, n_b = buffer layer, n^{++} = semiconducting, n_i = semi-insulating and n_k^{++} = contact layers. The thickness and the electrical parameters of some device structures are shown in Table I.

Table I

Parameters	Structures									
	$n^{++}-n_b-n$			$n^{++}-n_b^+-n-n_k^+$				$i-n_b^- -n$		
Carrier conc. [at/cm^3]	$2\cdot10^{18}$	$>6\cdot10^{17}$	$1.5\cdot10^{15}$	$2\cdot10^{18}$	$>6\cdot10^{17}$	$1.8\cdot10^{17}$	$8\cdot10^{17}$		$<10^{15}$	$7\cdot10^{16}$
Thickness of layers [μm]	300	3–4	~10	300	3–4	~11	2–3	300	5–6	<0.3
Hall mobility 300 K	2500		7300	2500		7 300		–	–	4085
Hall mobility 77 K			41 000			43 000		–	–	

In the first experiment series the change of the growth rate was examined as a function of the deposition and the source temperatures, meanwhile the gas-flows were kept constant.

The rate of deposition increases with the source temperature and/or with the temperature of the wafers. Above a certain value of the deposition temperature the growth rate, however, starts to decrease and the curves belonging to constant source temperature are nearly symmetrical (see Fig. 2a). This means that the optimum deposition rates, corresponding to different source temperatures, nearly overlap therefore the ideal deposition temperatures found between 1003—1013 K can easily be determined. Using these experimental data the lowest possible growth rates can be chosen to control the thickness of submicron layers.

Figure 2b shows the dependence of the charge carrier concentration on the deposition temperature at different source temperatures. Based on this relation the most suitable deposition temperatures, resulting in the lowest background impurity level, can be selected for each source temperature.

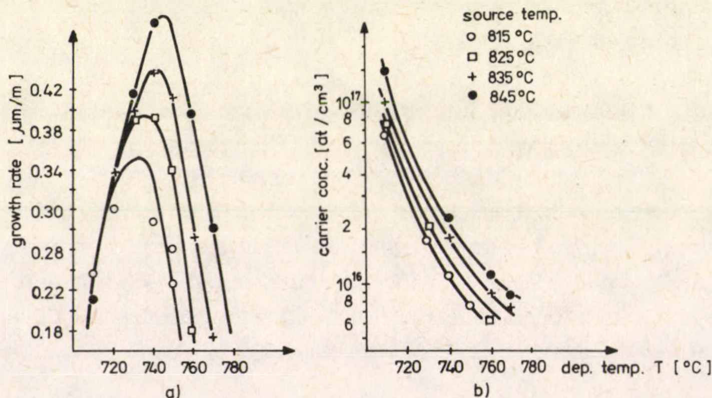


Fig. 2. Growth rate (a) and free carrier concentration (b) vs deposition temperatures

By maintaining constant source and deposition temperatures, one can obtain the changes in the deposition rates (a) and in the carrier concentration (b) as a function of the $H_2:AsCl_3$ ratio (see Fig. 3). The gas mixture entering the reactor tube was varied by using different H_2 flow rates in the by-pass line.

The effect of H_2+AsCl_3, flowing in the etching line, on the growth rate and the carrier concentration are illustrated in Fig. 4/a and Fig. 4/b, respectively.

Comparing Fig. 3 with Fig. 4, it can be seen that the growth rate and the background impurity level are much less influenced by the flow of hydrogen in the by-pass line than by the amount of H_2+AsCl_3 fed through the etching line. Even a small quantity (50—100 cm^3) of H_2+AsCl_3 added through the etching line can considerably

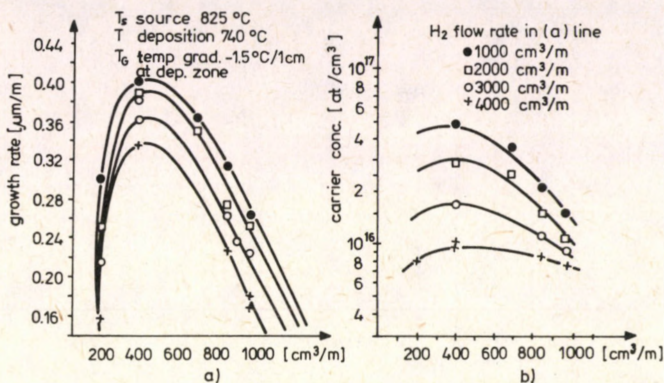


Fig. 3. Growth rate (a) and free carrier concentration (b) vs $AsCl_3-H_2$ flow rate

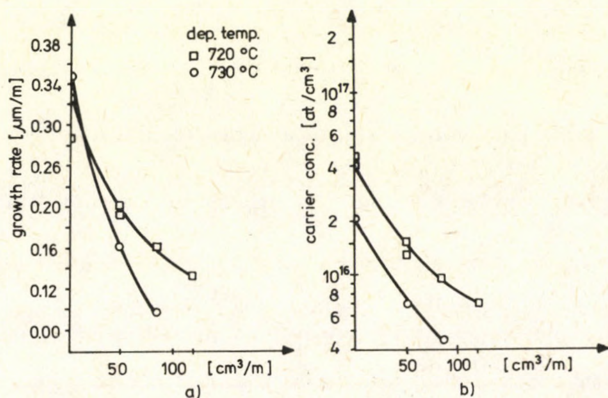


Fig. 4. Influence of by-pass $AsCl_3$ concentration on growth rate (a) and carrier concentration (b)

decrease both parameters, meanwhile — which is very important — the thermodynamic parameters of the system remain practically unchanged.

Utilizing the technological relationships discussed above, one can easily select the optimum growth conditions of individual device structures.

In Fig. 5 the distribution of layer thicknesses (a) and carrier concentrations (b) representing a Gunn structure series is shown in yield diagrams. The planned carrier concentration and thickness were: $Na = 1.5 \cdot 10^{15}$ at/cm^3 and 10 µm, respectively. The average carrier concentration effectively found in 46 samples was $1.25 \cdot 10^{15}$ at/cm^3, within the required 30% limits (18%). The carrier concentration of 9 wafers was exactly the planned value. This is 19% and can be taken as a good result. The layer thicknesses measured in 21 wafers gave an average of 10.14 µm showing only 1.4% difference.

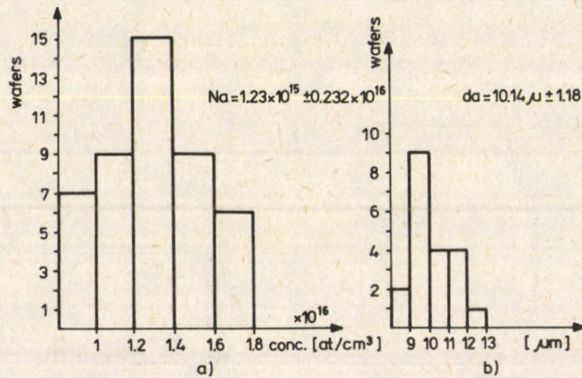


Fig. 5. Concentration (a) and thickness (b) distribution of GaAs epilayers grown for Gunn devices

Measurements

Applying mainly capacitance type measurements, (photo-) electrolytic etching and DC Hall measurements, we have been able to control the most important physical properties of the layer structures: the carrier concentration and the mobility of carriers in the active layer.

Using Schottky barriers made from tiron electrolyte it was possible to realize continuous etching and simultaneous capacity measurements.

The continuous modulation type technique used allowed the determination of the capacitances as a function of the depth and their dependence on the DC bias voltage. From C and dC/dV values the carrier concentration data could be calculated. The samples with Schottky contact were etched under the continuous radiation of a slightly higher energy than the GaAs gap, and the integral of the currents — characterizing the etching process — was proportional to the removed thickness. These profiles allow not only the determination of the carrier concentration but also the layer thicknesses.

The measurements serve two main purposes:

— to help in understanding the technological processes,

— to qualify the structures grown for device purposes.

Figure 6 shows the concentration profiles vs depth of different structures. The typical distribution profile of a multi-layered structure prepared for technological investigations can be seen in Fig. 7. The main informative layers were grown under different conditions to provide a means of following the influence of the individual technological parameters on the carrier concentration and the growth rate. Between the layers there are thin markings with a high doping level to separate the informative layers from each other.

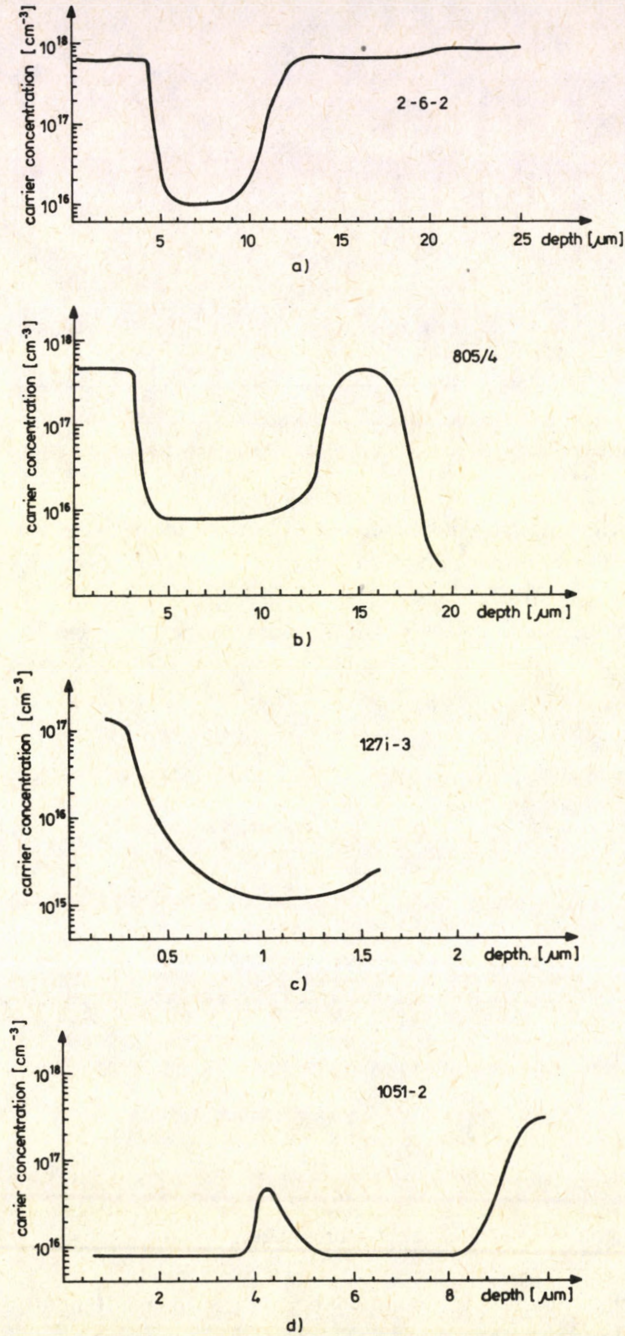


Fig. 6. Carrier concentrations vs depth profiles
a) Multilayer structure on highly doped substrate
b) Multilayer structure on semi-insulating substrate
c) Submicron structure on semi-insulating substrate
d) Structure with a highly doped built-in layer

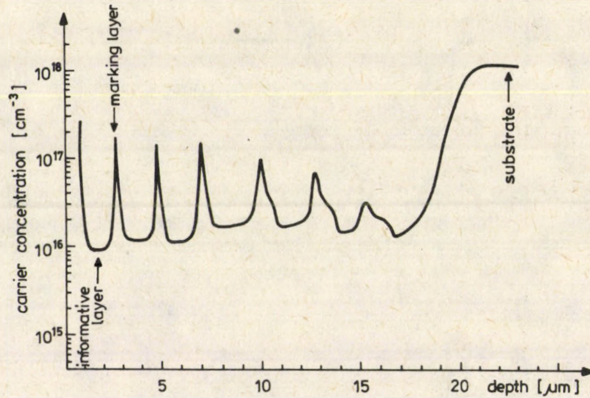


Fig. 7. Concentration profile of a multilayered structure for technological experiment

The typical concentration profile of a Gunn structure is shown in Fig. 8. In order to determine the mobility of the carriers in the active layer, a similar layer was deposited on a semi-insulating substrate in order to enable Hall measurements to be performed. The mobility values found at room temperature are 7700—7900 cm^2/Vs.

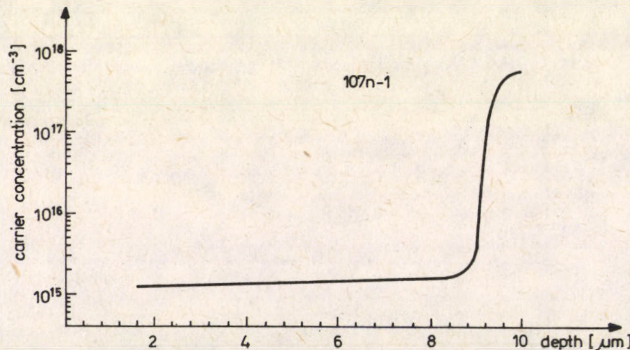


Fig. 8. Concentration profile of a common Gunn purpose structure

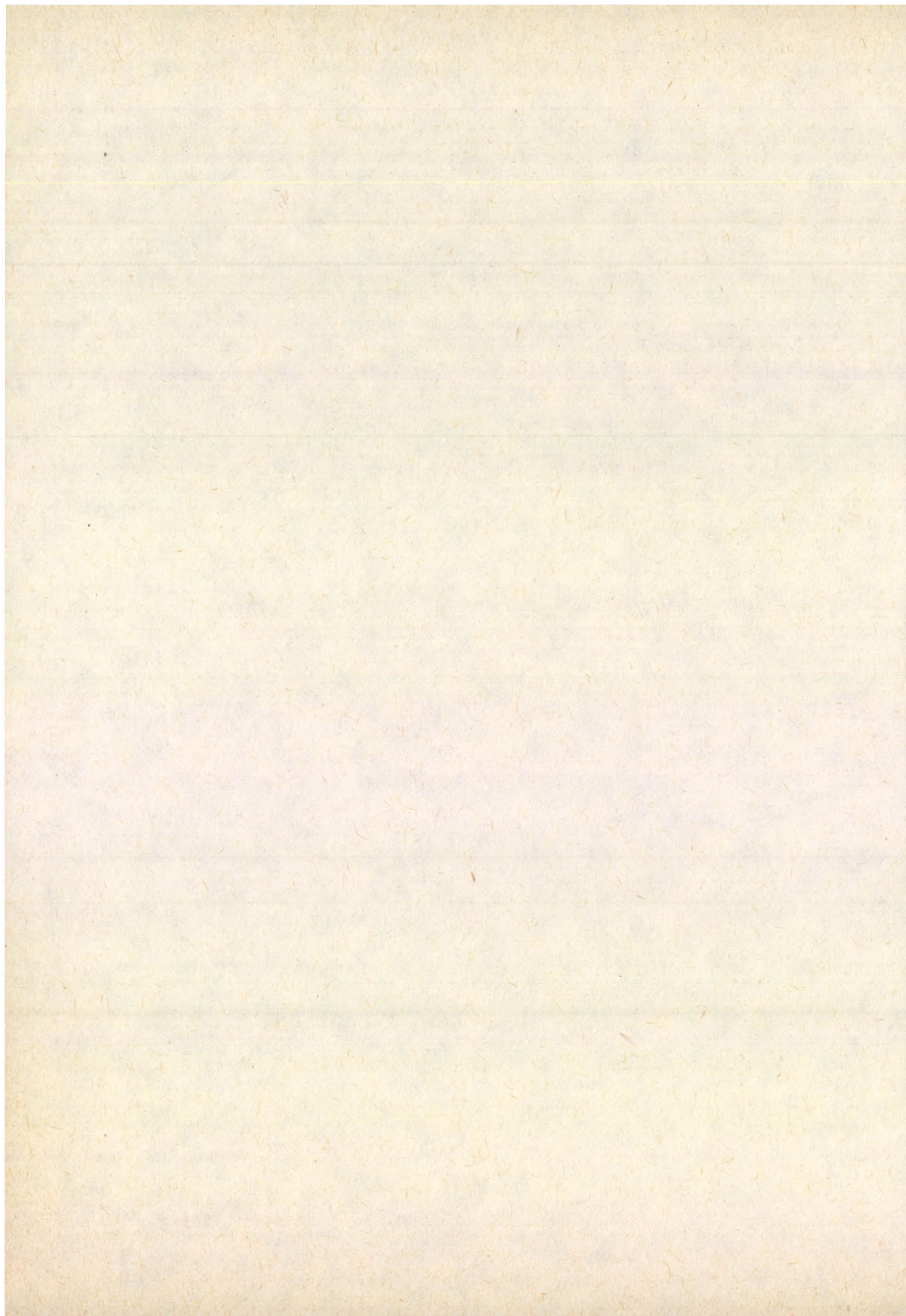
Conclusions

The technological experiments and related measurements have shown the possibility of producing practically all types of GaAs layer structures, with a great variety of carrier concentrations and thicknesses. The wide range of growth rate obtained allows the preparation of submicron structures, too. In addition, using by-pass $AsCl_3 + H_2$, both the growth rate and the background impurity concentration can

be reduced. The measuring techniques applied give a good possibility for the rapid control of the technology and the layer structures. The balanced technology results in good reproducibility and reliability of the structures from which different devices of high quality may be fabricated.

References

1. L. C. Luther and I. V. Dilorenzo, J. Electrochem. Soc., *122*, 760, 1975.
2. J. Korec and M. Heyen, J. Cryst. Growth, *60*, 297, 1982.
3. T. Görög and E. Lendvay, Acta Phys. Hung., *44*, 13, 1978.



Acta Physica Hungarica, 57 (3—4) pp. 233—235 (1985)

GROWTH OF PbSe CRYSTALS FROM LEAD MELT

V. RAKOVITS and M. HÁRSY

Research Institute for Technical Physics
Hungarian Academy of Sciences
1325 Budapest, Hungary

Using the growth from solution technique, PbSe was synthetized in presence of excess lead and crystallized from the metallic melt in a closed, evacuated quartz ampoule. It has been found that the preferred morphological shape of the crystals is the cube with lead enclosures in it.

PbSe, mainly in the form of a microcrystalline thin layer, is a widely used active component of different optoelectronic devices [1, 2]. Its single crystalline properties are less known because the preparation from melt is difficult, requiring a high pressure technique.

The slow cooling (SC) growth of PbSe from nonstoichiometric melts eliminates this problem. Table I shows the solubility of Se in lead [5].

Table I

Se solubility in Pb

Melting point t [°C]	at %
1006.2	40.01
982.1	35.0
969.1	30.0
952.40	25
926	20.5
825.9	10.0
710	2.31
675	1.59

Lead and 15 at% of selenium, both of 5N quality, were weighed into a quartz ampoule and sealed under 10^{-3} Pa pressure. The ampoule was then placed into the homogeneous temperature zone of a vertical furnace, and was kept at 880 °C for three hours during which time the formation of the compound was completed. After this the melt was cooled to 400 °C at a rate of 30 degrees per hour.

The bulky PbSe crystals were thickly covered with lead so they were placed on a quartz plate and the majority of the metal was melted and it flowed away in a hydrogen atmosphere.

Figure 1 illustrates a conglomeration of PbSe crystals grown from a lead solution. It can be seen that crystallization favours the formation of cubes; however, due to the uneven supersaturation and diffusion limited conditions, the growth of the edges and the corners was more rapid than that of the plate centres, resulting in raised edges with some ghostlike metal enclosures in between.

It is of interest to mention that other semiconducting compounds such as the Zn and Cd chalcogenides or the III—V intermetallics in similar conditions, i.e. SC grown from metallic melts, usually follow a dendritic mechanism caused by the development of sufficiently high temperature gradients on the growing surfaces, and form H-dendrites [3, 4].

In the case of PbSe the conditions of thermal conductivity should be different.

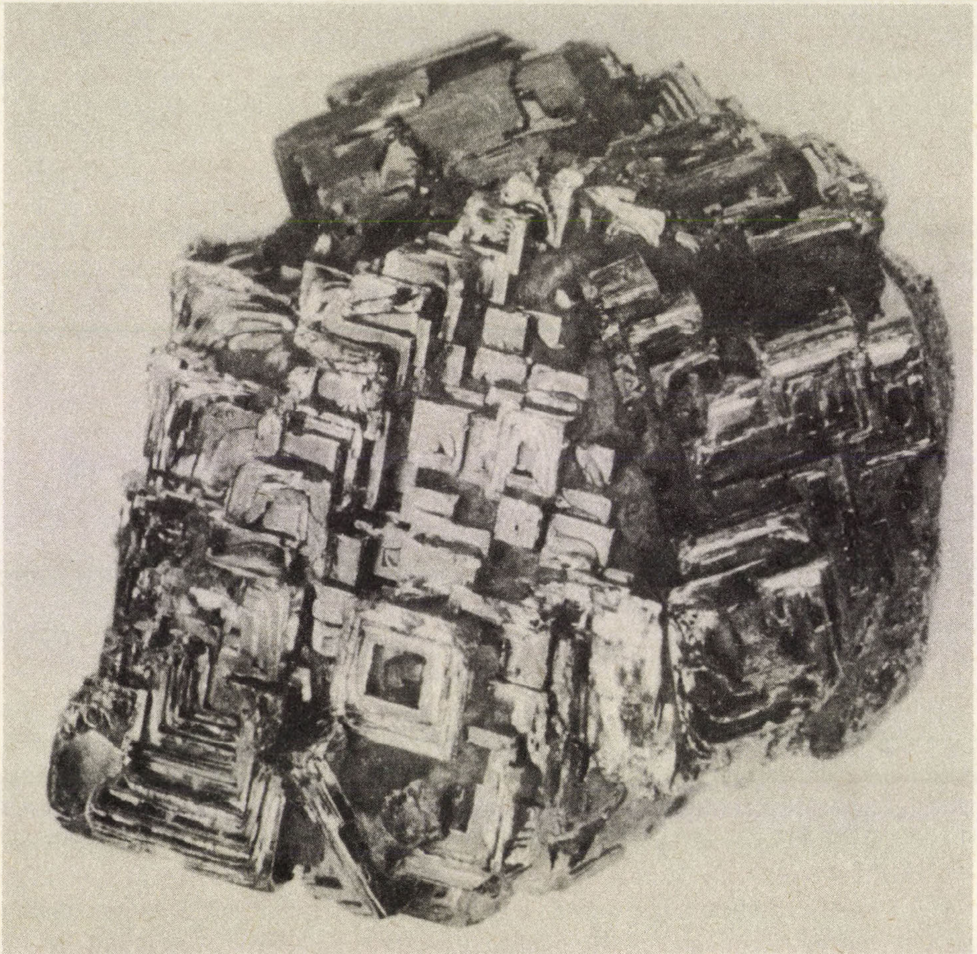


Fig. 1. PbSe grown from lead

References

1. E. H. Putley, Materials used in Semiconductor Devices, John Wiley, New York, 1965.
2. J. W. Faust and H. F. Hohn, J. Phys. Chem. Solids, *25*, 1407, 1964.
3. M. Hársy, Kristall u. Technik, *2*, 447, 1967.
4. M. Hársy, Mat. Res. Bull., *3*, 483, 1968.
5. E. Pelzel, Metall, *10*, 717, 1956.

Acta Physica Hungarica, 57 (3—4) pp. 237—244 (1985)

SELECTIVE EPITAXIAL GROWTH OF $Al_xGa_yIn_{1-x-y}Sb$

E. LENDVAY, L. PETRÁS and V. A. GEVORKIAN*

Research Institute for Technical Physics
Hungarian Academy of Sciences
1325 Budapest, Hungary

Using polycrystalline GaSb substrates and LPE techniques the orientation dependence of the growth rate was studied for AlGaInSb quaternary III—V alloys. On differently oriented substrate grains different growth rates have been observed. In extreme cases a definite inhibition of growth on particular orientations were detected. On twinned regions on one side thick epitaxial layer formed, however, on the other side no epitaxial layer was formed at all. Revealing the substrate grain structure by chemical etching and determining the orientations for both the substrate and epitaxial layer, one-to-one correspondence was found between the substrate structure and the overgrown epitaxial layer.

1. Introduction

Recently the interest in quaternary III—V alloys has increased considerably owing to their application in optoelectronics. These alloys have certain advantages over the pure binary compounds, however, only one quaternary, GaInAsP has thoroughly been examined. Quaternary alloys can be lattice matched to III—V binaries or ternaries and this factor is extremely important in optoelectronics, where the devices used are minority carrier type ones. The lattice matched heterojunction devices are of interest for light-sources and detectors, as well. Different devices operating in the 0.7—1.7 μm spectral range have been developed using the heteroepitaxial principle. These efforts are also supported by the fact that quaternaries have more degree of freedom than binaries or ternaries so the "taylor-made" production of a material with a proper forbidden band gap and lattice constant is easy. In practice, however, the preparation of the homogeneous solid phases of these alloys is rather complicated because of the phase relations, segregation and inmiscibility effects.

One of the most interesting groups of these materials is the pseudo-ternary group. In these alloys three different ions share a single sublattice in the sphalerite lattice. Unfortunately, in the cationic pseudo-ternary group AlGaInP and AlGaInAs can be prepared only in limited composition regions. The third material, AlGaInSb has a complete miscibility over the whole composition region. This material was recently described [1—3]. As it is seen in Fig. 1 AlGaInSb can be lattice matched to AlSb,

* Permanent address: Erevan State Univ., USSR

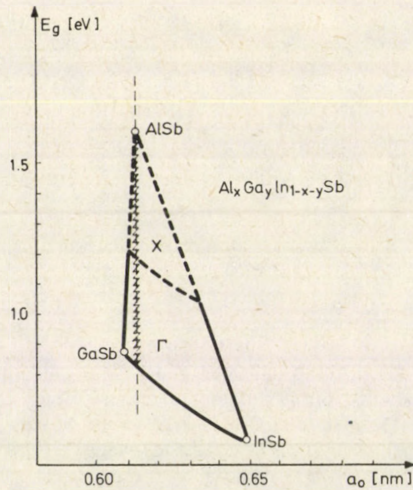


Fig. 1. Forbidden bandgap (E_g) vs lattice parameter curve in the pseudo-ternary $Al_xGa_yIn_{1-x-y}Sb$ system. The shaded area presents the indirect band gap region.

however, the lattice constant of GaSb is also very close to the lattice parameter of the quaternary alloy in a very wide composition range. As the possible bandgap values show, this system is a candidate for developing devices designed for the optical range of 0.2—2.0 μm. According to previous investigations, AlGaInSb has a direct bandgap in the 0.18—1.5 eV region, the indirect bandgap region covers an area falling into the 1.1—2.1 eV range [1].

Although in the whole compositon range there is a complete miscibility, LPE growth of this material is difficult due to the large distribution coefficient of Al. The first successful LPE growth of this material has recently been described [2]. The epitaxial layer formation was studied in different (Ga- and In-rich) melts. The role of the growth parameters (supercooling, growth temperature) and the substrate (misfit) were also studied [4], but the substrate orientation dependence of the growth rate has not yet been studied in detail. This dependence has to be determined in order to control accurately the layer thickness and to develop the theoretical model for calculating the growth rate, too.

It has been reported earlier that the LPE growth rate of InGaAs on the (100) face of InP is larger than that on the (III)B face [5, 6]. Similar orientation dependence was found for InGaAsP/InP heterosystem [7]. In these systems not only the growth rate, but the composition also change with the substrate orientation. For the pseudoternaries no similar data are available. In this paper the dependence of the growth rate on the substrate orientation is studied in the AlGaInSb–GaSb system using a LPE method. AlGaInSb layers were grown from In-rich melts on the GaSb (III)B face and on polycrystalline GaSb surface to determine whether the above mentioned growth rate — orientation dependence exists.

2. Experimental

In order to study the substrate effects $Al_xGa_yIn_{1-x-y}Sb$ layers were grown onto GaSb wafers from quaternary melts. The LPE apparatus consisted of a horizontal furnace system and a multiwell graphite boat of a sliding type. The LPE growth was performed in Pd-purified hydrogen atmosphere. Materials used were semiconductor grade GaSb, InSb as well as Ga, In, Al and Sb of 6N grade. The composition of the growth solution was $X^1_{Ga} \sim 0.07$; $X^1_{In} \sim 0.40$; $X^1_{Al} \sim 0.37$ and $X^1_{Sb} \sim 0.15$ where the X^1_i values represent the atomic fractions. Prior to growth the components were homogenized at 850 °C for about 3 hours, and the melt was cooled down to room temperature. The substrate was placed into the graphite holder at room temperature, and then the system was heated up to the growth temperature (T_g). At the start of the growth run the melt was supercooled first. Reaching the necessary supercooling (ΔT) value the melt was contacted with the substrate and then it was cooled with a constant rate. ΔT supercooling values of 2—10 °C were used. The applied cooling rate was in the 0.1—1.2 °C min^{-1} range. The growth was terminated after cooling down the system 10—20 °C by removing the melt.

It was found that using macrocrystalline GaSb substrates one can determine in one experiment numerous data of the growth rate. The method is very simple and effective. Preparing macrocrystalline GaSb wafers with a known texture [8, 9] after the LPE growth on several, differently oriented grains the difference in growth rate can directly be observed. A further advantage of the method is that the growth rate dependence is measured under the same conditions, so the experimental circumstances are strictly the same for all measured data. The crystallites in the substrate contact the same melt, therefore, the segregation effects can completely be excluded from the seed orientation dependence.

The substrate grain and the epitaxially overgrown layer orientations were measured by X-ray diffraction technique using Cu K_α reflexions. The grain structure was revealed by chemical etching [8, 10] on the back side of the GaSb substrate wafer. Both the substrate grains and the epitaxial layer grown onto the wafer counterpart were examined by Laue technique. The solid composition at the epitaxial surface was measured with a JEOL JSM—35 type scanning electron microscope using both wavelength and energy dispersive X-ray spectrometers. The growth rates were directly determined measuring the layer thicknesses on cleavage surfaces by scanning electron microscopy.

3. Results

The surface morphology of the quaternary $Al_xGa_yIn_{1-x-y}Sb$ LPE layers were examined in relation to the growth rate. Figure 2 shows the typical surface morphology of the epitaxial AlGaInSb layer grown onto polycrystalline GaSb substrate. As it is clearly seen, only a definite part of the substrate is covered by the epitaxial layer. The

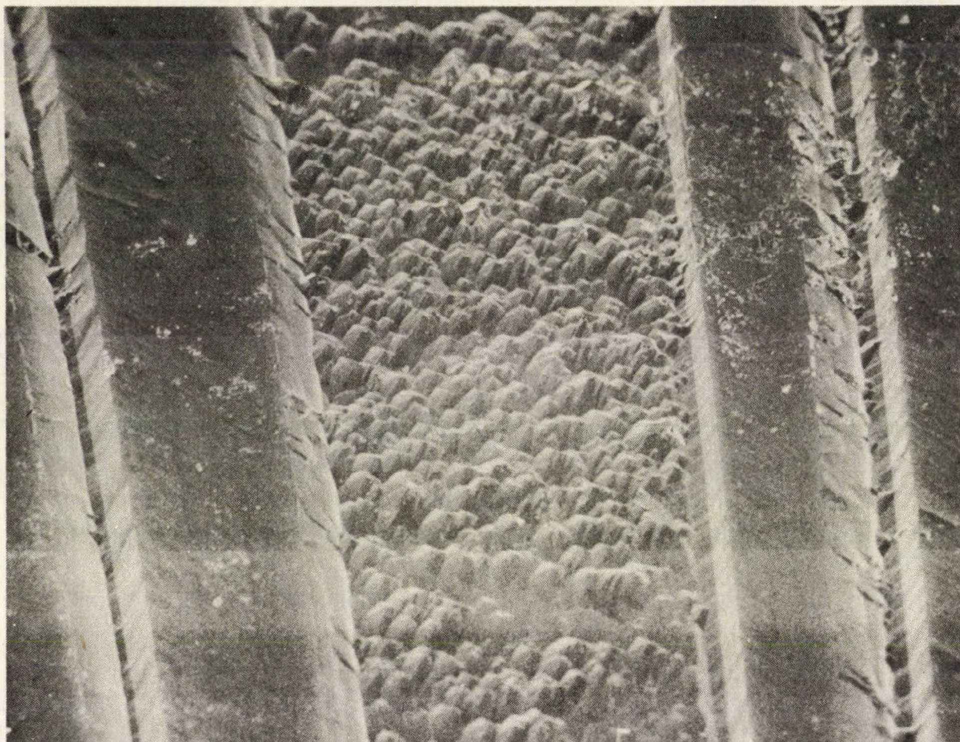


Fig. 2. LPE growth of AlGaInSb onto polycrystalline GaSb wafer from In-rich quaternary melts. ($T_g = 650\,^{\circ}C$, $\Delta T = 5\,^{\circ}C$ and $dT/dt = 1\,^{\circ}C\ min^{-1}$)

as-grown layer of AlGaInSb had normally a rough surface topography and thicknesses on different substrate grains were different. On the samples grown at low supercooling hillocks were regularly seen (Fig. 3). On cleaving through the samples a very strong anisotropy of growth rate has been observed. On the substrate surface there are regions where no epitaxial overgrowth can be found. In Fig. 4 the cleavage cross section of such a region is shown. At the interface there are no traces of substrate dissolution which was observed earlier for ternary antimonide/GaSb systems (see e.g. [11]). Even at the free substrate surfaces neither back melting nor surface roughening could be observed. In order to determine the orientation relations the substrate backside was chemically etched to reveal the grain structure [8]. The substrate thickness being typically 200 μm, the counterpart of the wafer had the same grain structure. Fig. 5 shows both the etched GaSb surface and the epitaxially grown AlGaInSb. It can clearly be observed that a one-to-one correspondence exists between the grain structure and the epitaxial topology. Measuring the orientation of the different GaSb substrate grains, two characteristic species can be found. Some grain surfaces have nearly (112) orientation

(light regions in the Figure. These crystals belong to the [110] zone), the other ones belong to the same zone, but their ⟨111⟩ direction differs from the previously mentioned with an angle of 39°. The two orientations form a twin with a twinning plane of (111). Comparing the two lateral surfaces it is obvious that an epitaxial layer formed only on the first type surface. Previously, for other materials it was found that the saturated composition of the liquid depends on the substrate orientation used and the liquid–solid distribution coefficients also depend on the seed orientation. This means that the equilibrium phase diagram is thus not completely adequate for the exact determination of the LPE conditions. Therefore, the perturbing effects of the substrate on the liquidus and solidus compositions must be considered in the LPE growth, too. The cooling rates used were also fast to achieve near equilibrium conditions. However, for III—V alloys it was described that fast cooling has an unimportant role in determining the orientation dependence of distribution coefficients [12]. Even when the solution with the same composition is used, the a_0 lattice constants of a layer grown on differently oriented faces are different. For InGaAs e.g. the layer thicknesses on the different faces were found different regardless of the lattice misfit [13]. For the explanation of this orientation dependence in a simple case (eg. for (100) and (111) faces)

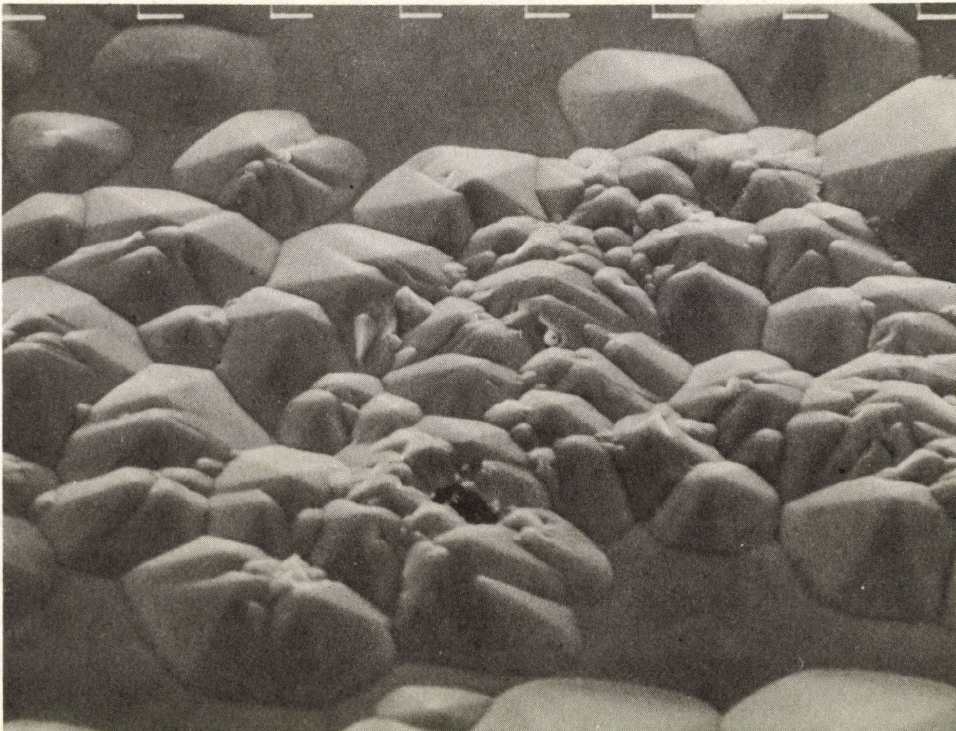


Fig. 3. Growth hillocks on AlGaInSb epitaxial layer surface

Fig. 4. Cleavage cross section of a AlGaInSb epitaxial layer grown onto polycrystalline GaSb wafer

Fig. 6 represents the pseudo-ternary phase diagram. A denotes the initial melt composition at temperature T_1 ($T_1 = T_g$). Paths AB and AC represent the trajectories of the melt compositions during the LPE process. The terminal points (B and C) are on the liquidus isotherm at T_2 ($T_1 > T_2$) depending on the seed orientation. As the solubility of the AlSb component at B is larger than that at C, the amount of AlSb incorporated into the layer at B would be smaller than that on the face presented at C. It means that the growth rate on the B face should be smaller than that on the C face giving a growth rate — orientation dependence. Using polycrystalline substrate the solution is approximately uniform throughout its volume and the composition changes along AD path take place diminishing the effect of the distribution factors. This means that the observed anisotropy cannot be caused by segregation effects alone. This assumption was also proved by a previous work [13], where two, differently oriented substrates were used in the same run set in the graphite LPE boat. During the growth in this experiment only the orientation dependence without the segregation effects can be observed for InGaAs. There is another factor, the so called surface

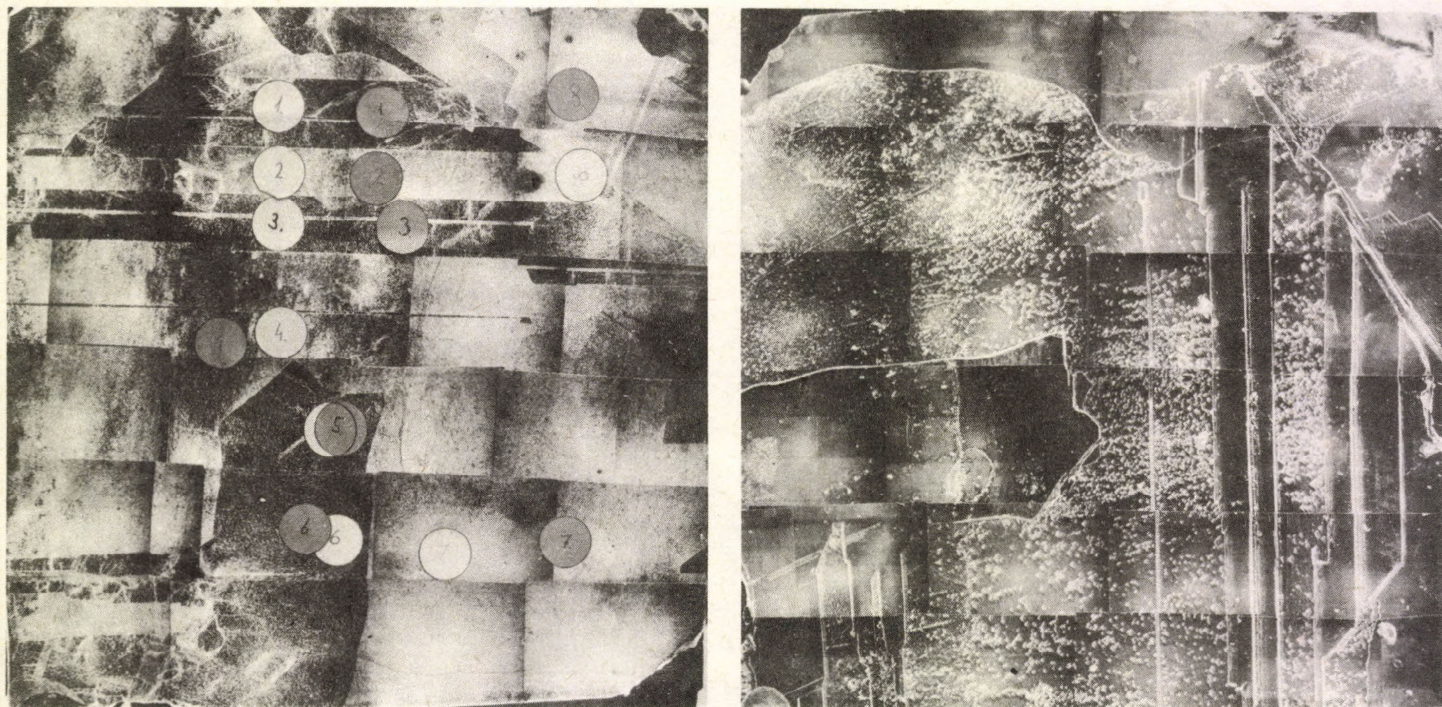


Fig. 5. Heteroepitaxy of AlGaInSb on polycrystalline GaSb substrate:
a) grain structure of the polycrystalline wafer revealed by chemical etching;
b) epitaxial overgrowth of AlGaInSb. Dots denote the spots of Laue measurements (black dots-quaternary, light dots-substrate)

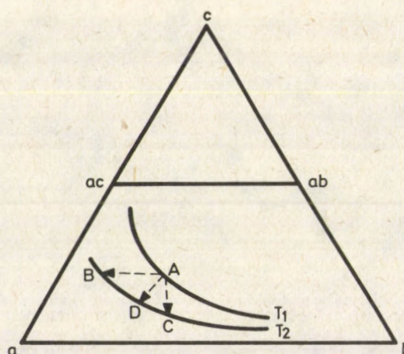


Fig. 6. Pseudo-ternary phase diagram (ac and ab are binary phases) and liquidus curves at T_1 and T_2 temperatures

roughness factor (α) on which the nucleation probability on different faces linearly depends [14]. For sphalerite lattice

$$\alpha = \frac{Z'}{4} \frac{\Delta H}{6.023 \times 10^{23}} \frac{1}{1.38 \cdot 10^{-23} T} = Z' \frac{\Delta H}{33.2T},$$

where ΔH is the fusion enthalpy (per atom) and Z is the number of nearest neighbour sites in the 2 dimensional (2D) nucleus, α must differ from 0, so the effect must be due to some other factors. Similar extreme growth anisotropy (epitaxy on one side of a twin and no overgrowth on the other side) has not yet been found among the III—V heteroepitaxial systems. Most probably the effect is a consequence of the chemical difference existing between the polar faces.

References

1. K. Zbitnew and J. C. Woolley, J. Appl. Phys., *52*, 6611, 1981.
2. E. Lendvay, Electronics Letters, *18*, 407, 1982.
3. E. Lendvay, in "Defect complexes in semiconductor structures", Lecture Notes in Physics Vol. 175, Springer Verlag, Berlin, 1983.
4. E. Lendvay, V. A. Gevorkyan, L. Petrás, I. Pozsgai and A. L. Tóth, J. Cryst. Growth (to be published).
5. K. Nakajima, T. Tanakasi, K. Ahita and T. Yamaoka, J. Appl. Phys., *50*, 4975, 1979.
6. T. P. Pearsall, M. Quillec and M. A. Pollack, Appl. Phys. Lett., *35*, 342, 1979.
7. K. Oe and K. Suqiyama, Appl. Phys. Lett., *33*, 449, 1978.
8. M. Hársy, T. Görög, E. Lendvay and F. Koltai, J. Cryst. Growth, *53*, 234, 1981.
9. F. Koltai, M. Hársy and E. Lendvay, Crystal Res. and Technol., *18*, 1513, 1983.
10. M. Hársy, F. Koltai, I. Gyuró, T. Görög and E. Lendvay, Acta Phys. Hung., *53*, 133, 1982.
11. A. Sasaki, A. Ohishi, E. Sogawa, S. Mizugaki, Y. Taheda and S. Fujita, in "GaAs and Related Compounds" Inst. Phys. Conf. Ser. No. 63., Japan, 1981.
12. T. A. Pearsall, M. Quillec and M. A. Pollack, Appl. Phys. Lett., *35*, 342, 1979.
13. K. Nakajima and K. Akita, J. Electrochem. Soc., *129*, 8603, 1982.
14. K. A. Jakson, J. Cryst. Growth, *24/25*, 130, 1974.

Acta Physica Hungarica, 57 (3—4) pp. 245—250 (1985)

HEAT CONDUCTION vs CRYSTAL DEFECTS IN GaSb INGOTS GROWN BY THE VERTICAL BRIDGMAN METHOD

M. HÁRSY, F. KOLTAI and T. GÖRÖG

Research Institute for Technical Physics,
Hungarian Academy of Sciences
1325 Budapest, Hungary

GaSb growths were performed in a vertical Bridgman arrangement to clarify the role of the radial heat loss, as the source of the radial component of thermal elastic stresses, in the formation of dislocations. Using identical ingot shapes, the variation of growth conditions — including a space experiment — has proved that correct isolation of the ampoule wall from the convection currents of the furnace can considerably decrease the dislocation density. It is concluded that the existence of radial stresses in the growing ingot has to be taken into account in both the theoretical calculations and the development of crystal growing methods.

Nowadays, because of the ever increasing demand for more and more perfect semiconducting materials, the sources and the formation of crystal defects are the subject of many recent publications. In general, the dislocations are supposed as originating from impurity defects, vacancy clusters or as a result of thermal elastic stresses. As for the grain boundaries they can be produced when migrating dislocation lines form a plane or they may also be caused by the wall effect and spontaneous nucleation [1—3].

In certain growth conditions, e.g. Floating Zone Melting and Czochralski Pulling, most of the factors that produce defects can virtually be eliminated. Thermal elastic stresses, however, remain a constant problem. During the widely used techniques based on growth from the melt, the temperature difference between the interface and the cooler parts of the growing ingot initiates stresses which can play a dominant role in creating defects. Usually the major part of the crystallization heat is conducted axially away from the ingot and, due to the temperature gradient, stresses also arise in the axial direction. This fact is the main assumption of many works studying the correlation between the thermo-elastic stresses and the formation of dislocations; on the other hand, the radial component of the heat loss has scarcely ever been taken into account.

Nevertheless, our earlier experiments showed that the radial stress could not be neglected. GaSb ingots grown by the vertical Bridgman method showed an approximately longitudinal texture. The (110) type texture axis subtended an angle of 5—30° to the boule axis. This (110) direction was parallel in every grain while the

crystallographic coordinate systems of the grains were rotated around it. This type of grain structure can only be explained as the effect of radial elastic stresses.

It was supposed that the radial component of stresses was produced mainly by the cooling effect of the air around the growth ampoule. If, therefore, one were to reduce the radial heat transport a decrease in the density of the crystal defects could be expected.

Experimental

GaSb ingots of identical shape (8 mm in diameter, and 50 mm in length) were grown in a vertical Bridgman arrangement [6]. The furnace temperature was kept to within 0.5 °C by using a stabilized DC voltage supply and running the experiments in a room of even temperature. The starting material used was a pre-synthesized polycrystalline GaSb of $7 \cdot 10^{-2}$ ohmcm resistivity and $1.3 \cdot 10^{17}$ at/cm^3 impurity concentration. Two growth series were performed either by varying the lowering rate or by the ampoule construction. The different ampoule shapes can be seen in Fig. 1. One of the growths using an E-type ampoule was carried out in 1980 under microgravity conditions as a part of the Eötvös Project of the Intercosmos Program during the joint Soviet—Hungarian space flight.

Samples cut perpendicularly and axially from the ingots were chemically polished and etched to reveal grain boundaries and dislocations. The Electron Channelling Method was used to determine the low and high angle boundaries [4, 5]. The dislocations were counted on photographs taken in a scanning electron microscope.

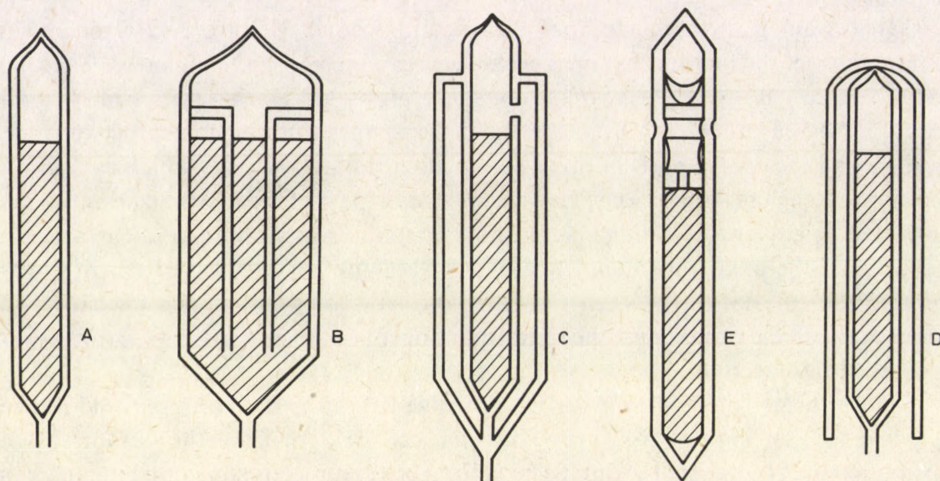


Fig. 1. Quartz ampoules used for GaSb growths

Results

The dislocation densities and the number of grain boundaries found in samples grown in conventional ampoules with different lowering rates are summarized in Table I. Table II presents the same parameters measured on samples grown in different ampoule constructions. Figures 2 and 3 show the cross-sectional fine structure of a space-grown sample and that of the boule which was grown on Earth under the same conditions.

Table I

Sample N°	Ampoule type	Lowering rate [mm/hour]	Dislocation density $\times 10^5/cm^2$	Number of low angle boundaries	Number of high angle boundaries
1	A	0.3	17 ± 10	4	2
2	A	0.6	24 ± 10	6	3
3	A	11.3	65 ± 60	15	9
4	A	18	118 ± 100	16	11

Table II

Sample N°	Ampoule type	Lowering rate [mm/hour]	Dislocation density $\times 10^5 \cdot cm^{-2}$	Number of low angle boundaries	Number of high angle boundaries
2	A	0.6	24 ± 10	6	3
5	B	0.6	36 ± 30	15	7
6	C	0.6	16 ± 5	9	5
7	D	0.6	10 ± 2	4	1
8	E	11.3	5 ± 4	4	0–3
3	A	11.3	65 ± 60	15	9

Discussion

The members of the first growth series (Table I) differ from each other in the lowering rate. At higher cooling speeds the number of defects increases. This effect, of course, may be due to the higher axial or to the higher radial stresses. The radial distribution of defects, however, is not even. It can be seen in Fig. 3, which shows the fine structure of Sample 3 grown at a relatively high lowering rate, that the grain boundaries are concentrated mainly around the centre of the ingot. According to the dislocation counts the densities near to the surface may be a magnitude lower than those in the core of the boule.

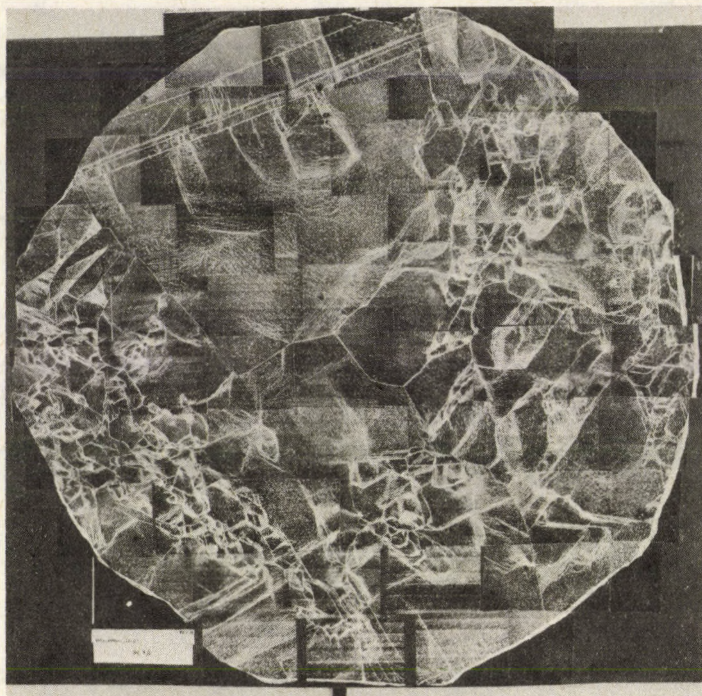


Fig. 2. Fine structure of space grown ingot

Sample 8 (Fig. 2) was grown at the same speed under space conditions. There was only minimal contact between the quartz and ridged surface of the ingot, and there was no air convection around the ampoule. Under these conditions the radial heat loss was highly decreased resulting in an ingot of much better quality with an even dislocation distribution.

Table II represents our efforts to minimize the radial heat loss mainly due to the convection current in the furnace. Sample 5 is the central section of a large ingot of 30 mm in diameter, separated by an inner quartz tube from the rest of the material. This type of isolation, however, has not given any improvement. Comparison of Sample 2 and Sample 5 shows the latter to have such an average dislocation density which corresponds to the central area of Sample 2.

Sample 6 was placed into an evacuated outer ampoule. This vacuum envelope decreased the conduction in every direction and resulted in an ingot of slightly better quality than Sample 2.

The simplest ampoule modification — a conventional ampoule covered with a quartz bell — has given the best terrestrial result. The bell seems to reduce effectively the air convection around the ampoule. The dislocation density found in Sample 7 is less than half of that in Sample 2.

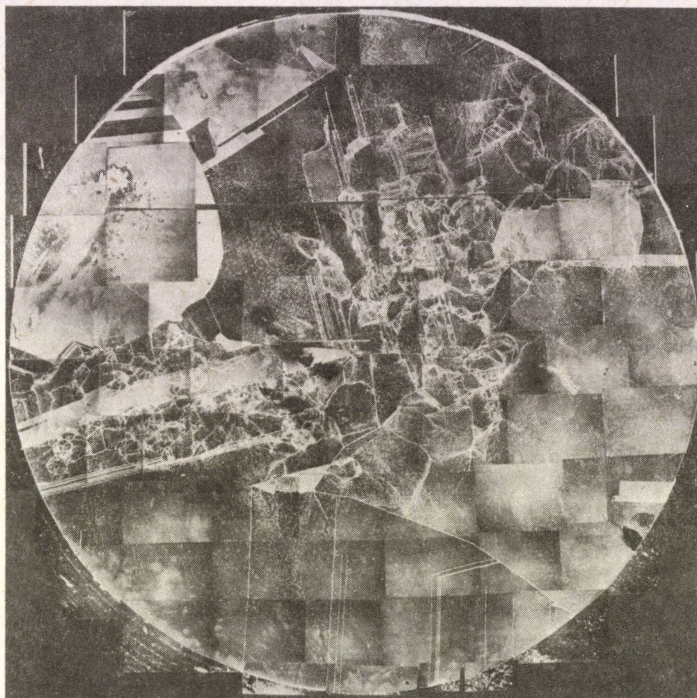


Fig. 3. Fine structure developed on terrestrial control GaSb

The number and the distribution of dislocations show a definite correlation with the cooling effect of the convection current. When the chimney effect was effectively restricted by the application of the quartz bell (Sample 7), or even in the case of the space experiment (Sample 8), the ingots were characterized by the low density and the uniform distribution of dislocations. Using conventional set-ups, the distribution is uneven; higher densities can be found in the centre rather than near to the surface of the ingot.

Similarly to the density of dislocations, the number of grain boundaries found in the different samples is influenced in the same way by the lowering rate and the ampoule construction. (See Tables I and II.) Their distribution, however, is not so unambiguous. In samples grown at a relatively high lowering rate the wide angle boundaries seem to be concentrated in the centre of the ingot (Fig. 3) whereas in other cases they show a random distribution. Usually the lower part of the ingot contains less boundaries than the upper section which solidified later.

Conclusions

The growth experiments performed in different ampoule constructions have proved the effect of the radial heat loss on the dislocation formation. Correct isolation of the growing ingot from convection currents results in a considerable decrease in the dislocation density. The lack of convection is one of the reasons why crystals of good quality can be grown in space conditions. Using a simple bell cover in terrestrial experiments, it nearly halves the number of dislocations. Therefore, the radial component of the thermal elastic stresses as a defect forming factor has to be considered together with the axial component. The reduction of the radial heat loss by using a proper shielding can be utilized in other crystal growing techniques too.

While the appearance of low angle boundaries can be related to elastic stresses, the grains misorientated with wide angles are probably generated by random nucleation with the wall effect having a minor role.

References

1. S. V. Tsivinsky, Krist. u. Technik, *10*, 5, 1975.
2. S. V. Tsivinsky and L. A. Maslova, Krist. u. Technik, *15*, 123, 1980.
3. L. G. Eidelman, Cryst. Res. and Technol., *18*, 405, 1983.
4. M. Hársy, F. Koltai, I. Gyuró, T. Görög and E. Lendvay, Acta Phys. Hung., *53*, 133, 1982.
5. F. Koltai, M. Hársy and E. Lendvay, Crystal Res. and Technol., *18*, 1513, 1983.
6. M. Hársy, T. Görög, E. Lendvay and F. Koltai, J. Cryst. Growth, 53, 234, 1981.

Acta Physica Hungarica, 57 (3—4) pp. 251—261 (1985)

X-RAY AND SEM INVESTIGATION OF CRYSTAL DEFECTS IN N-TYPE, P-TYPE AND SEMI-INSULATING GaAs

P. Franzosi and G. Salviati

MASPEC — C.N.R. Institute, 43100 Parma, Italy

In the present work X-Ray topography (XRT) and Transmission Cathodoluminescence (TCL) observations were performed for investigating crystal defects in GaAs single crystals grown by the Liquid Encapsulated Czochralski (LEC) method; n-type Si, S and Te-doped, p-type Zn-doped, semi-insulating Cr-doped and undoped crystals were tested at 300 K.

Single dislocations, dislocation arrangements, such as slip lines, cellular structures and precipitates have been imaged by using both techniques under various experimental conditions. XRT pictures were taken using scanning transmission and reflection geometry, which make use of different diffraction vectors; on the other hand TCL experiments were performed at different accelerating voltages. Consequently it was possible to compare the XRT and TCL images using similar penetration depths and geometrical resolution.

Efforts were made to interpret the results obtained by evaluating the different mechanisms responsible for the defective contrast of the two techniques. The complementary characteristics of the XRT and TCL, which allow for a satisfactory characterization of the defective state of GaAs single crystals, were evidenced.

Introduction

One of the best established techniques for investigating extended crystal defects in single crystal specimens is X-Ray Topography either by using Reflection (RXRT) or Transmission (TXRT) modes. The main advantages of the XRT are that it is possible to obtain both comprehensive images of relatively large samples and information on Burgers vectors of the dislocations and the deformation field at either the precipitates or inclusions. The main disadvantages here are the length of time that the exposure requires, the rather poor resolution, of the order of magnitude of a few microns, and the fact that some features of the semiconducting materials, for example the growth striations, are sometimes not detectable owing to the low anisotropic stress induced by the doping variation.

In as far as the semiconducting materials are concerned, especially the III—V compounds, the Scanning Electron Microscope (SEM) using the TCL method [1, 2], can be employed in order to characterize the defects in these materials with quite a good resolution (~ 1 μm). It must therefore be noted that one of the limits affecting the TCL resolution is the minority carrier diffusion length of the specimen investigated.

In this paper we will compare the XRT and TCL images, by using similar penetration depths and geometrical resolution, in order to study single dislocations,

dislocation arrangements, such as slip lines, cellular structures and precipitates. Both techniques were used under various experimental conditions: i) XRT pictures were taken by scanning transmission and reflection geometry and different diffraction vectors; ii) TCL experiments were performed at different accelerating voltages.

The aim of this work is to interpret the results obtained by considering the different mechanisms responsible for the defect contrast of the two techniques and to demonstrate the validity of the TCL method for studying semiconducting materials which have a doping level of 10^{16} cm^{-3} or lower.

Experimental

XRT experiments were performed using the Lang technique. The samples, mounted on a conventional Lang camera, were impinged by a collimated X-ray beam with a horizontal divergence of about two minutes of arc. The geometrical resolution of the experimental apparatus was estimated to be in the range of a few microns because of the finite height of the X-ray tube focal spot. Cu $K\alpha_1$ radiation was used for taking topographs either by anomalous transmission scanning geometry or by the reflection one. Ilford L4 nuclear plates were used to record X-ray topographs in both cases.

As for the TCL observations, a solid state Si detector and its preamplifier circuit were placed just under the specimen within the vacuum chamber of a Cambridge Stereoscan 250 SEM [2]. Owing to the close proximity of the sample to the detector, a good signal to noise ratio and a high collection efficiency (70—80%) were obtained even if measurements were made at 300 K. An active area of 0.2 cm^2 made it possible to image relatively large areas of the specimens. The samples were thinned by grinding them mechanically with a diamond abrasive up to a thickness of about 100—150 μm, polished to a mirror-like finish on both sides and the residual work damage was removed by usual chemical polishing etching procedures.

Single crystals of GaAs n-type Si, S and Te-doped, p-type Zn-doped, semi-insulating (SI) Cr-doped and undoped were investigated. The electrical and crystallographic properties of the crystal samples investigated are shown in Table I. With the exception of the Zn-doped specimen, all the others were grown at the MASPEĊ Institute using the LEC method.

Results and discussion

Figure 1 shows a TCL micrograph and a TXRT image of the same area of a Si-doped GaAs specimen ($N_D = 1.6\ 10^{17}$ cm^{-3}). The orientation of the surface studied was the (001). A particular scratched zone of the specimen (*a* in Fig. 1a and *b* in Fig. 1b), was selected to emphasize how the two images correspond. In the TCL image it is possible to see growth striations and dislocations in the two [110] and [$1\bar{1}0$] directions. It can

Table I

Electrical and crystallographic properties of the investigated crystals

Type	Dopant	Carrier concentration $[cm^{-3}]$	Dislocation density $[cm^{-2}]$	Orientation
n	Si	$3-5\ 10^{18}$	10^2	111
n	Si	10^{17}	10^4	111
n	Si	$1.6\ 10^{17}$	10^4	100
n	Si (not intentionally doped)	$2 \cdot 10^{16}$	10^5	111
n	S	10^{18}	$0-10^3$	111
n	Te	10^{18}	10^2	110
p	Zn	10^{18}	10^3	110
S.I.	Cr	—	10^4	111

be emphasized that the observed strong TCL contrast, due to the dislocations, is connected with their known non-radiative recombination centres behaviour [3, 4]. The black areas in the corners of the picture, due to geometrical conditions and corresponding to apparently non active zones, are caused by the circular shape of the solid state detector placed under the specimen.

In the TXRT image, it is possible to see dislocations in the [110] direction and some scratches due to work damage. The reason for this apparent difference between the two images can be easily explained:

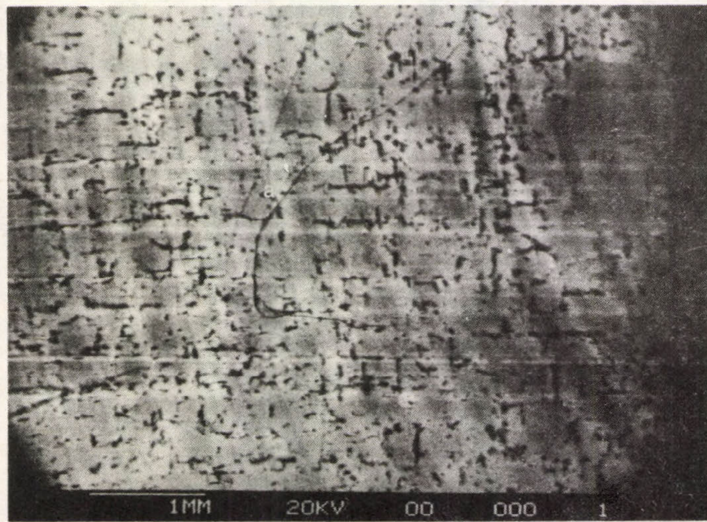
i) because the dislocations, for which $\bar{g} \cdot \bar{b} = 0$, do not show XRT contrast and, since the dislocations in Fig. 1 were proved to have $\bar{b}$ normal to their directions, it was possible to see only those dislocations which were perpendicular to $\bar{g}$.

ii) because a TXRT geometry was employed and since the thickness of the specimen was ~120 μm, it was also possible to see some features of the underside of the specimen (c_1, c_2, c_3 in Fig. 1b).

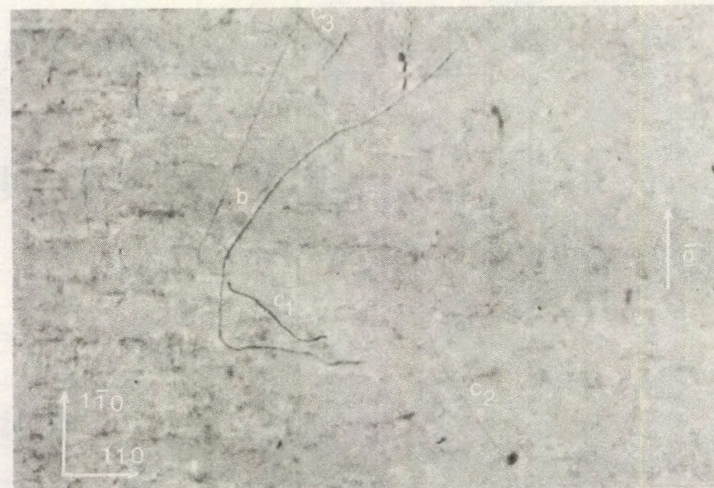
On the other hand, it is impossible to obtain the same results with the TCL mode, because only information can be obtained within a thickness of 3—5 μm when, as in this case, the technique is used with an accelerating voltage of 20 KV [5].

It can be noted that in the TXRT image the growth striations are not visible and this is probably justified by the fact that growth striations, in this case, are due to a small variation in the doping level. Consequently, the above variation does not justify the presence of an anisotropic stress within the specimen which is large enough to give a XRT contrast.

On the contrary, in the TCL image it is possible to see growth striations owing to the fact that this technique is also sensitive to small local doping level variations [6]. Therefore, the reason for the apparently inexact correspondence between the two pictures is due to a simple contrast feature. In Fig. 2 it is possible to see a TCL and a RXRT image, of a (111) oriented GaAs – SI Cr-doped specimen at 300 K grown by the

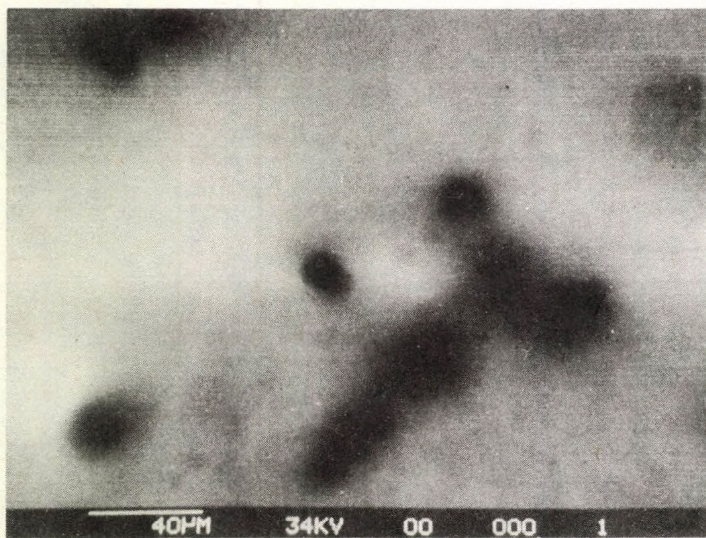


a)



b)

Fig. 1. TCL micrograph a) and TXRT image b) comparing a scratched zone of a (001) GaAs Si-doped specimen ($N_D = 1.6\,10^{17}$ cm^{-3}). Growth striations and dislocations are revealed



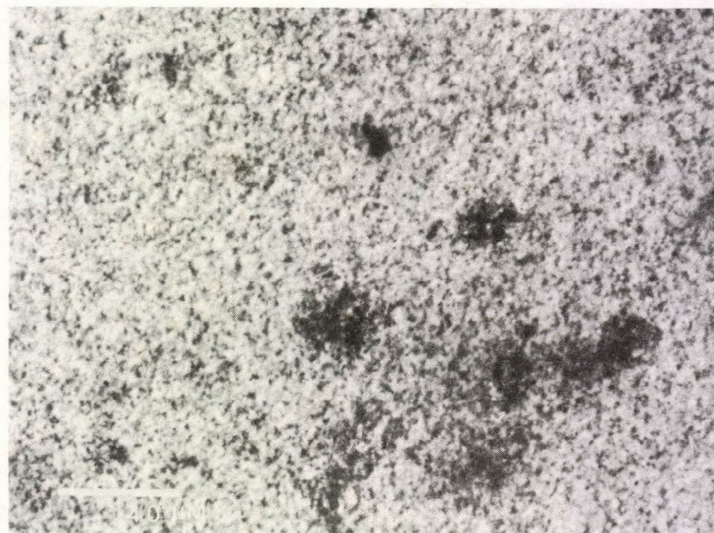


Fig. 2. TCL and RXRT images of dislocations in a (111) oriented SI – GaAs Cr-doped. The one-to-one correspondence between the two images is evidenced

LEC method. RXRT images were obtained using Cu Kα_1 radiation and 333 symmetric reflection. An X-ray penetration depth of 8.7 μm was connected with the X-ray diffraction. Because the geometrical resolution was estimated to be about 2 μm, a direct correlation between TCL and RXRT images can be made.

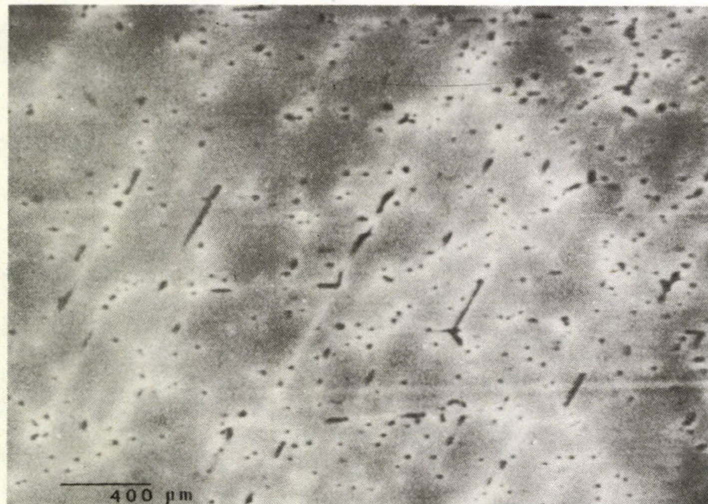
This Figure shows two enlarged micrographs of the same area in which the two imaging methods are compared. In spite of the weak contrast of the XRT image, due to the experimental conditions, dislocations, some of which emerged along inclinated directions, while others were nearly perpendicular to the surface, were imaged and a one-to-one correspondence between TCL and RXRT images was confirmed. Therefore, it can be concluded that the dark contrast in the TCL images was really produced by dislocations [7].

Comparing the pictures of Figs 1 and 2, it is possible to see that both the transmission and the reflection XRT mode confirm the validity of the TCL technique as a valid tool for imaging dislocations in GaAs single crystals ranging from heavily doped to semi-insulating samples.

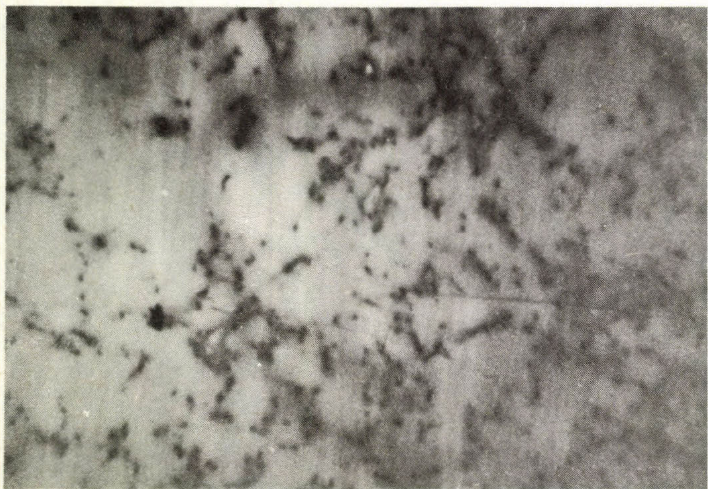
In Figure 3 three TCL micrographs of (111) oriented GaAs Si-doped crystals with different doping levels are shown. In Fig. 3a it is possible to see a typical cellular structure with dislocations lying in the $\langle 110 \rangle$ directions and emerging at different angles with respect to the surface in a specimen with $N_D = 10^{17}$ cm^{-3}. In the same picture it is possible to note both the impurity gettering effect at dislocations [3] and the different luminescence efficiency due to the different defect concentration [8]. Figure 3b shows dislocations (black spots surrounded by white halos) and growth striations in a specimen with $N_D = 3$—$5\ 10^{18}$ cm^{-3}. The horizontal white lines are not due to the contrast but only to the noise from the TCL apparatus amplifier. Figure 3c shows a TCL image of a crystal with a carrier concentration of $N_D = 2\ 10^{16}$ cm^{-3}; the surface studied was always the (111) one and in this case no additional pertinent information is given even if irregular cellular structures surrounding dislocation-free areas are evidenced. In any case the different arrangement of the defects in the three pictures and their varying concentration, due to the different doping levels, are features to note.

In addition to the above, GaAs p-type Zn-doped specimens were also studied. The TCL image of a (110) Zn-doped specimen ($N_A = 10^{18}$ cm^{-3}) is shown in Fig. 4. Besides a number of dislocations, the interesting feature of this picture is that some areas of the specimen have an uneven contrast level as a result of the fact that some areas are lighter than the more homogeneous grey of the background. Until now it has not been clear if this is due either to a non-homogeneous dopant distribution within the ingot, or to a non-uniform absorbtion from the bulk of the transmitted radiation or to clusters of precipitates. In any case some more experiments will be carried out in our Institute to interpret this phenomenon.

The presence of microprecipitates was very evident in GaAs n-type Te-doped samples ($N_D = 10^{18}$ cm^{-3}). As shown in Fig. 5, besides a scratch and a few dislocations, the most relevant feature of this specimen is a very small scale precipitation of the

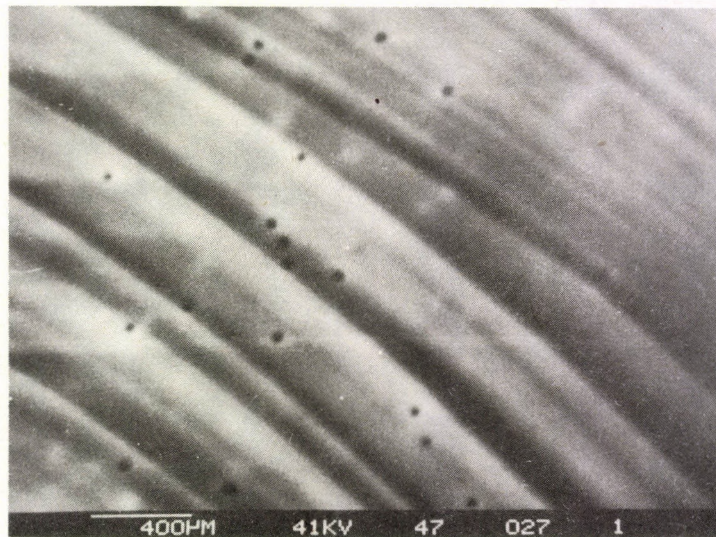


a)



400 μm

b)



c)

Fig. 3. TCL micrographs of (111) oriented GaAs Si-doped specimen. a) $N_D = 10^{17}\ \mathrm{cm}^{-3}$; b) $N_D = 3-5\ 10^{18}\ \mathrm{cm}^{-3}$; c) $N_D = 2\ 10^{16}\ \mathrm{cm}^{-3}$. The accelerating voltages in (a) and (c) are the same as in (b). For some more details see text

dopant, which TCL revealed to be a dense mass of small dark spots on the grey background. Impurities gettering around the dislocations is also well evidenced to be a precipitate-free region. It can be noted that the TCL images of dislocations and microprecipitates are clearly distinguishable owing to their different dimensions (8 μm for the dislocations and 1 μm for the precipitates).

In Fig. 6 a slice of a GaAs S-doped ingot ($N_D = 10^{18}$ cm^{-3}) normal to the ⟨111⟩ pulling direction can be seen. The picture shows dislocation slip arrays along a ⟨110⟩ direction, whose density seems to decrease as it approaches the centre of the slice. A non-homogeneous contrast distribution is also well evidenced. Figs 7a and 7b show TCL and RXRT pictures of the centre of the same slice as in Fig. 6. Once again the one-to-one correspondence of the two images which show a core effect of the LEC grown ingot is evident. In this particular case the TCL micrograph shows that the difference between the A and B regions in Fig. 7a is probably due to the presence of S precipitates in the inner region of the ingot. In this picture it is not possible to see growth striations only because of the fact that the stronger contrast within the A zone masks the weaker one within the B zone. The presence of precipitates is justified by the high doping level in the fuse and this interpretation of the TCL image also seems to be supported by the contrast of the XRT pictures.

Conclusions

This series of TCL, TXRT and RXRT micrographs shows that the Transmission Cathodoluminescence is a good technique for imaging crystal defects at room temperature in semiconducting materials, such as n-type, p-type and semi-insulating GaAs single crystals. This technique is a rapid, simple and sometimes more sensitive one compared with the more established and tedious XRT technique. Notwithstanding the fact that the two techniques are completely different as far as the contrast is concerned, a good comparison of the results obtained from the two different experimental approaches is possible. Thus complementary information on the types of defects investigated becomes available and this feature is particularly useful for the characterization of materials such as were studied in this work. Furthermore, in the case of GaAs this technique is not limited by the doping level or the type of dopants in spite of common beliefs. In conclusion, it might be possible to routinely use the Transmission Cathodoluminescence technique for a large scale investigation of semiconducting materials emitting in a range from 0.4 to 1.8 μm (in this case only a Si detector was used but, of course, it is possible also to use a Ge one). On the other hand, not only can this technique be used for characterizing bulk materials but, when combined with the emission Cathodoluminescence [9] it can also be used for epitaxial homo or heterostructures, for example LEDs and LASERs emitting in the above mentioned range.

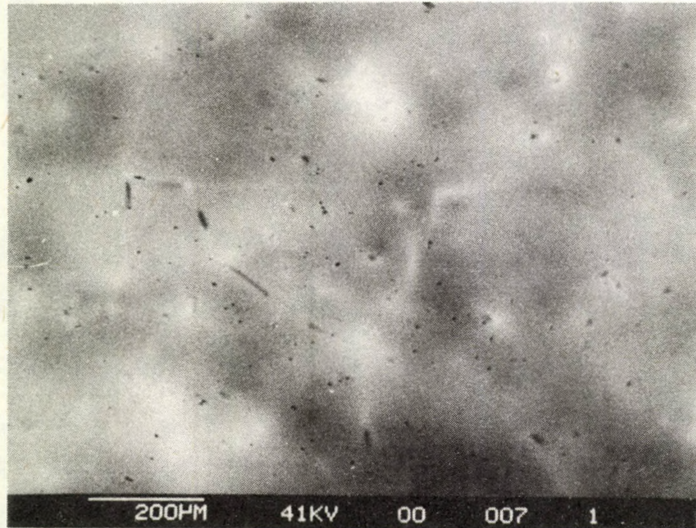


Fig. 4. TCL micrograph of a ⟨110⟩ oriented GaAs Zn-doped sample ($N_A = 10^{18}\ \mathrm{cm}^{-3}$). Dislocations and uneven contrast level are imaged

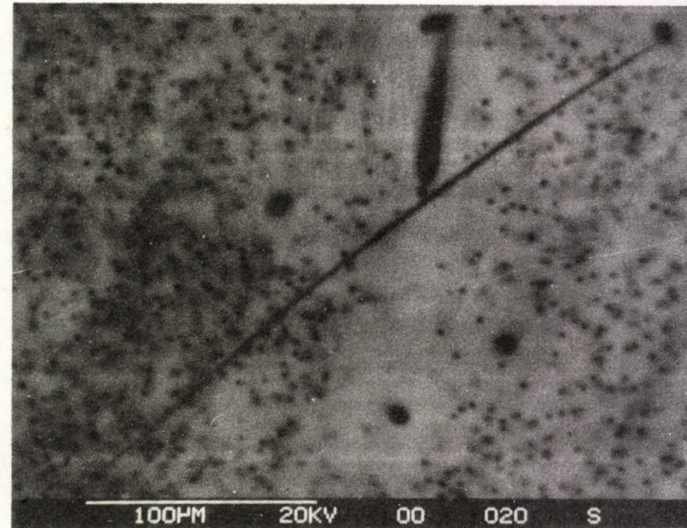


Fig. 5. TCL image of a (110) oriented GaAs Te-doped specimen ($N_D = 10^{18}\ \mathrm{cm}^{-3}$). A few dislocations and a high density of microprecipitates are shown

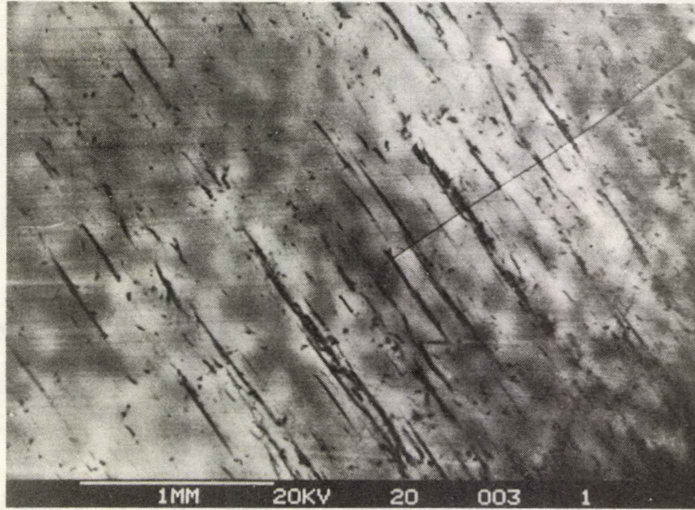
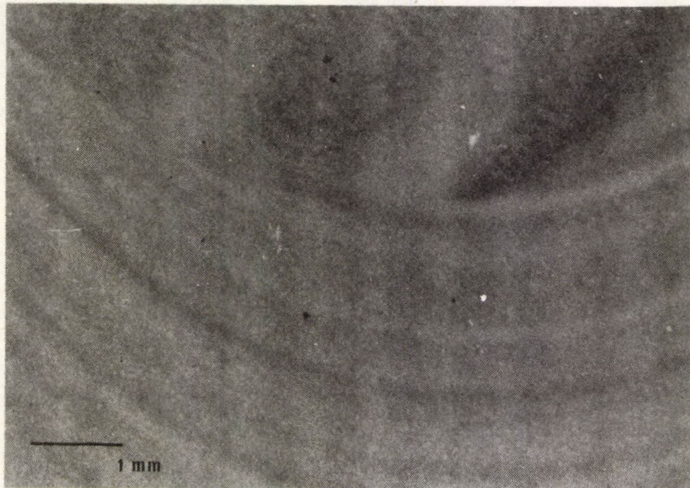


Fig. 6. Dislocation slip arrays and non-homogeneous contrast level in a (111) oriented GaAs S-doped specimen ($N_D = 10^{18}$ cm^{-3}) are shown. The pictures concern the slice periphery



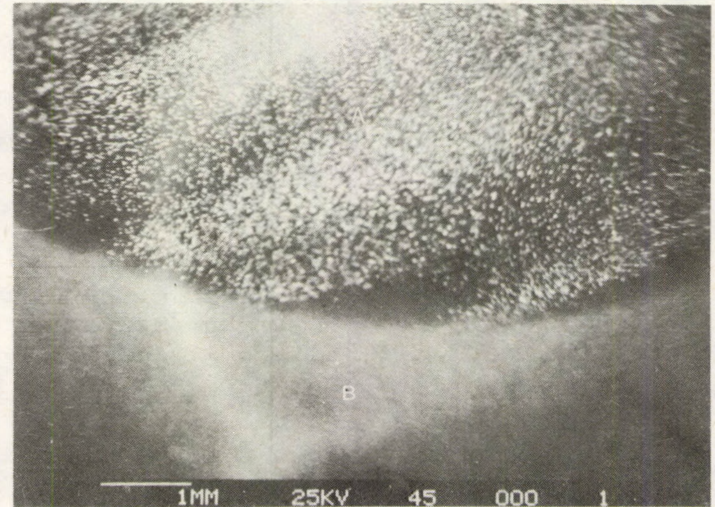


Fig. 7. TCL and XRT pictures of the centre of the same slice of the Fig. 6: an evident core effect is shown. Once again the one-to-one correspondence between the two images is evidenced

Acknowledgements

Thanks are due to Dr. L. Zanotti who grew the samples and to Mr. M. Scaffardi for technical assistance. The authors are also indebted to Mrs. A. Fontana for the supervision of the manuscript.

References

1. A. K. Chin, H. Temkin and R. J. Roedel, Appl. Phys. Lett., *34*, 476, 1979.
2. P. Franzosi and G. Salviati, Materials Chemistry and Physics, MAC 87, Vol 9, 1983.
3. D. B. Holt, Scanning Electron Microscopy, edit. by D. B. Holt, M. D. Muir, P. R. Grant and J. M. Boswarva, Academic Press, U. K., 1974.
4. A. K. Chin, J. Temkin, S. Mahiejan, A. A. Ballman, W. A. Bormer and A. G. Dentai, J. Appl. Phys., *50*, 5707, 1979.
5. K. Kanaya and S. Okayama, Proceedings VII-International Congress on Electron Microscopy, Grenoble, Vol. II. p. 159, 1970.
6. D. A. Cusano, Solid State Comm., *2*, 353, 1964.
7. P. Franzosi and G. Salviati, Journal of Crystal Growth, *63*, 419, 1983.
8. B. D. Chase and D. B. Holt, Phys. Stat. Sol., *a19*, 467, 1973.
9. M. Cocito, C. Papuzza and F. Taiariol, Proceedings of the Institute of Physics Conference Series, No. 67, 273, 1983.

Acta Physica Hungarica, 57 (3—4) pp. 263—270 (1985)

LEC-GROWTH AND THE MAJOR ELECTRICAL AND STRUCTURAL CHARACTERISTICS OF SEMI-INSULATING GALLIUM ARSENIDE

R. FORNARI, C. PAORICI, L. ZANOTTI and G. ZUCCALLI

MASPEC — C. N. R. Institute, 43100 Parma, Italy

Etch-pit density, mobility, Hall constant and sheet resistance have been measured in several GaAs ingots grown by LEC procedure. The main results are that: i) the dislocation density is influenced by the melt stoichiometry; ii) the sheet resistance varies monotonically along the crystal radius; iii) the mobility and Hall constant vary with temperature and behave as if one deep level (which changed its filling state according to T) were present. Moreover, it has been observed that PBN crucibles chemically react with the molten boron oxide, resulting in a contamination of the melt, which can be responsible for crystal twinning.

Introduction

Semi-insulating (SI) gallium arsenide has become important because of its applications in FET technology and in the field of high speed logic devices [1]. For these purposes, materials with high resistivity, high mobility, low concentration of defects and good thermal stability are desirable. To obtain GaAs with the above mentioned properties the following two methods may be used: i) growth of Cr-doped crystals by either Horizontal Bridgman or Liquid Encapsulation Czochralski (LEC) techniques, ii) direct synthesis and growth of undoped GaAs by the LEC method. However, since Cr-doped wafers can give rise to thermal instability effects, due to outdiffusion of the dopant during the annealing processes, necessary either for activating implanted ions or for growing epitaxial layers, undoped high resistivity GaAs seems to be more reliable. Therefore, during the past few years, a lot of research aimed at producing good quality semiinsulating GaAs, without intentional doping, has been developed; consequently the LEC technique has received increasing attention because of the possibility to obtain regularly shaped, large, undoped crystals [2, 3]. Recently, the melt stoichiometry has been identified as the key parameter determining the electrical behaviour of the grown crystal and a critical value of the Ga/As ratio was determined (Ga/As = 0.98), below which the GaAs exhibits high resistivity and mobility, and good thermal stability [4, 5].

A great deal of research has also been devoted to better understanding the compensation mechanisms which lead to a very low conductivity in GaAs. Actually, it is believed that a deep level located at mid-gap (called EL2 and probably related to a native antisite defect) whose concentration heavily depends on the impurity

background and the melt stoichiometry, is responsible for the greatly reduced number of free carriers [6, 7].

In this paper we shall describe the growth procedure that we used to prepare SI undoped ingots of about 300 grams, and we will examine their main electrical properties as determined by Hall effect measurements and dark-spot sheet resistance, with reference to melt stoichiometry. The same parameter will be employed in comparing the dislocation density in several ingots grown under different conditions and emphasis will be placed on the role of stoichiometry in EPD reduction.

Finally, the twinning probability is discussed taking into account the type of crucible used for the growth and the contamination induced by it.

Experimental part

All the pulling experiments listed in Table I were carried out within a high-pressure A. deLittle puller, by adopting the 'in-situ' synthesis and growth process first described by AuCoin et al [8]. The starting materials were arsenic and gallium, 6 N purity, purchased from MCP and Alusuisse, respectively. The crucible in which both the elements were placed had been carefully cleaned, and a pellet of boron oxide (of about 70 grams) was put on the Ga-As charge. Both PBN (pyrolitic BN) (Union Carbide) and silica (Heralux) crucibles were employed, whereas the encapsulant was, in a few instances, B_2O_3 produced by Johnson—Matthey with a water content less than 250 ppm, and in some other cases by Yamanaka with an unknown amount of moisture (certainly more than in the JM pellets).

Pulling rates of 8—15 mm/h were used and the ordinary synthesis and growth runs lasted about 12 hours. The stoichiometry of the melt was determined by weighing

Table I

Undoped GaAs crystals, ⟨111⟩ oriented, obtained by direct synthesis and pulling

Sample	Starting composition (As/As+Ga)	Crucible	Encapsulant	Diameter [mm]	Result
27	0.500	silica	Yamanaka	40–45	single
28	0.500	silica	Yamanaka	40–45	single
51	0.499	PBN I° (+)	J. M.	35–42	single
53	0.512	PBN III°	J. M.	35–40	twinned
54	0.507	PBN I°	J. M.	38–40	single
55	0.496	PBN II°	J. M.	35–38	twinned
56	0.498	PBN I°	J. M.	35–40	single
58	0.499	PBN II°	J. M.	40–45	twinned
59	0.482	silica	J. M.	35–45	single

(+) I°, II° and III° mean new crucible, and crucible already used for two or three runs, respectively

the initial Ga and As amounts, the seed, the crucible and the B_2O_3 pellet and then, at the end of the growth, by weighing the crystal and the crucible containing the residual charge and the solid encapsulant. In this way the exact quantity of arsenic which evaporated, as well as the precise Ga/As ratio in the melt were found.

Specimens for electrical measurements were taken from different zones of the ingots, and their dimensions were about $8 \times 8 \times 0.5\,\mathrm{mm}^3$. Ohmic contacts were obtained by evaporating a gold–germanium alloy (17 : 3, w : w) and by annealing the samples in an inert atmosphere at 350 °C for a few minutes. In some cases, the contacts were not perfect at room temperature, and the ohmic behaviour appeared only above 50 °C. Hall effect measurements versus temperature were carried out using an equipment developed for high resistivity samples. The magnetic field was constantly kept at 4500 Gauss, and the current was varied between 10 nA (at RT) and 5 μA (at 240 °C) to achieve a measurable voltage. Dark-spot resistivity profiles were obtained along the slice diameter, by means of a movable mask, similar to those described in [9], and by a 1000 W white light source.

To investigate the dislocation density the samples were etched in molten KOH (300 °C) for 30 minutes and were then examined using optical and electron microscopes. X-ray topography and transmission cathodoluminescence were also performed.

Results

In Table I a series of crystals with their relative stoichiometry and growth parameters is shown; it can be seen that three of those ingots twinned. We will discuss the probable meaning of this evidence in the next Section.

As for the structural perfection, Fig. 1 shows the results of an investigation on the dislocation density in several ingots pulled in both $\langle 111 \rangle$ and $\langle 100 \rangle$ directions, and obtained by direct synthesis or from commercial polycrystals. It must be stressed that in the $\langle 111 \rangle$ orientation, the density of dislocations tends to be lower than in the $\langle 100 \rangle$ direction. Furthermore, the direct synthesis is always accompanied by a noticeable reduction of dislocations, especially when the liquid phase is nearly stoichiometric. Measured resistivities in our undoped crystals were in the 5×10^6 to $8 \times 10^7\ \Omega \times \mathrm{cm}$ range. Mobility values versus temperature were rather scattered and difficult to analyze correctly, however, the mobility very often exhibited an increase in the RT-100 °C range and had a maximum around 100 °C. From the absolute values we deduced that the conduction in our samples was mixed. In Fig. 2 we report some typical mobility characteristics in the RT-220 °C range; the relative Arrhenius plots are shown in Fig. 3.

In Fig. 4 the behaviour of sheet resistance along the wafer diameter is reported. Our characteristics are different from those shown in [9]. This has been attributed to a different growth procedure, i.e. to different thermal gradients and liquid–solid interface, which influences the local resistivity in the slice [10].

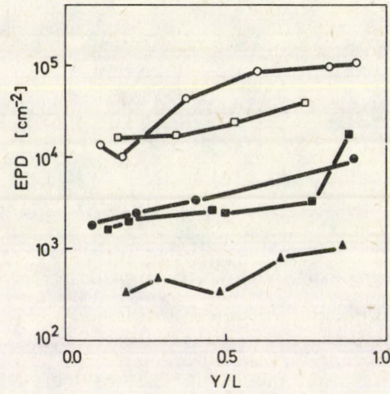


Fig. 1. Etch-pit density in several undoped LEC crystals: ○ ⟨100⟩ oriented, grown from commercial polycrystals; □ ⟨111⟩ oriented, grown from commercial polycrystals; ● ⟨100⟩ oriented, direct synthesis; ■ ⟨111⟩ oriented, direct synthesis; ▲ ⟨111⟩ oriented, direct synthesis, nearly stoichiometric melt. *Y/L* is the grown fraction of the GaAs ingot

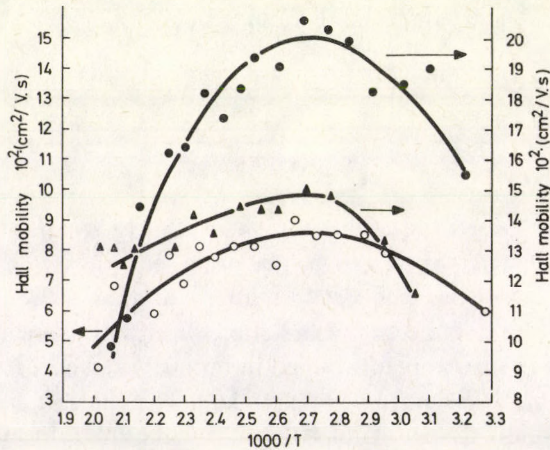


Fig. 2. Hall mobility measured in three samples taken from the top (●, ▲) and from the bottom (○) of a LEC boule

Discussion

A) Dislocations

When undoped material is grown, the hardening effect obtained when some particular dopants [11] are used cannot be exploited. Therefore, to decrease the dislocation density, a reduction of internal and external stresses must be reached. The

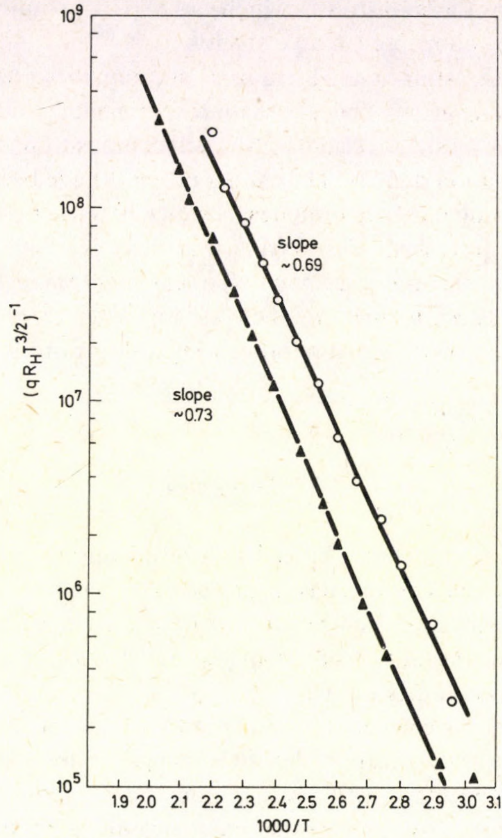


Fig. 3. Arrhenius plot for samples ▲, ○ of Fig. 2

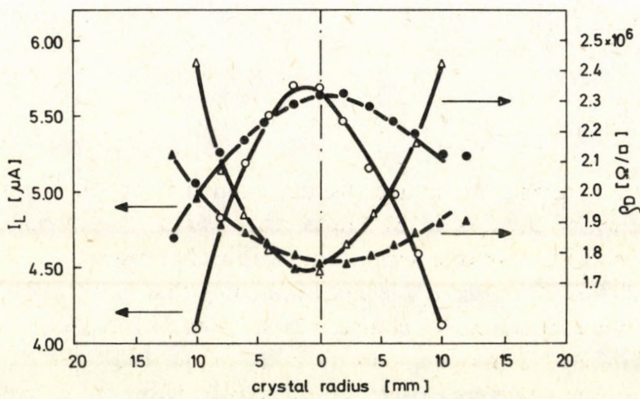


Fig. 4. Experimental characteristics of current (I_L) under illumination and relative sheet resistance (calculated as $\rho_D = V/I_L$) along the diameter of two wafers cut at $Y/L = 0.9$ (●, ▲) and 0.5 (○, △). Constant applied voltage V = 10V

latter ones are given by thermal stresses which can be partly eliminated by lowering the pulling rate [12], and adopting suitable shields.

In Fig. 1 all the crystals investigated were of comparable size and, except for the pulling orientation, were grown under the same experimental conditions, thus it can be assumed that the internal stress related to the melt composition is the dominant factor influencing the dislocation density. The results in Fig. 1 have been interpreted on this basis. In fact, by growing material from nearly stoichiometric melts, the concentration of vacancies can be decreased. This means that there is a smaller probability that dislocations will form through vacancy collapsing. Moreover, the stoichiometry is connected to local defect formation, such as, for instance, microprecipitates and complexes, whose role could be that of preventing dislocations from propagating into the crystal.

B) Twinning

The twinned crystals listed in Table I were always grown in PBN crucibles which had previously been used. This fact has suggested to us that, if the crucible is not new, a reaction between encapsulant and boron nitride takes place, giving rise to a heavy contamination. If the amount of boron in the liquid phase exceeds a certain limit, twinning and polycrystallinity occur.

As a matter of fact, we never obtained twinned crystals when either quartz or new PBN crucibles were used, despite the wide range of melt stoichiometry. These observations are in agreement with a recent study on the residual impurities in high-purity LEC grown GaAs [5]. On the other hand, it would seem that the role played by the stoichiometry in twinning is negligible with respect to the role of impurity content, as long as we remain in the $0.48 < \mathrm{As}/\mathrm{As}+\mathrm{Ga} < 0.52$ range. Perhaps, beyond these limits, the twinning probability quickly increases, according to a well known qualitative study [13].

C) Electrical properties

By testing a number of samples we determined, from the Arrhenius plots, the activation energy of carriers in SI GaAs; this was always around 0.70 eV. The material behaved as if only one level at mid-gap were present.

The Hall mobilities gave a clear indication that both holes and electrons contributed to the conduction, i.e. compensation ratios $\mathrm{EL2}/N_A - N_D \sim 2$. [6]. Most of the samples exhibited a maximum in mobility in the 100—150 °C range (Fig. 3), probably due to a larger electron contribution activated from the deep electron trap. At higher temperatures the mobility decreases because of the significant increase of carrier

scattering. To prove that the n/p ratio is responsible for the change in mobility let us consider the expressions of mobility and Hall constant for mixed conduction:

$$\mu_H = r_H \mu_n \frac{n - p(\mu_p/\mu_n)^2}{n + p(\mu_p/\mu_n)}, \tag{1}$$

$$(qR_H)^{-1} = r_H^{-1} \frac{(p + n(\mu_n/\mu_p))^2}{p - n(\mu_n/\mu_p)^2}, \tag{2}$$

$$np = p_i^2 \qquad \text{(function of } T\text{)} \tag{3}$$

where n, p, μ_n, μ_p have the usual meanings of the density of negative and positive carriers and their mobilities; r_H is the scattering factor and p_i is the intrinsic carrier concentration. For example in one sample we found $\mu_H(50\,°C) = 1100\,cm^2/V.s$ and $\mu_H(100\,°C) = 1500\,cm^2/V.s$; by substituting these values in Eq. (1), assuming that $r_H = 1$, $\mu_n/\mu_p = 15$ and $\mu_n = 5000$ and $4000\,cm^2/V.s$ at 50 and 100 °C, respectively, we found $n/p = 0.013$ at 50 °C and $n/p = 0.033$ at 100 °C. Moreover, since $p_i^2 = 8 \times 10^{15}$ and $4 \times 10^{18}/cm^3$ at 50 100 °C, respectively, we can find the absolute values of n and p for the two temperatures. With these values the Hall constant can be calculated; such $(qR_H)^{-1}$ values fit the experimental data very well. We observed that the Hall constant is negative even if p is larger than n, due to the difference between electron and hole mobility in GaAs. If there is a mixed conduction a quick method of understanding is to look at the Hall constant value at 100 °C. For $(qR_H)^{-1} \leq 5 \times 10^{10}/cm^3$ we have reduced mobility, i.e. mixed conduction. The mobility behaviours found in three samples cut at the top and bottom of an ingot confirm that the filling state of EL2 determines the slope of the initial characteristic of the mobility (Fig. 2). The very different values of mobility for those samples can be explained by the fact that the top and the tail portions of a crystal are grown with a different melt stoichiometry. Since the investigated crystal was pulled from a nearly stoichiometric melt (slightly poor in As at the beginning of the growth) we conclude that the larger the Ga/As ratio, the smaller the n/p ratio, thus the lower the mobility. This confirms Holmes et al's previous results [4].

The sheet resistance measurements showed a minimum resistance in the slice centre (Fig. 4). Similar measurements have shown a close correspondence between dislocation density and leakage current (equivalent to sheet resistance) [9, 14]. In our case it is difficult to directly correlate these entities, since the EPD along the diameter of many samples always followed a W shape (according to Jordan's model [12]), while the resistivity profiles were V-shaped. The shallow impurity accumulation, due to facet effect, in the wafer centre [10], could explain such a discrepancy, also if the trapping action of dislocations is considered.

Conclusions

We have measured the dislocation density in crystals grown from melts with different stoichiometry and we have found that the pulling from nearly stoichiometric melts (by in-situ synthesis) gives less dislocated ingots. This was attributed to a lower concentration of vacancies and to a particular interaction of defects which prevented the dislocations from propagating within the crystal. We also suggested that the contamination due to PBN crucibles is crucial in producing twinned crystals.

Furthermore the behaviours of $(qR_H)^{-1}$ and mobility versus temperature have been explained on the basis of a single deep level at the mid-gap, by taking into account its various filling states within the temperature range examined. Finally, a new figure for the sheet resistance profiles is reported; the minimum resistance is found in the slice centre and interpreted as being due to a higher concentration of shallow impurities in that wafer zone.

Acknowledgement

This work was partly supported by the Progetto Finalizzato CNR 'Chimica Fine e Secondaria'. Some of the crystals listed in this paper were grown with the help of Dr. V. Rakovics of the Research Institute for Technical Physics (Budapest). Thanks are due to Mr. M. Curti and Mr. R. Magnanini for their technical assistance.

References

1. H. Winston, Solid State Tech., January 1983, p. 145.
2. B. Lent, M. Bonnet, N. Visentin and J. P. Duchemin, Microelectr. J., *13*, 5, 1982.
3. H. M. Hobgood, G. W. Eldridge, D. L. Barret and R. N. Thomas, IEEE Trans. on Electron Dev., *ED 28*, 140, 1981.
4. D. E. Holmes, R. T. Chen, K. R. Elliott and C. G. Kirkpatrick, Appl. Phys. Lett., *40*, 46, 1982.
5. H. M. Hobgood, L. B. Ta, A. Rohatgi, G. W. Eldridge and R. N. Thomas, Proc. of Semi-Insulating III—V Compounds, Evian, ed. S. Makram-Ebeid and B. Tuck, Shiva, 1982, p. 28.
6. G. M. Martin, J. P. Farges, G. Jacob, J. P. Hallais and G. Poiblaud, J. Appl. Phys., *51*, 2840, 1980.
7. J. Lagowski, H. C. Gatos, J. M. Parsey, K. Wada, M. Kaminska and W. Walukiewicz, Appl. Phys. Lett., *40*, 342, 1982.
8. T. R. AuCoin, R. L. Ross, M. J. Wade and R. O. Savage, Sold State Tech., *22*, 59, 1979.
9. R. Blunt, S. Clark and D. J. Stirland, IEEE Trans. on Electron Dev., *ED 29*, 1039, 1982.
10. R. Fornari, L. Zanotti and G. Zuccalli, Mat. Chem. and Phys., *9*, 307, 1983.
11. G. Jacob, Proc. of Semi-insulating III—V Compounds, Evian, ed. S. Makram-Ebeid and B. Tuck, Shiva, 1982, p. 2.
12. A. S. Jordan, R. Caruso and A. R. Von Neida, Bell System Tech. J., *59*, 593, 1980.
13. A. Steinemann and V. Zimmerli, Solid State Electr., *6*, 597, 1963.
14. S. Miyazawa, T. Mizutani and H. Yamazaki, Jpn. J. Appl. Phys., *21*, L 542, 1982.

Acta Physica Hungarica, 57 (3—4) pp. 271—283 (1985)

ELECTROCHEMICAL ETCHING

CHR. NANEV and K. DICHEVA

Institute of Physical Chemistry, Bulgarian Academy of Sciences
1040 Sofia, Bulgaria

Silver faces with only a few exactly known dislocation emergence points were etched electrochemically. Deep pyramidal etch pits were observed on the screw dislocations whereas flat-bottomed pits arise on the edge dislocations. Qualitatively, this fact corresponds to the results obtained with natural, i.e. as-grown from the vapour phase, {0001}-faces of zinc single crystals. They were etched electrochemically by a new "droplet" technique. Quantitatively, the formation of two-dimensional negative nuclei during the etching of the edge dislocation emergence points is understood as an atomistic process. Electrochemical etching of zinc was compared with chemical etching. It was concluded that SO_4'' selective adsorption (from $ZnSO_4$ aqueous solution) changes significantly the polar diagram of the lateral dissolution rate on the basal face of the zinc single crystals and that the electrochemical etching reveals better the actual structure of these crystals.

1. Introduction

Etching is the most accessible method for revealing defects emerging on (close packed) crystal faces. The electrochemical variant of chemical etching has some advantages. First, it is easy to dose and to measure exactly the etching parameters (being electrical quantities in this case). Second, in some cases it can be applied *in situ* during the electrocrystallization process. This fact reveals the unique possibility not only to control the defects (and their "destiny" [1]) but also to investigate their impact on the processes of crystal growth and dissolution in the most simple and immediate way. An example how electrochemical etching can contribute to the growing of more perfect crystals gives the capillary technique [2] developed by Kaischew, Budevski, Bostanov et al. Crystal faces free of screw dislocation emergence points were prepared here under special growth conditions, elaborated also by means of electrochemical etching control.

Because of its advantages, electrochemical etching is the subject of our lasting interest.

2. Theory

The most simple explanation for the preferred dissolution of the crystal at the screw dislocation emergence point during its selective etching relies on the presence of the well-known non-vanishing step created by this type of lattice imperfection. In contrast, at dislocations having pure edge character an immediate, i.e. simple

geometrical, reason is lacking. However, in practice, experimentally the edge dislocation emergence points are etched selectively and approximately as fast as the screw dislocations themselves.

Cabrera [3] was the first to reveal the proper physical meaning of edge dislocation etching. He considered it as a process of successive generation (and spread) of two-dimensional negative nuclei preferably created at the dislocation emergence points where, due to the released deformation energy, the process is favoured. In this way a greater final decrease in the free energy for the whole system is achieved there compared with the unconstrained regions. In his considerations Cabrera took into account only the elastic strain energy accumulated in the crystal lattice surrounding the dislocation. However Schaarwächter et al [4, 5] indicated that for metals, and especially for copper, the nuclei are so small that they have to confine practically only the dislocation core region; they drew the conclusion that one has to consider only the dislocation core energy instead of the elastic energy.

Our experiments show that the creation of clearly distinguishable etch pits on the edge dislocations emerging on the $\{0001\}$-faces of zinc single crystals required surprisingly high anodic pulses (amplitudes as high as 400 to 800 mV were necessary). In order to grasp this fact let us calculate now the radius R of the two-dimensional negative nucleus formed on the perfect crystal face under these conditions. Especially under strong undersaturations these nuclei arise randomly distributed on the perfect crystal face regions. They have to have a larger radius R as compared with radius R_d of the two-dimensional negative nuclei arising at the edge dislocation emergence points under the same conditions. The well-known Gibbs—Thomson's equation is applicable for the calculations

$$R=\frac{\varkappa s}{\Delta\mu'}, \tag{1}$$

where $\varkappa$ denotes the peripheral energy, s is the surface of 1 mol from the substance and $\Delta\mu'$ means the undersaturation. In the electrochemical case

$$\Delta\mu'=zF\eta_c, \tag{2}$$

where F is Faraday's equivalent, z is the number of electrons exchanged during the electrochemical reaction, η_c means the crystallization part of the applied overvoltage. One simple estimation taking into account these 400 to 800 mV shows the inapplicability of the classical nucleation theory. Its applicability suggests the formation of one nucleus consisting of at least a hundred atoms. This is necessary in order to define one real $\varkappa$. For example, for a hexagonal two-dimensional cluster consisting of 127 zinc atoms (i.e. with only 6 atoms in the edge row and radius of the inscribed circle 1.61 nm, taking $\varkappa=2$ to 4×10^{-11} J/m (2 to 4×10^{-6} erg/cm) like silver [2] and with $s=3.4\times10^{4}$ m^2 (3.4×10^{8} cm^2) one can easily calculate a crystallization overvoltage $\eta_c=3.5$ mV. Despite the fact that in the case under consideration we do not know exactly the percentage of the crystallization overvoltage in the total experimen-

tally applied overpotential this η_c-value is very far below our expectations based on the experimental data. Thus, our experiments strongly support Schaarwächter's conclusion concerning R_d. But his considerations are not free of shortcomings, too:

1. Like Cabrera, Schaarwächter again applies one approach based on the continuum approximation despite the fact that, especially here, for the limited number of atoms in the dislocation core, the atomistic treatment is obligatory.

2. He accepts one linear function for the dependence of the dislocation core energy vs radial distance, as measured from the dislocation centre, which is rotary-symmetrical, i.e. one evidently non-realistic assumption, especially for the edge dislocations characterized by their pronounced non-circular core (see e.g. [6]).

This is the reason to try one consistently atomistic treatment of the problem of edge dislocation etching. (Schaarwächter himself also understood the fact that the formation of two-dimensional negative nuclei confined in the dislocation core region is one atomistic process [5. p. 26]).

The atomistic treatment of crystal etching is based on Stranski—Kaischew's theory. It does not operate with any macroscopic quantity, e.g. such as peripheral energy, and thus suggests an unequivocal way for calculating the two-dimensional nucleation work. (This is the value determining the deepening rate during the etching of the edge dislocation emergence points). The work for the formation of a two-dimensional negative nucleus is given by the maximum change of Gibbs' thermodynamic potential ΔG vs the number n of atoms in the complex. In the case under consideration

$$\Delta G = \sum_{i=1}^{n} (\varphi_i - \varphi_0) - n\Delta\mu, \tag{3}$$

where φ_i is the separation work for the corresponding atom (for the sake of convenience the atoms are numbered here starting from the very central atom of the edge dislocation) and φ_0 is the separation work of the atom in the half crystal position, $\Delta\mu$ being the undersaturation per atom.

One has to keep in mind that the separation work φ_i for the detachment of one atom in a constrained position is decreased as compared with the separation work φ from an unconstrained lattice position by the value of its deformation energy u_i^{def}. Besides, we can accept that nearly one lattice bond less has to be destroyed for the detachment of the very central atom of the edge dislocation as well as for the following few atoms (until the fifth or sixth one, for which the decrease in lattice bonding may become double and so on; as a matter of fact the coefficient a_i (see Eq. (4)) expressing this relation may depend in a more complicated way on the kind of crystal lattice under consideration). Thus, the third term in Eq. (4) holds for the decreased number of lattice bonds and u_i^r is the relaxation energy:

$$\varphi_i = \varphi - u_i^{\mathrm{def}} - a_i\psi + u_i^r. \tag{4}$$

Here ψ denotes the work done for the separation of two next nearest neighbours in the crystal lattice.

On replacing (4) in (3) one obtains

$$\Delta G_i = i(\varphi - \varphi_0) - \sum_{i=1}^{n} (u_i^{\text{def}} + a_i\psi - u_i^{r}) - i\Delta\mu. \tag{5}$$

As a matter of fact, only the main energetic term $(\varphi - \varphi_0)$ is exactly known. The whole expression $\sum_{i=1}^{n} (u_i^{\text{def}} + a_i\psi - u_i^{r})$ has to be estimated under some reasonable assumptions. One possibility in this respect will be shown here:

If we consider the etching of one edge dislocation emergence point on the {0001}-face of a zinc single crystal we have to have $\varphi = 9\psi$, $\varphi_0 = 6\psi$ and hence $\varphi - \varphi_0 = 3\psi$. Reliable values for u_i^{def} by zinc are unfortunately lacking therefore a rough estimate will be given here on the basis of the calculations performed by Cotterill and Doyama [7]. Using central force interaction described by the Morse potential function these authors evaluate the deformation energy values for every individual atom of the edge dislocation 1/2⟨110⟩ {111} in the copper lattice. It can be estimated (see Fig. 11 of [7]) that

$$\sum_{i=1}^{n} u_i^{\text{def}} = 0.27\psi + 0.26\psi \ln i. \tag{6}$$

Due to the constrained lattice state the relaxation mode of the top-most crystal lattice plane is also altered. It is obvious that the lattice tends strongly to relax and to close the gap (it is especially pronounced after the detachment of the very central atom but in an ever decreasing way also after the separation of the following atoms from the dislocation centre). As far as our task is to create a hole there we have to counteract with energy u_i^r. Evidently u_1^r, i.e. the work done in order to prevent the disappearance of the single vacancy introduced (forced) in the dislocation centre, has the same value and the opposite sign of the relaxation energy which would be gained by closing the gap, i.e. approximately the energy for the creation of one elementary jog [8]

$$u_1^r \approx 0.1 G b^3, \tag{7}$$

where G is the shear modulus and b is Burgers vector of dislocation. With the values for $G_{\text{Zn}} = 3.4 \times 10^{10}$ N/m (3.4×10^{11} dyn/cm^2) and $b = 0.26$ nm (2.6×10^{-8} cm) one can easily calculate $u_1^r = 0.37$ eV, that is, $u_1^r \approx 1.6\psi$. For the next atoms u_i^r decreases quickly approaching the relaxation energy value of a (single) vacancy in the non-constrained crystal lattice (about 0.1 eV [9], that is, $\sim 0.4\psi$). As a first approximation let us think that u_i^r decreases linearly between $u_1^r = 1.6\psi$ and $u_{9-12}^r = 0.4\psi$. (These values are proposed as a basis for reflection, until better figures turn up.) The numerical data estimated as a result of the above considerations are presented in Table I.

The maximum values of ΔG_i, i.e. the works for the two-dimensional nucleation, were evaluated graphically by assigning definite values for $\Delta\mu$, expressed in ψ units (with $\psi = 0.23$ eV in the case under consideration). In this way one can determine the undersaturation intervals presented in Table II in which the corresponding critical complexes i^* can exist, i.e. clusters of vacancies, situated in the edge dislocation centres,

which serve as two-dimensional negative nuclei, are stable. In the fourth column of the same Table one can see the corresponding critical undersaturation values η_c (calculated according to Eq. (2)). They agree with our experimental results concerning the pulse amplitudes necessary to develop etch pits on the edge dislocations emerging on $\{0001\}$-faces of zinc single crystals (see para 3).

Table I

Energy sum values

$$\Sigma(\varphi_i - \varphi_0) = i(\varphi - \varphi_0 - \psi) - 0.27\psi - 0.26\psi \ln i + \sum_{i=1}^{n} u_i^r$$

i	1	2	3	4	5	6
$\Sigma(\varphi_i - \varphi_0)$	3.3ψ	6.6ψ	9.9ψ	13.0ψ	16.0ψ	18.9ψ

Table II

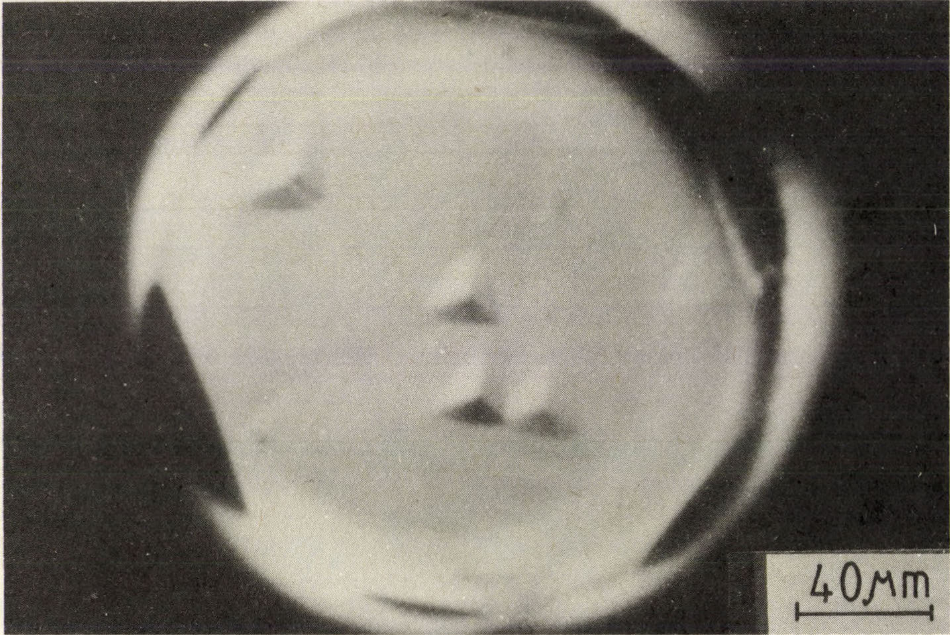
Undersaturation intervals in which complexes of i^* vacancies appear on the edge dislocation emergence point

i^* [number of atoms]	μ_{cr} [ψ]	$\Delta\mu_{cr}/kT$	η [mV]
3	$3.3 \div 3.1$	$29 \div 27.5$	$380 \div 357$
4	$3.1 \div 3.0$	$27.5 \div 26.7$	$357 \div 345$
5	$3.0 \div 2.9$	$26.7 \div 25.8$	$345 \div 334$

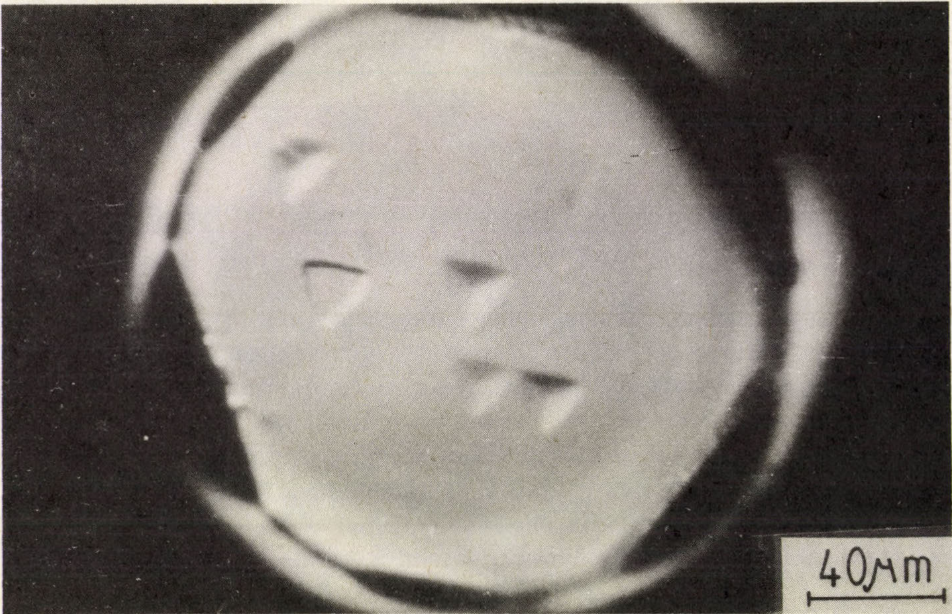
3. Electrochemical etching of separate faces of silver single crystals grown by capillary technique [10]

The capillary technique as developed by Kaischew et al [11] and Bostanov et al [12] consists in the forced growth of a silver single crystal in a glass capillary tube. By electrocrystallization under special conditions the front crystal face grows larger filling up the whole capillary cross section. This process is accompanied by the progressive diminishing of the number of screw dislocations emerging on the face. By slow growth and modulation of dc with an ac of much higher amplitude the appearance of new screw dislocations was avoided. The explanation is that as a result of the reverse the growing crystal face is cleaned from foreign particles (occasionally) embedded in the crystal lattice [10]. Since the existing screw dislocations lie inclined to the growth direction [1] the crystal face gradually becomes liberated from them.

Figure 1 shows an octahedral face of a silver single crystal prepared by the capillary technique. By means of a short cathodic pulse small growth hillocks are produced on the emergence points of the screw dislocations (Fig. 1a). This cathodic

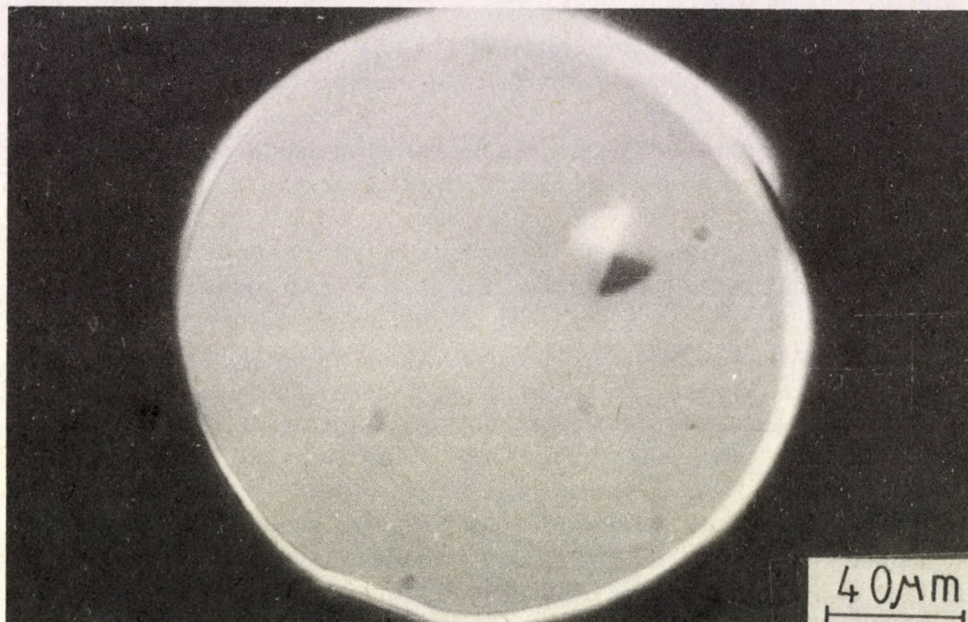


a)

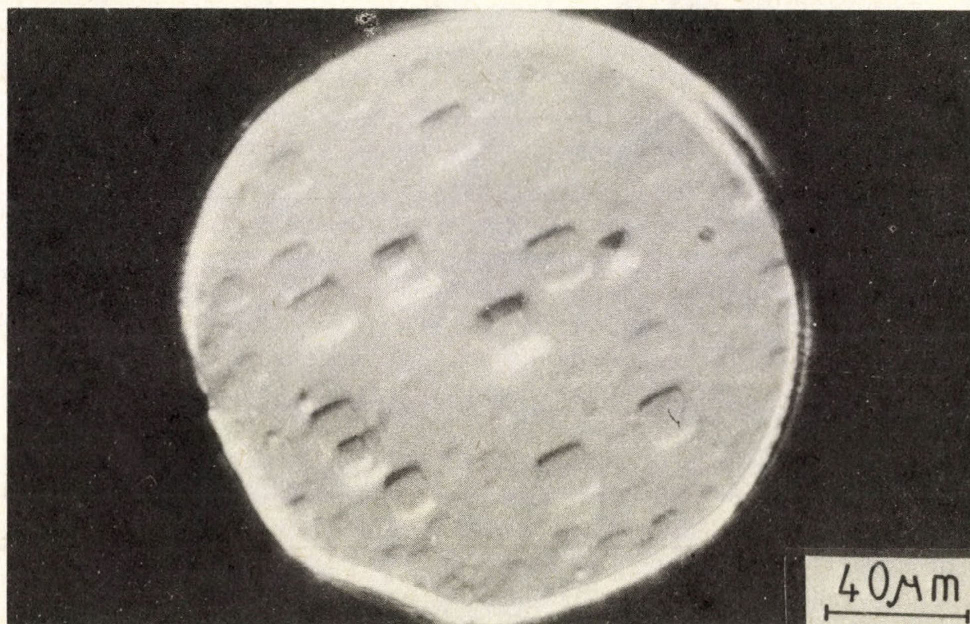


b)

Fig. 1. Octahedral face of a silver single crystal (microphotographs taken by means of a Nomarski interference contrast)
a) cathodic picture (pulse parameters 28 mV, 90 ms)
b) anodic picture (pulse 28 mV, 90 ms)



a)



b)

Fig. 2. Cubic face of a silver single crystal (same contrasting technique as Fig. 1)
a) cathodic pulse 30 mV, 40 ms
b) anodic pulse 30 mV, 40 ms. Generally, flat-bottomed etch pits are larger (for explanation see para 4)

picture has been compared with the anodic one. After an appropriate smoothing of the surface, accomplished by growing the crystal with very low current density, we apply an anodic pulse. It results in the creation of relatively deep pyramidal etch pits (Fig. 1b) at the screw dislocation emergence points (as marked by the small growth pyramids). In addition, one other defect is etched here. A flat-bottomed shallow etch pit is observed (Fig. 1b). By repeating the etching procedure several times, each time smoothing the surface, it was proven that this is also a line defect, i.e. probably one edge dislocation. Like the pyramidal etch pits on the screw dislocations also this etch pit appears at every run. (Occasionally other flat-bottomed etch pits appeared as well. Due to their statistically random character it was concluded that these reveal defects which do not deeply penetrate the crystal lattice, i.e. they are foreign particles embedded in it during the growth of the crystal for smoothing the face). If the face is free of screw dislocation emergence points no pyramidal, but eventually flat-bottomed, etch pits appear.

Exactly the same results were obtained with the cubic face of the silver single crystals (Fig. 2). Thus, these experiments are very instructive, but restricted only to qualitative and semi-quantitative explanations. The restriction is due to the fact that the periphery of the crystal face is also attacked during the anodic pulse (Fig. 1b, Fig. 2b). In order to avoid this "edge" effect and to perform quantitative measurements on the etching parameters we developed the so-called droplet technique.

4. Electrochemical etching of zinc [13]

The experiments described here were performed with "natural", i.e. as-grown from the vapour phase, {0001}-faces of zinc single crystals (Fig. 3) [14]. As compared with the usually used {0001} zinc cleavage planes [15] they are smoother. On the other hand, we were able to establish that most of the more mobile screw dislocations slide out of the crystal as a result of its cleavage.

The basic principle of the droplet technique can readily be seen in the scheme presented in Fig. 4. A small droplet of the electrolyte (1n $ZnSO_4$ aqueous solution, pH $= 1.8$ to 2.0) has been pressed to a desired point, sufficiently far from the edges of the relatively large basal face of the zinc single crystal. The experimental arrangement enables us to scan over the surface, i.e. to use one and the same crystal several times etching different regions on its basal face. Besides, we have the possibility to observe the etching process *in situ* (by means of an interference microscope). Etch pits were produced, only on the wetted portion of the basal face, by an applied anode pulse. Well-pronounced etch pits were observed as a result of single shots with the parameters: amplitude ranging from 400 to 800 mV and durations from 100 to 40 ms, respectively. No pits were observed by pulse amplitudes lower than 400 mV.

With zinc we have no possibility to reveal the screw dislocation emergence points via cathodic pulses, i.e. to create growth pyramids. Therefore we applied the classical double pulse potentiostatic technique in order to distinguish between screw and edge

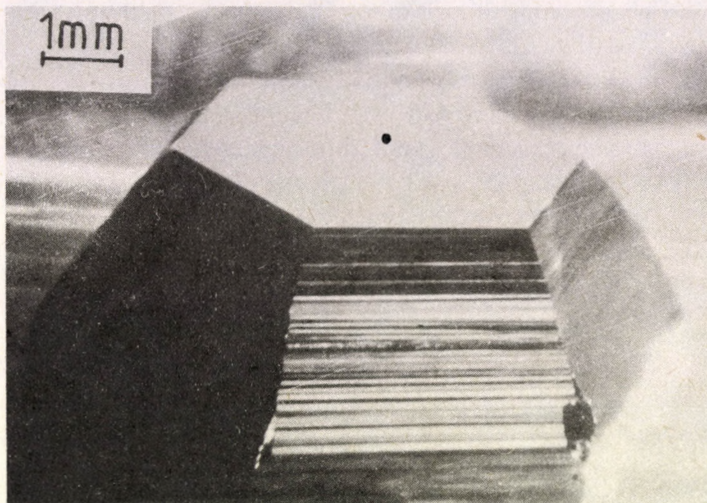


Fig. 3. Zinc single crystal grown from vapour phase

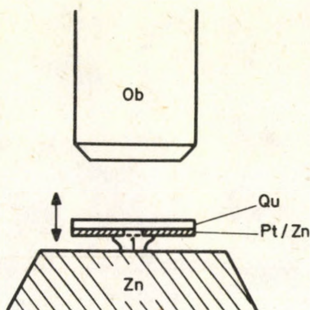


Fig. 4. Schematic representation of the principle of the droplet method (1 droplet, Zn zinc single crystal (anode); Ob — microscope objective, Pt/Zn — zinc-plated platinum plate cathode, Qu — quartz-glass plate)

dislocations by means of electrochemical etching. The anodic pulse was shaped in this case in the following manner (Fig. 5). The first pulse is of a high amplitude (400 mV) and short duration (30—40 ms). The second one is many times longer (2—3 s) but only 30—50 mV high. As shown by our preliminary experiments this value is one order of magnitude less than the undersaturation required for two-dimensional nucleation at the edge dislocation emergence points. In other words, the small hole created at the edge dislocation emergence point during the first (shorter) pulse only spreads to visible dimensions during the second one. In contrast, the screw dislocation emergence points dissolve without any barrier during the whole time. Really, there we observe pin-bottomed pyramidal etch pits (Fig. 6). Due to the presence of the non-vanishing step the deepening continues here with nearly the same intensity also during the lower second

pulse. If this self-perpetuating step disappears for some reason the pit takes the form of a truncated pyramid. In several cases we observed such pits (Fig. 7). Flat-bottomed pits were created by means of the double-pulse technique much more frequently. It is worth noticing that, as a rule, they are considerably larger compared with the pyramidal etch pits (on the same scale of magnification). In complete accordance with our previous considerations this fact can be attributed to the different dissolution mechanisms (compare also Fig. 2b). Thus, as with silver, the flat-bottomed pits are obviously due to the edge dislocations or/and foreign particles embedded in the crystal lattice.

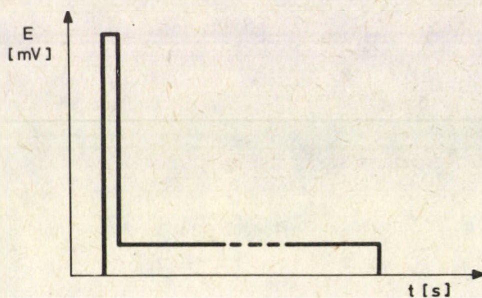


Fig. 5. Double potentiostatic pulse shape (for parameters see text)

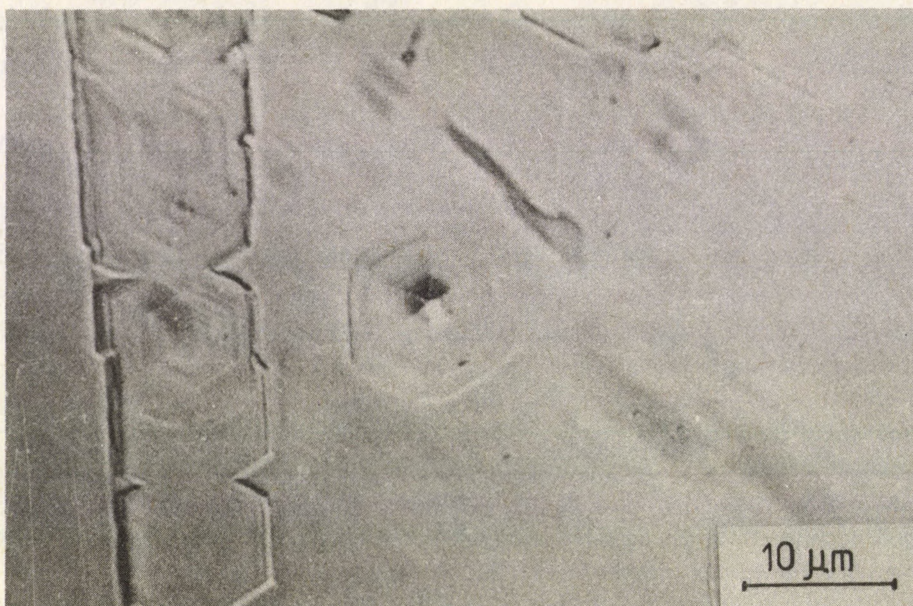


Fig. 6. Pyramidal etch pit produced on basal face of zinc single crystal by means of a double pulse (interference contrast by differential shearing; microscope objective 50×)

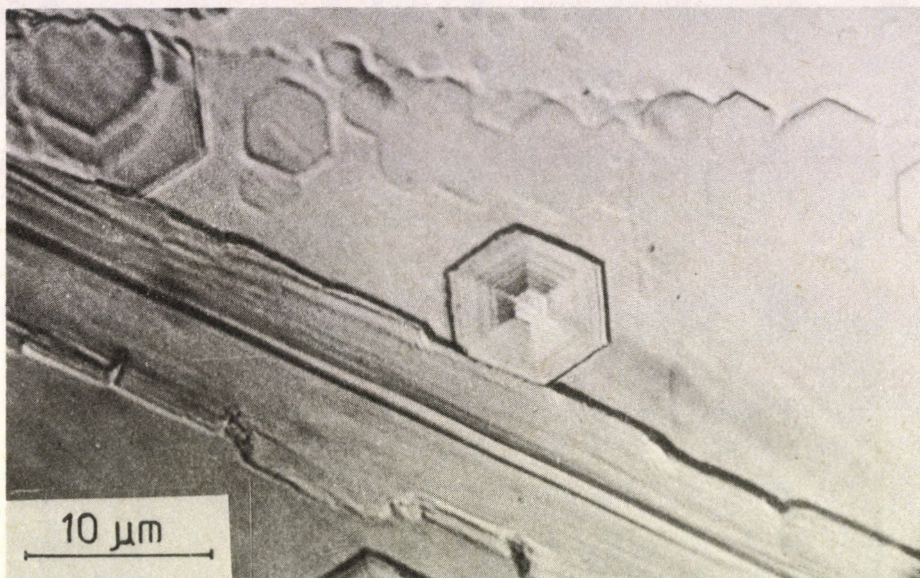


Fig. 7. Truncated pyramid resulting from a double pulse (differential interference contrast, objective 50 ×)

5. Comparison with the chemical etching of zinc [13]

Electrochemical etching was compared with the classical chemical one. As a chemical etchant we used an alcoholic solution of hydrochloric acid [16] (in our experience this is the best selective etchant for zinc). Two kinds of hexagonal etch pits (relatively deep pyramidal forms and shallow flat-bottomed pits) are observed also by chemical etching (Fig. 8) but the similarity finishes here. The first difference was found by the azimuthal orientation of the etch pits. Their edges run in the $\langle 10\bar{1}0\rangle$ crystallographic direction with chemical etching and in $\langle 11\bar{2}0\rangle$ by electrochemical etching. It was by no means easy to reveal the cause of this phenomenon. Adding the components of both solutions to each other we were able to find that the SO_4''-anions are the species which are predominantly adsorbed on the steps of the basal face during the etching. Thus, SO_4''-adsorption changes the polar diagram of the azimuthal dissolution rate on the basal face of the zinc single crystals and causes the $\langle 11\bar{2}0\rangle$-orientation of the etch pits.

The second difference is perhaps more important. By means of X-ray diffraction topography it was found [17] that our zinc single crystals contain screw dislocations with Burgers vector $1/3\langle 11\bar{2}3\rangle$ (besides those with Burgers vector $\langle 0001\rangle$). But as a result of the chemical etching we observed only etch pits in the form of regular pyramids, i.e. due to $\langle 0001\rangle$-dislocation etching. In contrast, from electrochemical

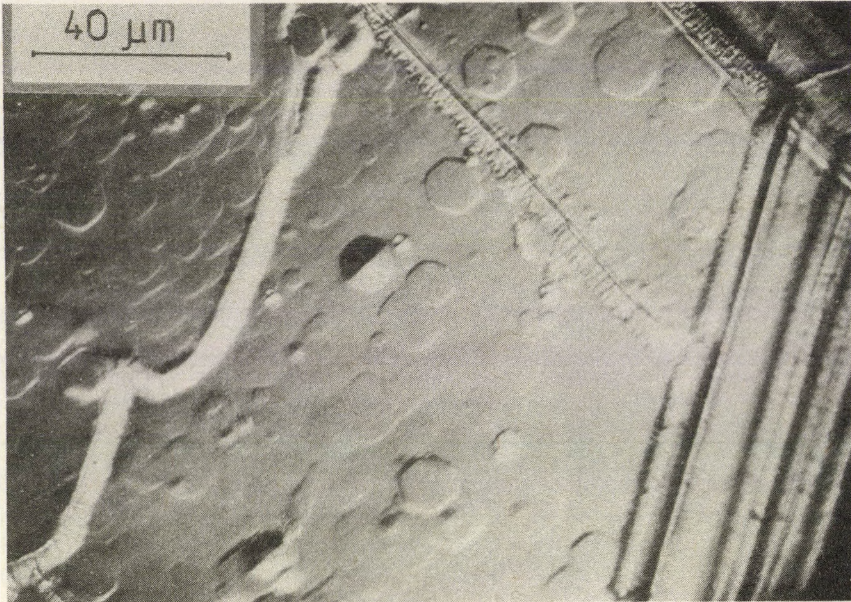


Fig. 8. Edges of chemical etch pits are perpendicular to edge of basal face of zinc single crystal

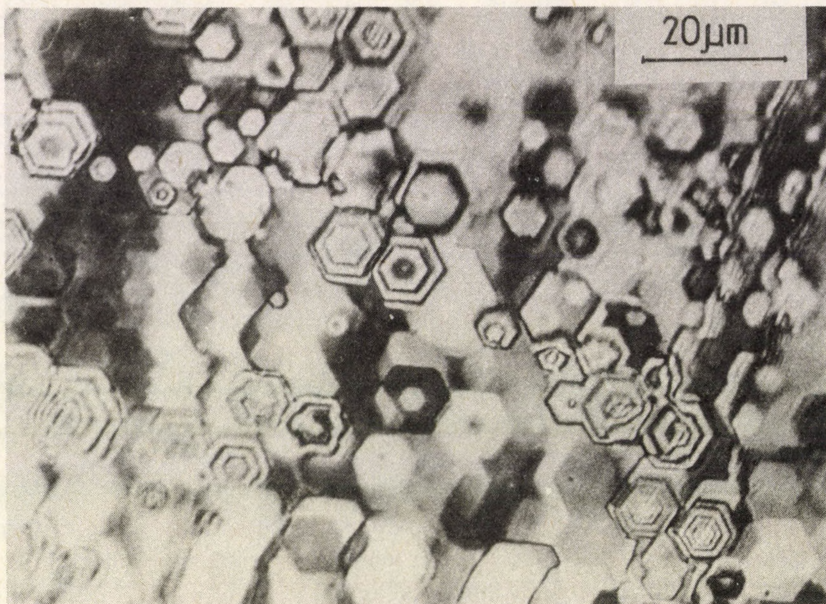


Fig. 9. Symmetric and non-symmetric (at top left) etch pits. Shearing method, microscope objective 50×

etching, pits in the form of inclined pyramids were observed additionally. It was measured interferometrically (Fig. 9) that the dissolution direction in this case is very close to the $\langle 11\bar{2}3\rangle$ one. Thus, our conclusion is that electrochemical etching reveals better the actual structure of (the basal face of) zinc single crystals.

References

1. V. Bostanov, Chr. Nanev, R. Kaischew, E. Budevski, Kristall und Technik, *2*, 319, 1967.
2. R. Kaischew, E. Budevski, Contemp. Physics, *8*, 479, 1967.
3. N. Cabrera, J. Chem. Phys., *53*, 675, 1956.
4. W. Schaarwächter, Phys. Stat. Sol., *12*, 375, 1965.
5. W. Schaarwächter, L. Jasper, K. Lücke, Der Einfluß der Versetzungsstruktur auf die Kristallauflösung, Westdeutscher Verlag, Köln and Opladen, 1967.
6. J. P. Hirth and J. Lothe, Theory of Dislocations, McGraw-Hill Book Company, New York, 1968.
7. R. M. J. Cotterill and M. Doyama, Phys. Rev., *145*, 465, 1966.
8. D. Hull, Introduction to Dislocations, Pergamon Press, Oxford, 1965.
9. H. G. Van Bueren, Imperfections in Crystals (Chapter 2, Section 2.1), North-Holland, Amsterdam, 1960.
10. Chr. Nanev, El. Acta, *25*, 1383, 1980.
11. R. Kaischew, G. Bliznakow, A. Scheludko, Bull. Acad. Bulg. Sci. Série Phys., *1*, 147, 1950.
12. V. Bostanov, A. Kotzeva and E. Budevski, Bull. Inst. Chim. Phys. Acad. Bulg. Sci., *6*, 33, 1967.
13. Chr. Nanev and K. Dicheva, El. Acta, *28*, 933, 1983.
14. R. Kaischew and Chr. Nanev, Phys. Stat. Sol., *10*, 779, 1965.
15. M. B. Ives and D. D. McAusland, Techn. Rep. No. 17, Contr. No. 3925(00), 1968.
16. H. S. Rosenbaum and M. M. Saffren, J. Appl. Phys., *32*, 1866, 1961.
17. Chr. Nanev and M. Wilkens, Phys. Stat. Sol., *23*, 663, 1967.

Acta Physica Hungarica, 57 (3—4) pp. 285—293 (1985)

INVESTIGATION OF MORPHOLOGY AND SURFACE STRUCTURE OF SMALL NaCl PARTICLES

M. RÖSLER

Institute of Solid State Physics and Electron Microscopy
Academy of Sciences of GDR, DDR-4020 Halle (Saale) GDR

The morphology and the surface structure of small NaCl particles, produced by gas-evaporation-technique (GET), were investigated by means of the decoration method. It was shown how the size and the morphology of the particles from the inner zone of the convection flow depend on the temperature of the evaporation source and the gas pressure. Details of the surface structure of spherical, facetted spherical and cubic particles are correlated with the experimental conditions of particle production. The first stage of sintering of cubic particles was detected from the step structure around the contact region. Evaporation structures on the particle surface were produced and described. The advantages of the decoration method in comparison with direct transmission of the particles were shown also for other NaCl aerosol generation methods.

1. Introduction

In the past years the properties of small particles have been investigated by means of various electron-microscopical methods. Thus, Uyeda and co-workers [1] examined the crystallography and the habit of diverse metal aerosol particles with the help of transmission electron microscopy (TEM). Particles of many substances, however, are not suitable objects for direct TEM investigation because of their high sensitivity to electron radiation damage. In their TEM investigations of the morphology and the size distribution of ionic crystals Washida and Kobayashi [2] as well as Podzimek [3] imaged only the profiles of the particles. In these studies certain changes of the profiles were observed which are due to radiation damage inside the electron microscope. In such cases a replica technique in connection with the decoration method appears to be a more appropriate means of obtaining information on the surface structures, the morphology and the size distribution of the particles.

2. Experimental procedure

Small particles of NaCl were produced by the gas-evaporation-technique (GET) [1, 4].

NaCl was selected as a model substance for aerosol formation in laminar buoyant-flows for the following reason: The small absorption of IR-radiation by NaCl prevents the particles from being directly heated by thermal radiation from the

evaporation source. Moreover, NaCl is not subjected to chemical reactions with most of the usual carrier gases even if they contain some impurities of oxygen. Later, the NaCl surfaces can be investigated in molecular dimensions by means of the decoration technique.

The experiments were carried out in a bell jar of a conventional vacuum system (Fig. 1). Via a dose valve (1) a certain amount of nitrogen (99.99 vol %) or helium (99.99 vol %) was let into the chamber after it had been evacuated. The actual gas pressure was measured by a membrane vacuum meter (2).

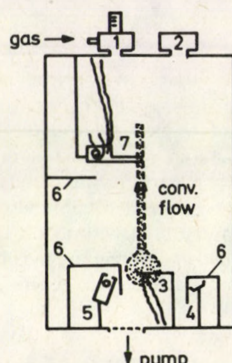


Fig. 1. Work section for producing NaCl aerosols by GET (1: valve, 2: membrane vacuum meter, 3: NaCl evaporation source, 4: Au evaporation source, 5: Pt/C source, 6: shutters, 7: NaCl finger for particle collection)

The NaCl evaporation source was a standard molybdenum boat (3), which was fixed a few cm above the bottom of the work section of the vacuum chamber. All the other components such as the gold evaporation source for the decoration technique (4) and the platinum/carbon source for the shadow-casting technique (5) were fixed below the NaCl evaporation source and were covered by shutters (6) during the aerosol formation.

The aerosol particles were collected from the inner zone of the laminar convection flow at a certain distance above the source by their settlement on narrow, cold NaCl fingers (7). These fingers were fixed with tungsten spirals, which made it possible to heat the substrate and the settled aerosol particles to the temperature necessary for decoration after evacuating the work section. In order to catch particles from the inner zone of the flow these fingers were movable in a plane perpendicular to the flow direction.

The temperatures of the NaCl evaporation source and of the NaCl finger were measured by a Pt/PtRh-thermocouple and a Ni/NiCr-thermocouple directly at the source and in the finger, respectively.

The surface of the NaCl finger together with the settled aerosol particles was investigated in TEM by means of a standard replication. It is important that the

particle concentration on the specimen surface should not be too high, otherwise the replica will break into small pieces during the preparation, or due to the repeated superposition of particle covers it will be impossible to find out the structure of individual particles.

3. Results

In electron microscopy appropriate characteristics of the aerosol particles are the mean size (size distribution), the morphology and the surface structure. The dependence of these characteristics on the experimental conditions can yield information about the generation mechanism of the aerosol particles in the convection flow.

3.1. Dependence on the evaporation source temperature

In these experiments the temperature of the evaporation source T_q was varied from 1000 K to 1373 K. Below 1000 K the convection flow was not visible and therefore it was not sure that the flow was laminar. Above 1373 K, evaporation proceeded so rapidly that the evaporation substance was consumed before the flow became stationary. The pressure of the helium carrier gas was $2.65 \cdot 10^4$ Pa in all experiments.

In the whole temperature interval small cubic particles occur the mean size of which lies between 40 and 180 nm. With increasing source temperature the mean size of these cubic particles increases only slowly. At source temperatures above 1150 K there appear additional spherical and facetted spherical particles. The mean diameter of these spherical particles increases with increasing source temperature and reaches a value of 500 nm at 1373 K.

3.2. Dependence on carrier gas pressure

In order to ascertain the role of the gas pressure in the particle formation process the He pressure is varied from $1.33 \cdot 10^2$ to $2.65 \cdot 10^4$ Pa at a source temperature of 1183 K. At low pressures the measurement of the mean size becomes inaccurate because of the frequent superposition of particles and their tendency to build up network-like structures. Therefore instead of the exact mean values Fig. 2 shows the size intervals in which particles were observed at the respective pressure. Below $1.33 \cdot 10^3$ Pa only small, strongly agglomerated cubic particles occur. Between $1.33 \cdot 10^3$ Pa and $6.7 \cdot 10^3$ Pa spherical particles are additionally formed, the size interval of which increases with increasing gas pressure. Above $6.7 \cdot 10^3$ Pa the size interval of the particles formed does not change markedly. The diameter of the biggest particles is about 280 nm.

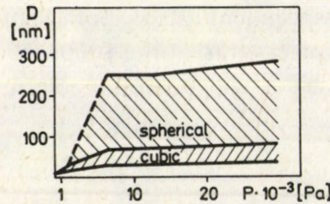


Fig. 2. Size intervals of the particle diameters D depending on He gas pressure P, $T_q = 1183$ K, distance between evaporation source and collector: 8.5 cm

3.3. Surface structure of aerosols

The fine structure of the NaCl particles was observed by means of surface decoration. The preferred objects of investigation are cubic particles because the surface steps on a flat surface can be resolved better than those on rounded surfaces, which are characterized by a high density of steps.

Figure 3a shows a rectangular particle, which lies on the substrate. The substrates (NaCl fingers) were produced by cleavage in air, resulting in a much higher gold nucleus density on the substrate surface than on the fresh aerosol surface. The aerosol particles are visible due to the small density of gold nuclei, the absence of dissolution structures and the intensive contrast at the edges of the particle. Figure 3b and 3c shows groups of particles, the individuals of which collided during the transport to the substrate. They are oriented with their $\{100\}$ planes parallel to each other. The contact region represents mostly a high step (indicated by arrows). In the illustrations high steps of this origin can be seen to split up by emission of a system of lower steps. This is the first stage of the sintering process of the particles. Since the particles are transported with the convection flow into regions with lower temperatures these step systems around the contact zone are frozen-in and can be decorated.

Figures 4a and 4b represent the rare case of sintering of nonisometric particles. The process of sintering starts with the formation of feet, which consist of narrow piled-up steps whose height is approximately 6 nm (indicated by arrows).

Investigation of the surface structure of spherical NaCl particles shows that there are two types of morphologies: the nearly perfect sphere (Fig. 5a), which has flat regions only in a small area around the $\langle 100 \rangle$ directions, and the facetted sphere (Fig. 5b), which has extensive octagonal $\{100\}$ planes and rough $\{110\}$ and $\{111\}$ planes. The irregularities in Fig. 5b are due to deformation of the replica film. Particles with a diameter smaller than 300 nm mostly show only the $\{100\}$ planes and bands connecting the $\{100\}$ planes in $\langle 100 \rangle$ directions.

At sizes of spherical particles smaller than approximately 100 nm one can often observe only disc-shaped $\{100\}$ planes, because due to the high curvature in other regions the decoration technique cannot resolve further step structures. Figure 5c shows a particle with a diameter of 90 nm where the diameter of the pole plateau

(indicated by circles) is 40 nm. Besides the regular particles described above, twinned NaCl cubes occur. Analysis of angular projections in the image plane shows that these are {111} twins. All of these twinned particles were found in the size interval of 80 to 160 nm edge length.

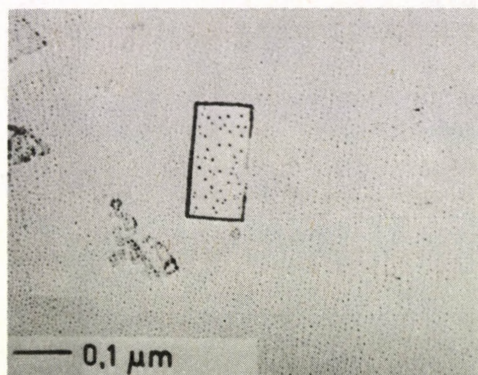
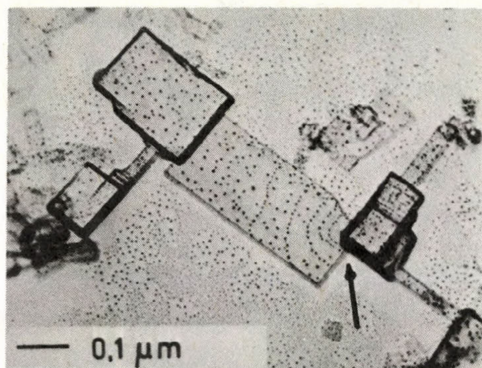
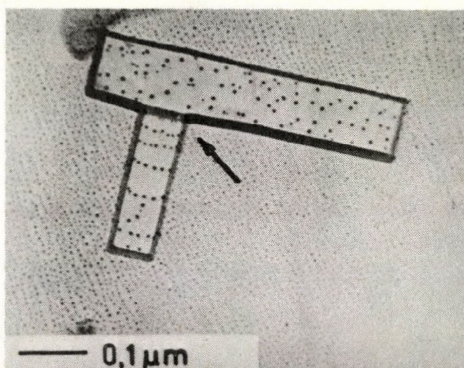


Fig. 3a. Rectangular NaCl particle

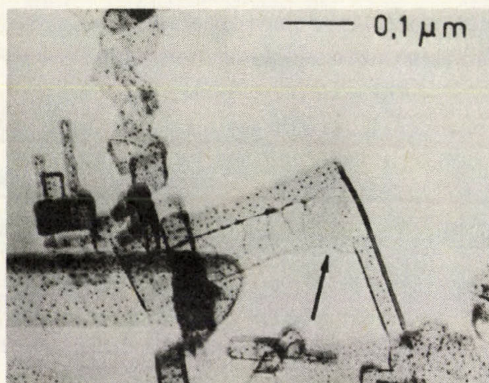


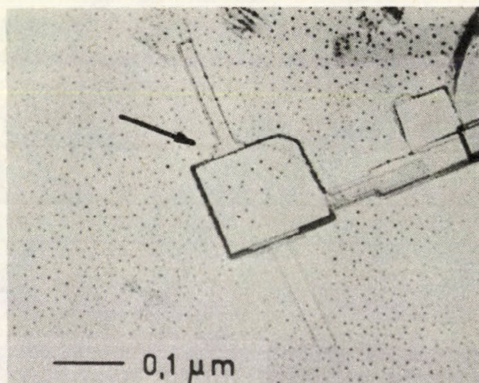


Figs 3b and 3c. Decorated steps around the contact region of collided aerosol particles

3.4. Evaporation structures of NaCl aerosols

In order to compare the surface structure of as-grown particles produced by GET with defined evaporation structures, NaCl particles were heated for 15 min at a temperature of 360 °C after they had settled on the substrate. Because of the different heat contact with the substrate they had different temperatures. The network-like structure of the particle ensemble guarantees that the substrate does not influence the particle surface structure.





Figs 4a and 4b. Formation of feet during the process of sintering of nonisometric particles

The main features of the evaporation structures are: the formation of lochkeims on the {100} planes, the existence of step ensembles over the whole particle surface (Fig. 6) as well as sometimes the appearance of {110} and {111} planes. Such structures have never been found on particles which are not especially heated, neither in laminar nor in turbulent flows.

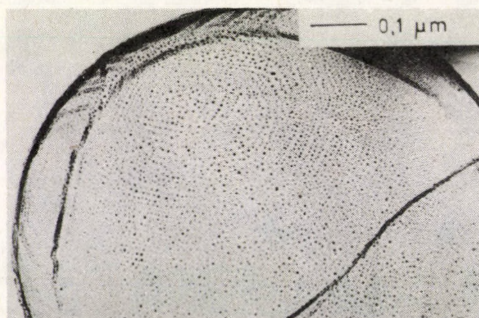


Fig. 5a. The nearly perfect sphere

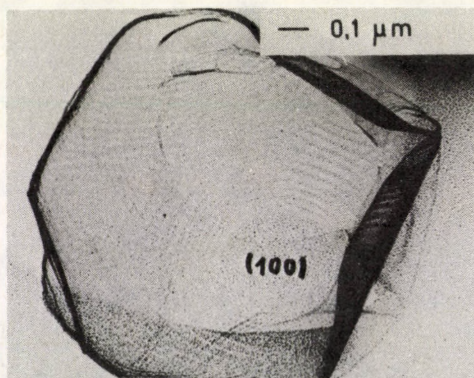


Fig. 5b. The facetted sphere with {100], {110] and {111] regions

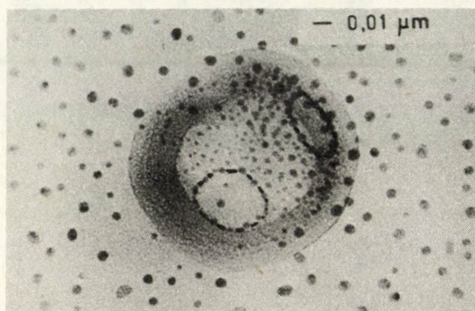


Fig. 5c. Small facetted sphere with circular {100} planes

Fig. 6. Evaporation structure: Steps on the whole particle surface

4. Discussion and conclusions

The dependence of particle sizes on source temperature and carrier gas pressure is qualitatively in agreement with the results of Washida and Kobayashi [2] for KCl. But the use of the decoration technique besides the conventional TEM yields additional information about the change of the growth morphology with increasing source temperatures. In the inner zone, there is a transition from the cube over the cube with steps along all the edges, and the facetted sphere to the nearly perfect sphere. Experimental investigations of the temperature profile along the flow direction and calculations on the actual partial vapour pressure in laminar flows have shown [5] that three typical situations may arise:

i) At high source temperatures nucleation, growth and coalescence appear in the liquid phase. After solidification of the particles there is no molecular material in the flow for further growth. This leads to spherical particles (see Fig. 5a).

ii) At medium source temperatures nucleation and growth also appear in the liquid phase, but after solidification there is enough molecular material for further growth. The particles have facetted forms (see Fig. 5b).

iii) At low source temperatures nucleation, growth and coalescence will appear in the solid state. The cubic form as well as various compound forms will arise (see Fig. 3a, 3b and 3c).

No evaporation structures were found on the surface of as-grown NaCl aerosols. The fall-out of particles back into the hot region of the convection flow can therefore be

excluded. Moreover, the re-evaporation, which was proposed by Thölen [6] for metal aerosols in the inner zone of the flow cannot be confirmed for NaCl particles.

It was proved experimentally (see Appendix) that the decoration method is useful also for other techniques of NaCl particle production and yields far more information on morphology and the structure of growth, coalescence and dissolution than the usual TEM.

Appendix

By means of a Rapaport—Weinstock generator and a heat drift tube NaCl particles were produced from a saturated solution of NaCl in water. The use of the decoration method allows one to distinguish between coalescence structures of cubic particles which form rounded compound particles (Fig. 7a), the dissolution structure of cubic particles (Fig. 7b) and really rounded single particles (Fig. 7c).

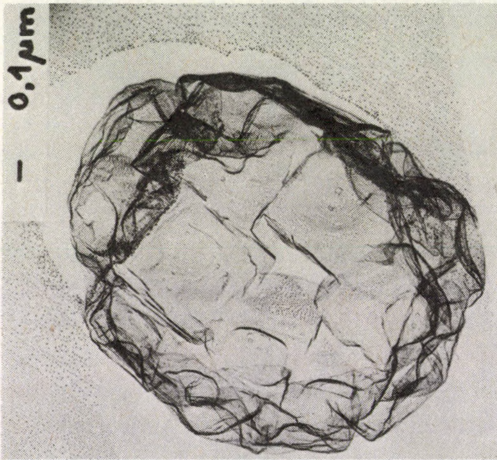


Fig. 7a. NaCl compound particle produced by the Rapaport—Weinstock generator

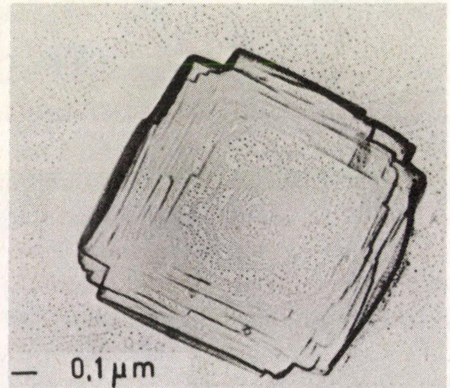


Fig. 7b. NaCl aerosol particle with dissolution structures

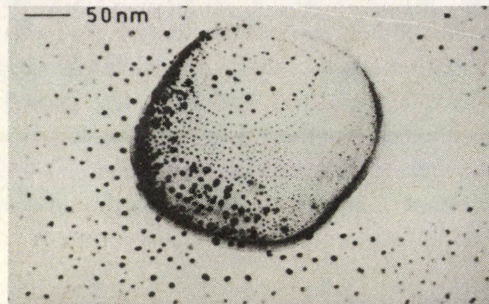


Fig. 7c. Small rounded aerosol particle

In such a way it is possible to analyse natural giant salt aerosols with respect to formation by coalescence as well as natural aerosols with a size below 0.3 μm for which it is difficult to work with aerosol size spectrometers.

Acknowledgements

The author would like to express his sincere thanks to Prof. H. Bethge and Dr. M. Krohn for their helpful discussions and critical reading of the manuscript.

References

1. R. Uyeda, J. Cryst. Growth, *24/25*, 69, 1974.
2. H. Washida and K. Kobayashi, Jap. J. Appl. Phys., Suppl. 2, Pt. 1, 1974.
3. J. Podzimek, J. Rech. Atmos., *11*, 227, 1977.
4. W. Romanowski, Hochdisperse Metalle, Akademie-Verlag, Berlin, 1982.
5. M. Rösler, Thesis, Halle 1985.
6. A. R. Thölen, Acta Metall., *27*, 1765, 1979.

Acta Physica Hungarica, 57 (3—4) pp. 295—301 (1985)

THE GROWTH OF TeO_2 CRYSTALS FROM THE BOILING SOLUTION

L. ÖTSI, E. HARTMANN and B. VAJNA

Research Laboratory for Crystal Physics
Hungarian Academy of Sciences
1112 Budapest, Hungary

A closed circulating system is described in which the circulating solvent transported the material from the nutrient to the seed crystals growing by evaporation at the boiling point. By means of this technique tetragonal paratellurite crystals have been grown from nitric acid. The morphology and kinetics of paratellurite crystals have been investigated.

K. Nassau [1] described eleven years ago a novel technique by which iodate crystals had been grown from solutions at their boiling point. Recently Rodriguez-Clemente et al [2] applied this technique to the growth of KH_2PO_4, NaCl and KCl crystals.

The single crystals of paratellurite have excellent acoustooptical properties due to their low ultrasound velocity. The single crystals of paratellurite are usually grown from the melt by the Czochralski method [3]. There were attempts to grow these crystals from aqueous solutions of hydrochloric acid which is an effective solvent for growing paratellurite single crystals under hydrothermal conditions [4, 5].

We tried to grow paratellurite crystals from aqueous solution of nitric acid at normal pressure at the boiling point. One of the main parts of the growing apparatus was a modified Thielepape extractor (Fig. 1). Refluxing solvent from the condenser passed through the extraction thimble containing pieces of TeO_2. The resaturated solution dripped again into the boiling solution. In this way a closed circulating system has been attained in which the circulating solvent transported the material from the nutrient to the seed crystals growing by evaporation at the boiling point. The circulating system was necessary because the solubility of TeO_2 in aqueous solutions of nitric acid is rather low. The solution in the spherical flask was heated with a two-zone resistance furnace controlled by a variable autotransformer. The temperature of the boiling solution was 120 °C. The reflux of the resaturated solution was 5—6 drops/min. The seed crystals of different orientation were cut out of paratellurite crystals grown from the melt by the Czochralski method. The seed crystals were mounted on platinum wires. The growth period lasted 2—3 weeks and the crystals attained dimensions of 15—20 mm.

It was questionable whether the crystals grown from the solution have the tetragonal paratellurite (SG: $P4_12_12$) or the orthorhombic tellurite (SG: Pabc) structure [6]. To decide this, Bragg angles of the crystal faces were searched for, back

reflection Laue photographs were taken from several directions and a powder photograph was compared with that of a Czochralski-grown paratellurite crystal. The Bragg angles of the crystal faces could have been assigned to paratellurite as well as tellurite, however, one of the Laue patterns showed a 4-fold rotation axis (Fig. 2) which

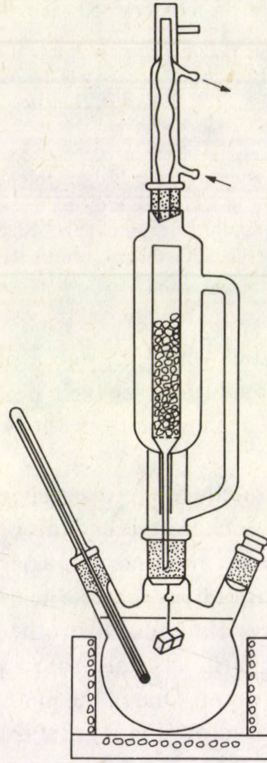


Fig. 1. Thielepape extractor

means that the point symmetry of the structure cannot be the point group mmm of tellurite. Two other Laue patterns taken from directions having an angle of 45° between them and being orthogonal to the 4-fold axis showed 2-fold rotation axes. With these symmetries a structure has to have tetragonal Bravais lattice and may only be described by one of the following point groups: 422, 4/m2/m2/m, 4mm and $\bar{4}$2m. The parameters of lattice calculated from the powder photograph are $a=0.4811$ nm, $c=0.7623$ nm being in good agreement with the literature data ($a=0.48122$ nm, $c=0.76157$ nm). Comparing the powder pattern of a solution grown crystal with that of paratellurite (Fig. 3) a perfect agreement is found.

On the basis of these evidences the structure of our solution grown crystals has been determined as paratellurite. Our paratellurite crystals were usually bounded by tetragonal dipyramids and tetragonal prism faces. The habit of the crystals is presented

in Fig. 4, where d denotes the $\{101\}$, m the $\{110\}$ and n the $\{112\}$ faces. The growth rates of the faces were:

$$v_{110}=0.89\times10^{-2}\ \mathrm{mmh}^{-1},$$

$$v_{112}=0.47\times10^{-2}\ \mathrm{mmh}^{-1},$$

$$v_{101}=0.40\times10^{-2}\ \mathrm{mmh}^{-1}.$$

On the $\{101\}$ faces growth layers (Fig. 5) while on the $\{112\}$ faces growth hillocks (Fig. 6) could be observed.

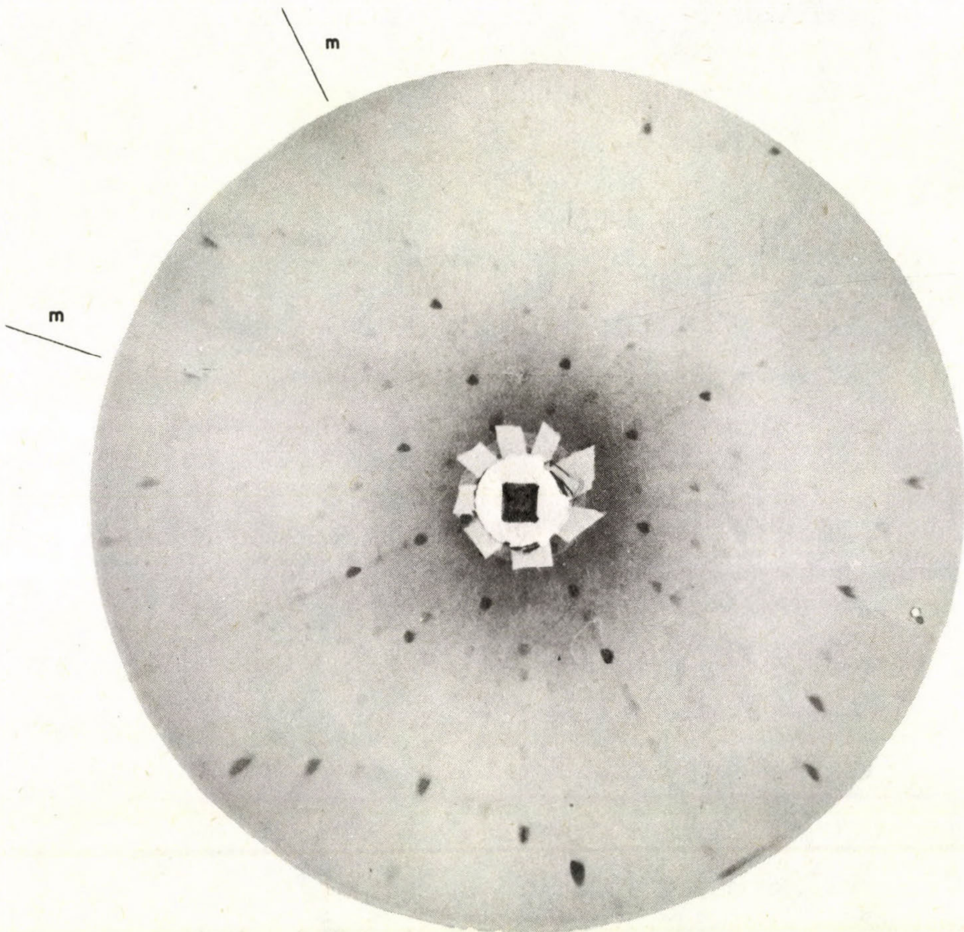


Fig. 2. The back reflexion Laue photograph of a TeO_2 crystal grown from solution. The 4-fold symmetry axis shows that the crystal belongs to the point group 422 of paratellurite and not the point group mmm of tellurite

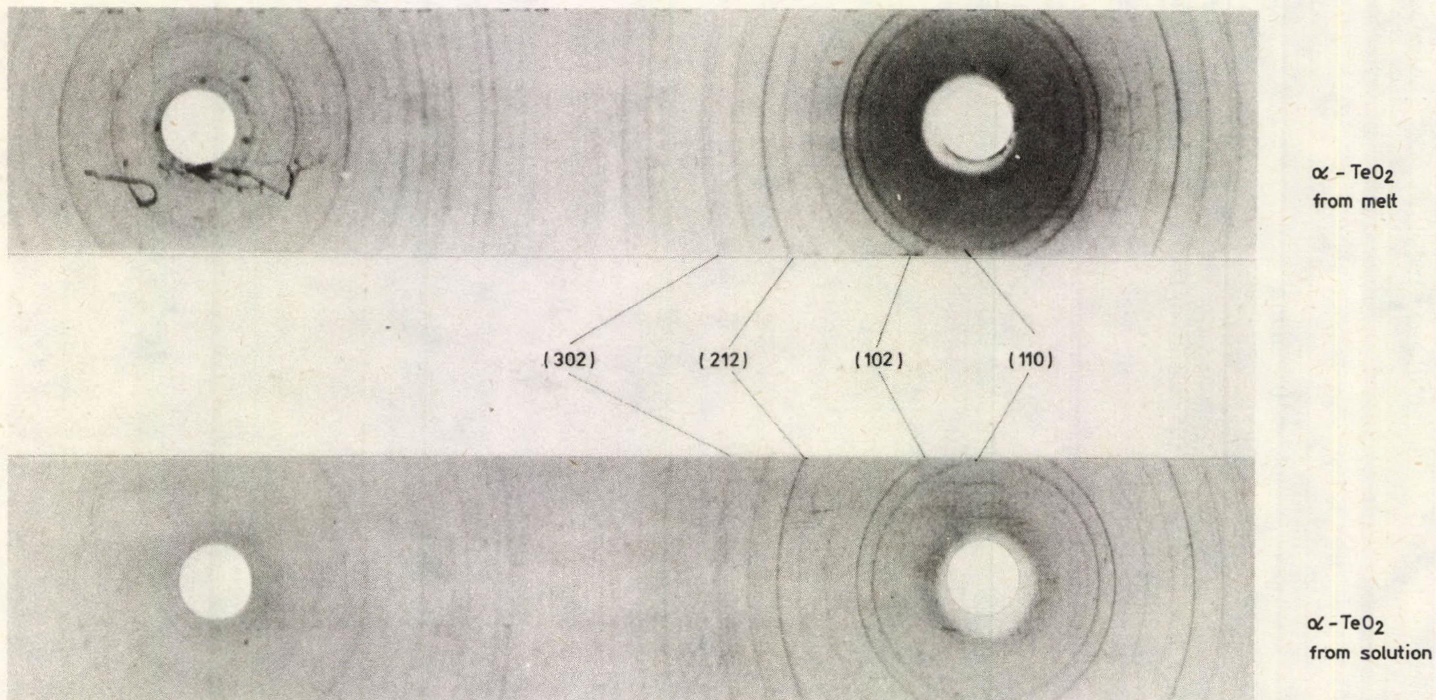


Fig. 3. Powder photographs of paratellurite grown from melt by Czochralski method and TeO_2 grown from solution. Their similarity proves the identity of the two phases

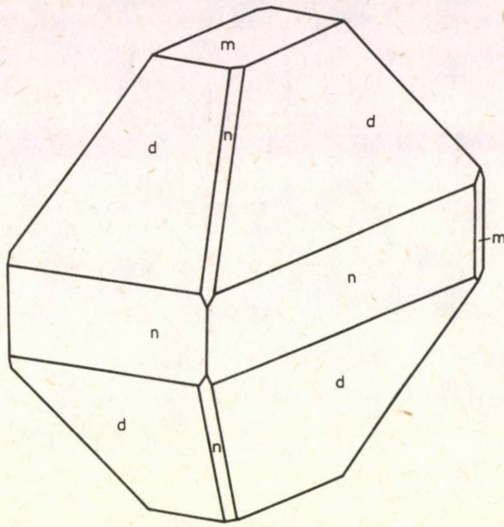


Fig. 4. The habit of TeO_2 crystal grown from solution

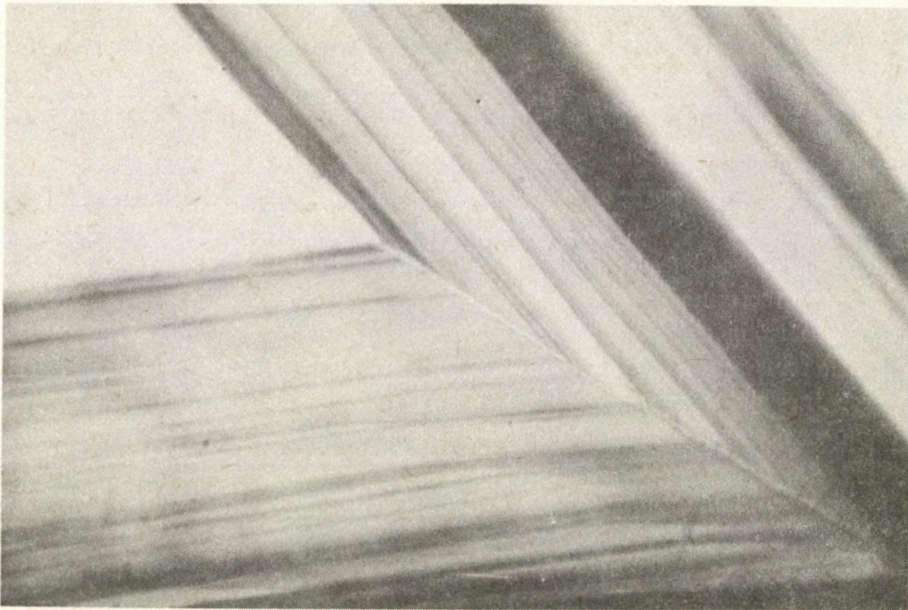


Fig. 5. Growth layers on the (101) face. Magnification: 20 ×

The transparence of TeO_2 layer grown from solution on seed crystal was quite good (Fig. 7).

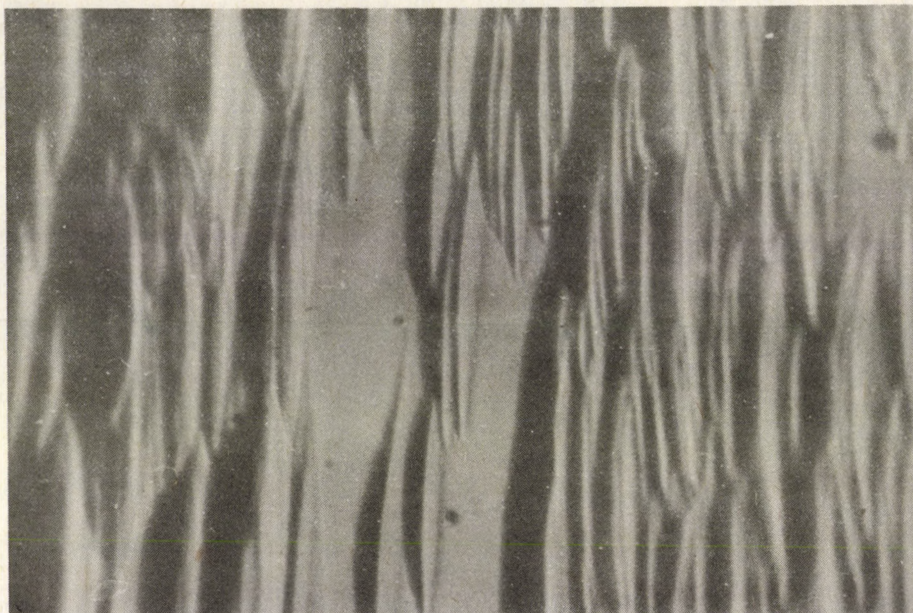


Fig. 6. Growth hillocks on the (112) face. Magnification: 10 ×

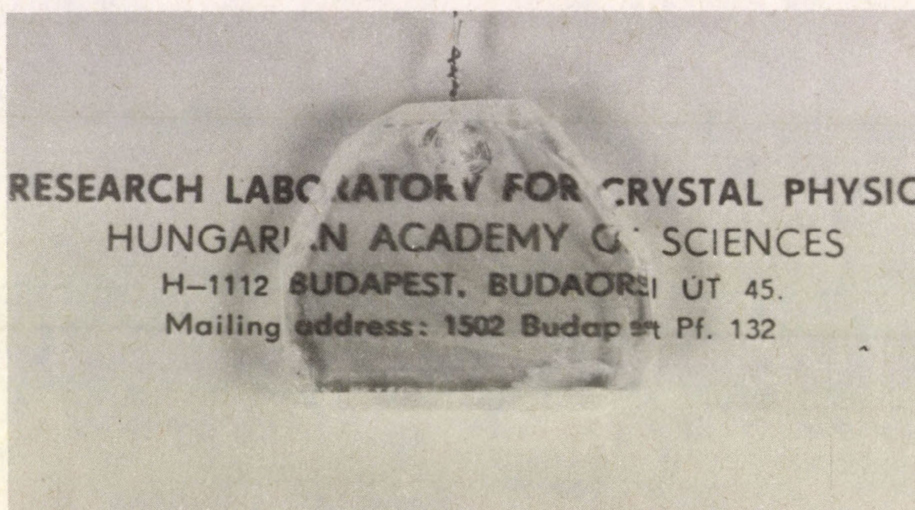
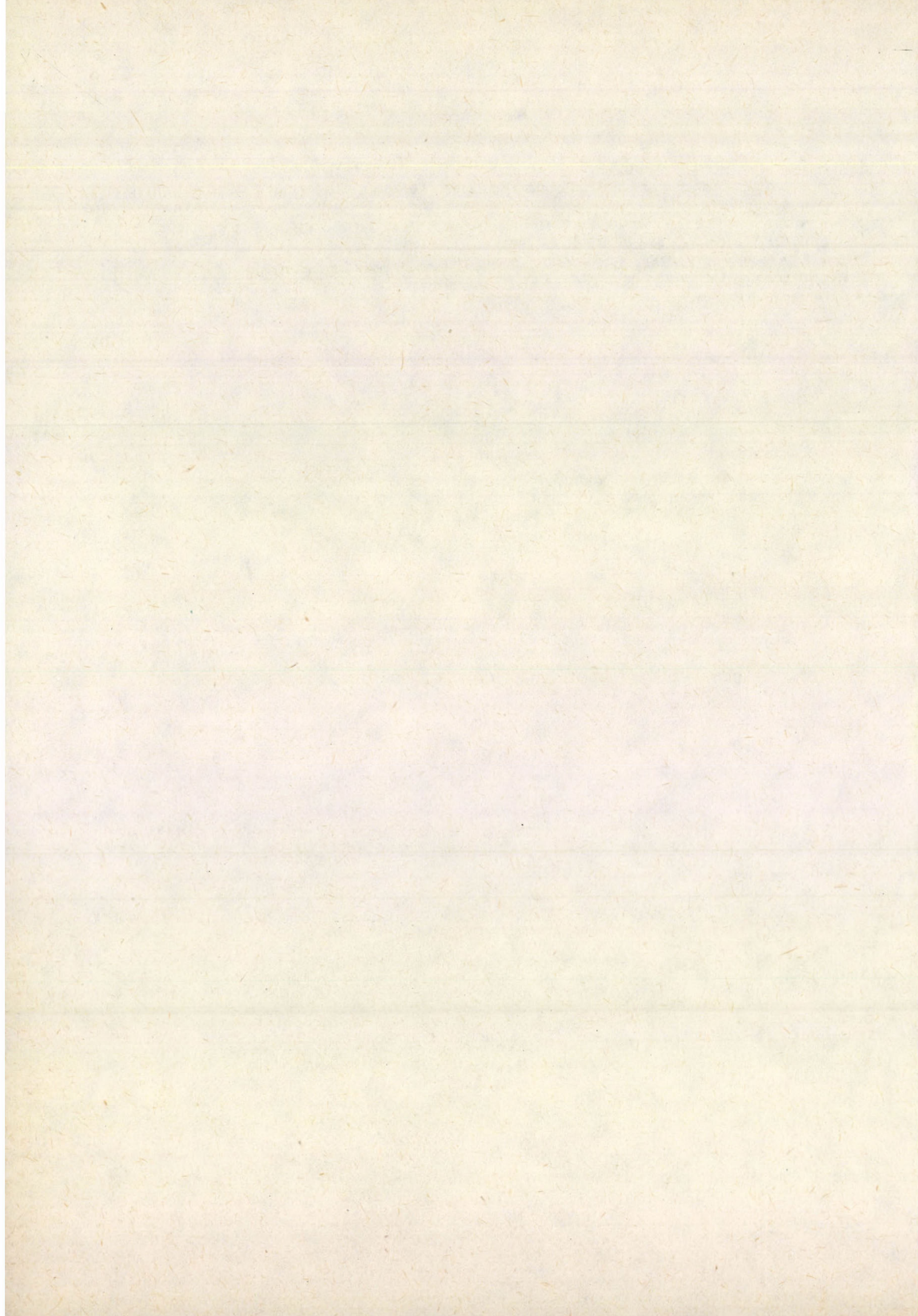


Fig. 7. Transparence of TeO_2 layers grown from solution on seed crystal

References

1. K. Nassau, J. Cryst. Growth, *15*, 171, 1972.
2. R. Rodriguez-Clemente and S. Veintemillas-Verdaguer, European Meeting on Crystal Growth '82. Abstracts, p. 249.
3. I. Földvári, K. Raksányi, R. Voszka, E. Hartmann and Á. Péter, J. Cryst. Growth, *52*, 561, 1981.
4. E. D. Kolb and R. A. Laudise, Mater. Res. Bull., *8*, 1123, 1973.
5. V. I. Popolitov, A. Ya. Shapiro and A. N. Lobachev, Soviet Physics, Crystallography *27*, 375, 1982.
6. S. A. Malyutin, K. K. Samplavskaya and M. Ch. Karapetyanc, Zh. Neorg. Himii, *26*, 1745, 1971.



Acta Physica Hungarica, 57 (3—4) pp. 303—308 (1985)

GROWTH AND LUMINESCENCE PROPERTIES OF GaSb SINGLE CRYSTALS

W. LEWANDOWSKI, M. GAJEWSKA, B. PASTUSZKA and M. BUGAJSKI

Institute of Electron Technology
*02—668 Warsaw, Poland**

The Horizontal Bridgeman Method has been used to grow undoped GaSb single-crystals. It is demonstrated that device quality material can be grown by this method.

Low temperature photoluminescence revealed excitonic structure of the spectra comparable to those observed in high purity LPE layers. Besides the edge emission, luminescence due to the Ga_{Sb} antisite defect has been observed. The mechanism of the antisite formation and the influence of background donors on its concentration is discussed.

1. Introduction

The ternary and quaternary III—V alloys based on GaSb are commonly used for fabrication of optoelectronic as well as high speed electronic devices as: the low threshold $In_xGa_{1-x}Sb$ Gunn oscillators [1], the low noise $Al_xGa_{1-x}Sb$ APDs for 1.3 μm band [2], $Ga_{1-x}Al_xAs_{1-y}Sb_y$ LEDs and LDs for 1.7 μm band [3], AlSb/GaSb superlattices for new kinds of LDs [3] and InAs/GaSb superlattices for high speed electronic devices [4]. These alloys are grown epitaxially on GaSb substrates whose quality appears to be crucial for device performance. There are several methods used for obtaining GaSb single crystals, i.e. Czochralski pulling [5—8], Bridgeman method [9], Travelling Heater Method THM [10] and solution growth methods [11—13] with either stoichiometric or nonstoichiometric melts. To date the best results have been obtained with THM method. It suffers, however, from the very low growth rate (several millimeters per day) which makes it impractical for industrial purposes. On the other hand, similarly as in the GaAs case, the Bridgeman method seems to be the most promising for obtaining device quality GaSb single crystals. It appears to be relatively easy to implement and does not require sophisticated equipment.

2. Crystal growth

In this work a modification of the Bridgeman method, i.e. Horizontal Bridgeman Method (HBM) has been applied to grow GaSb single crystals. The narrow heat zone characteristic of the THM method was used. The schematic drawing of the crystal growth apparatus is shown in Fig. 1. Direct synthesis of stoichiometric GaSb has been

* Address: Al. Lotników 32/46, 02—668, Warsaw, Poland

carried out in quartz boat with inner walls covered with pyrolitic carbon. The process was run at $T = 740$ °C for 3h in hydrogen atmosphere. After this time the synthetized material was cooled down to $T = 600$ °C at the rate of about 1.5 degree/min and then to room temperature. Prior to crystallization the polycrystalline ingots were several times purified by using the floating zone method. After purification the tail ends of the

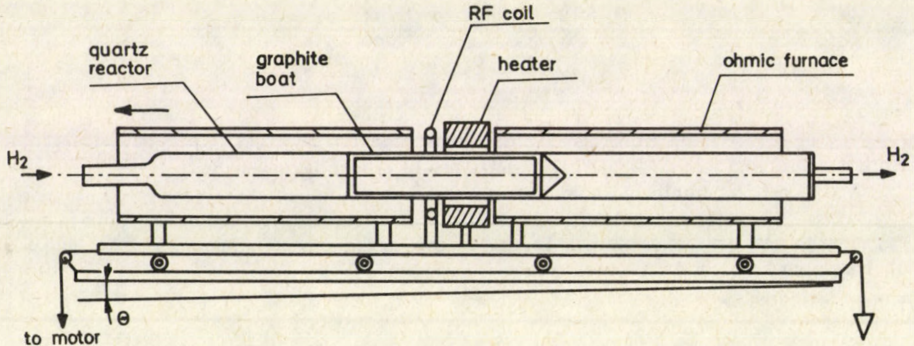


Fig. 1. Growth apparatus for Horizontal Bridgeman Method

polycrystalline ingot were cut out and the ingot was etched in the solution $HF:HNO_3:CH_3COOH$ (1:9:20) [6]. The crystallization of GaSb was carried out in the furnace with radio frequency (RF) coil suitable for obtaining a narrow heat zone. On both sides of the RF coil ohmic heaters were placed, providing the necessary preheating of the synthetized material as well as allowing to control the cooling rate of the crystallized material. Close to the coil there was an extra heater which was used to determine the front of crystallization. The reactor was inclined at an angle of 7° to avoid mass transport during the growth process. The crystallization itself was initiated by a seed material with (111) B plane oriented perpendicularly to the growth direction. At the first step of the process the seed was slightly etched by the melt to remove the surface defected layer. For the crystallization the reactor was moved with the speed of 11 mm/h. The temperature of the molten zone was kept at $T = 720$ °C. The dimensions of the grown crystals were determined by the size of the quartz boat. In the GaSb growth processes a boat of 13 mm in diameter and 150 mm in length has been used. The resulting material was p-type with a hole concentration at room temperature of the order of 10^{17} cm^{-3}.

3. Photoluminescence

The quality of the grown GaSb single crystals was characterized by photoluminescence. The photoluminescence measurements were carried out in the temperature range from liquid helium to room temperature. The samples were mounted on a cold finger of liquid helium cryostat. As excitation source an argon ion

laser was used. The emitted radiation was analyzed with a grating monochromator and detection was made using PbS cell.

It was shown [14, 15] that low temperature spectra of undoped GaSb crystals are strongly dependent both on the quality as well as on the growth conditions. The typical low temperature spectra consist of two emission bands; an edge emission and the low energy band labelled as A. The spectra for the temperature range from 17 K to 77 K are shown in Fig. 2. In our samples we have observed a fine structure of the edge emission.

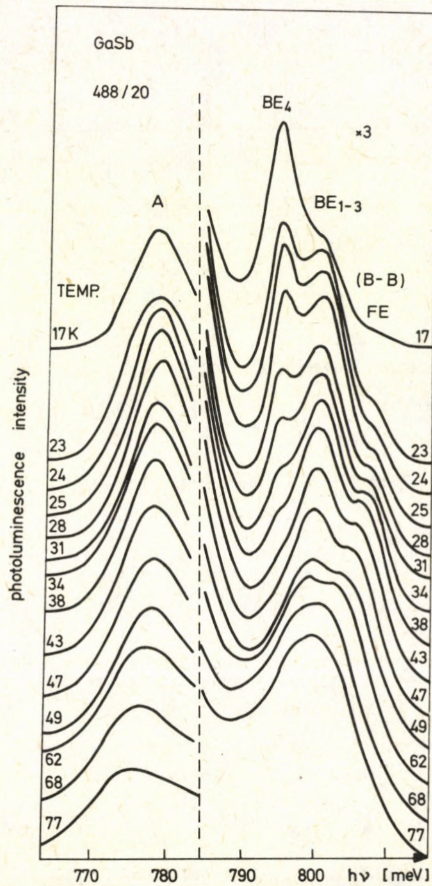


Fig. 2. Photoluminescence spectra of as grown GaSb in the temperature range from 17 K to 77 K. FE — free exciton recombination, BE_{1-3} and BE_4 — bound exciton recombination, A — band to acceptor recombination

This structure has been attributed to the recombination of free excitons (FE) and excitons bound to residual impurities. The bound exciton emission is not completely resolved, however, by comparison with the spectra of high quality LPE layers we were able to identify the observed peaks with BE_{1-3} and BE_4 transitions (notation follows

[15]). Such recombination transitions were observed so far only in the crystals grown by THM or in the LPE layers but not in the crystals grown by the Bridgeman method [14]. Since bound exciton recombination is effectively quenched by crystal inhomogeneities and residual stress its presence is a direct sign of the high crystal quality of the Bridgeman grown material. The A band situated at about 35 meV below the energy gap is connected with conduction band to acceptor recombination. With the increase of temperature above 77 K the A band becomes gradually weaker and at 300 K completely disappears (see Fig. 3). The thermal rejection of the hole, trapped at the

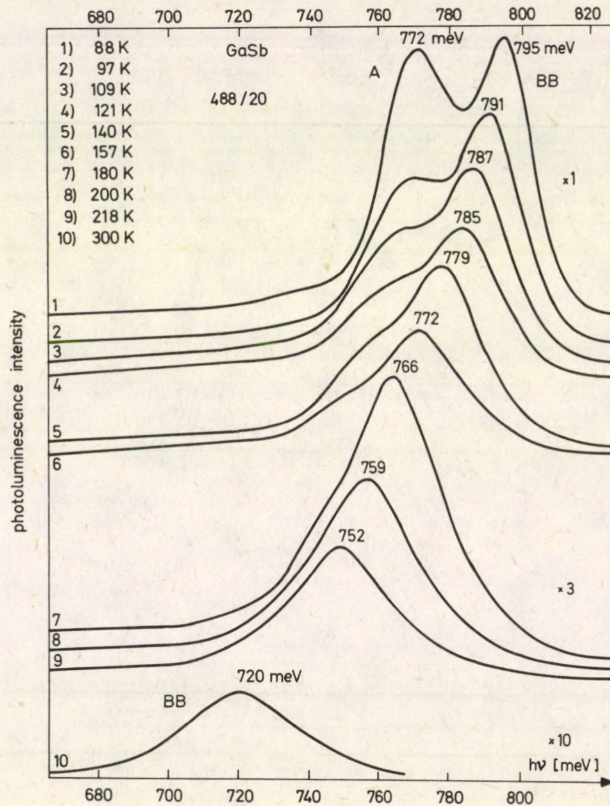
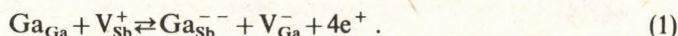


Fig. 3. Photoluminescence spectra of as grown GaSb in the temperature range from 77 K to 300 K. BB — band to band recombination, A — band to acceptor recombination

acceptor level, into valence band is responsible for that behaviour. This acceptor is directly responsible for the p-type conductivity of undoped GaSb crystals. Technological experiments have shown that stoichiometry of the crystals is responsible for such property of GaSb. It has been observed that the hole concentration in the crystals depends critically on the antimonium content in the melt [14]. On the other hand annealing experiments have indicated that the lack of antimonium is responsible for

p-type conductivity of as grown GaSb [16]. All these experiments support a model of antisite defect as being the origin of A acceptor in GaSb. For being the acceptor this defect should include gallium on antimonium site (Ga_{Sb}) but the problem whether it is a simple antisite defect or a complex including antisite defect and gallium vacancy is still unresolved.

This defect is similar to the major deep level (EL2) in GaAs [17]. At is has been stated before the concentration of Ga_{Sb} antisite depends on the number of factors. The role of these factors can be understood in terms of the following model of the antisite formation.



$$Ga_{Ga} + V_{Sb}^{+} \rightleftarrows Ga_{Sb}^{--} + V_{Ga}^{-} + 4e^{+} . \qquad (1)$$

The concentration of $[Ga_{Sb}]$ antisite defects produced according to the reaction (1) is

$$[Ga_{Sb}^{--}] = K \frac{[V_{Sb}^{+}]}{[V_{Ga}^{-}]} p^{-4} , \qquad (2)$$

where K is the equilibrium reaction constant for the reaction. The intensity of the A band in the luminescence spectrum should be proportional to the $[Ga_{Sb}^{--}]$ antisite concentration and therefore should depend (via Eq. (2)) on the free hole concentration. To check this we have compared the spectra of three GaSb crystals with slightly different background donor concentrations and resulting different free hole concentrations. The spectra referring to the temperature of 80 K are shown in Fig. 4. The hole concentrations for each sample have been measured by a conventional van der Pauw method. It can be seen that with increasing hole concentration the A band intensity drops which is equivalent to the decrease predicted by Eq. (2). Summarizing we would like to stress that although the Ga_{Sb} antisite formation leads to the generation of free holes the background donor concentration through the charge neutrality condition can influence the rate of reaction (1) and thus the Ga_{Sb} concentration.

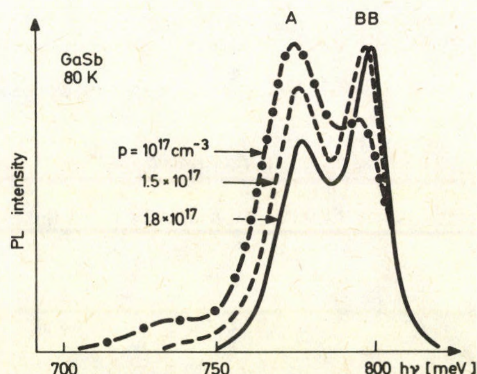


Fig. 4. Normalized 80 K photoluminescence spectra of GaSb samples with different hole concentrations

4. Conclusions

The Horizontal Bridgeman Method has been used to grow GaSb single crystals. The usefulness of this method for the production of device quality substrate material has been proved by detailed photoluminescence measurements. The low temperature spectra have revealed an excitonic structure comparable to those observed in high purity LPE layers. Since bound exciton recombination is effectively quenched by crystal inhomogeneities and residual stress its presence is a direct sign of the high crystal quality. Besides the edge emission, luminescence due to the Ga_{Sb} antisite defect has been observed. The intensity of this luminescence is proportional to the $[Ga_{Sb}^{--}]$ antisite concentration and is inversely proportional to the fourth power of free hole concentration. This behaviour suggests the following model for the antisite defect formation: (a) formation of antimonium vacancies during the solidification process, which is influenced by melt stoichiometry; and (b) defect migration to the nearest neighbour gallium site leading to the formation of double acceptor antisite $[Ga_{Sb}^{--}]$ defect. The antimonium vacancy migration occurs during the post growth cooling of the crystal. The antisite formation is described by Eq. (1).

References

1. K. Segawa, H. Miki, M. Otsubo and K. Shirahata, Electron. Letters, *12*, 124, 1976.
2. O. Hildozbrand, W. Kuebart, K. W. Benz and M. H. Pilkuhn, IEEE J. Quantum Electron., *QE—17*, 284, 1981.
3. G. Motosugi and T. Kagawa, Japan. J. Appl. Phys., *19*, 2303, 1980.
4. L. Esaki, J. Crystal Growth, *52*, 227, 1981.
5. M. Kumagawa, J. Crystal Crowth, *44*, 291, 1978.
6. S. Kondo and S. Miyazawa, J. Crystal Growth, *56*, 39, 1982.
7. M. A. Dasevskij, G. V. Kukuladze and M. S. Mirgalovskaja, Soviet Inorg. Matter., *5*, 1141, 1969.
8. M. Kumagawa, Japan. J. Appl. Phys., *16*, 2297, 1977.
9. M. Hársy, T. Görög, E. Lendvay and F. Koltay, J. Crystal Growth, *53*, 234, 1981.
10. K. W. Benz and G. Müller, J. Crystal Growth, *46*, 35, 1979.
11. F. J. Reid, R. D. Baxter and S. E. Miller, J. Electrochem. Soc., *113*, 713, 1966.
12. T. Migauchi and H. Sonomura, Japan J. Appl. Phys., *4*, 317, 1965.
13. G. Woff, P. H. Keck and J. D. Broder, Phys. Rev., *94*, 753, 1954.
14. K. W. Benz and G. Müller, Gallium Arsenide and Related Compounds 1978, Ind. Phys. Conf. Ser. No. 45; Chapter 2, 154, 1978.
15. W. Jakowetz, W. Rühle, K. Breuninger and M. Pilkhun, Phys. Stat. Sol. (a), *12*, 169, 1972.
16. J. Allegre and M. Averous, Defects and Radiation Effects in Semiconductors 1978, Inst. Phys. Conf. Ser. No 46; Chapter 1, 379, 1978.
17. J. Łagowski, H. C. Gatos, J. M. Parsey, K. Wada, M. Kamińska and Walukiewicz, Appl. Phys. Lett., *40*, 342, 1982.

Acta Physica Hungarica, 57 (3—4) pp. 309—312 (1985)

EVOLUTION OF IMPURITY CENTRES IN NaCl : Ca SINGLE CRYSTALS UNDER LASER RADIATION

M. V. GALUSTASHVILI and D. G. DRIYAEV

Institute of Physics, Academy of Sciences of the Georgian SSR, Tbilisi, USSR

The influence of multiple laser irradiation of the prethreshold intensity ($\lambda = 10.6$ μm) on the threshold of optical break-down $W$and the yield point τ of NaCl crystals with different concentration of Ca^{++} impurity (from 30 to 500 ppm) was studied. W and τ samples were measured simultaneously, which were annealed isothermally with the following hardening.

Comparison of the effect of laser and heat treatment on $W$and τ allowed to conclude that the multiple irradiation of NaCl crystals with high concentration of Ca^{++} impurity leads to dissolving the impurity complexes absorbing the radiation, causing the increase of the crystal optical strength.

In view of the current widespread use of NaCl crystals as elements of CO_2-laser optical systems, investigations of the dependence of optical breakdown threshold of these crystals on various factors are rather urgent.

It has been established that one of the mechanisms responsible for the optical breakdown of transparent dielectrics within the nanosecond range of laser radiation pulses is the thermal explosion on absorbing inclusions. From such a viewpoint, one can explain the dependence of the optical breakdown threshold in alkali halide crystals (AHC) on the type, concentration, and state of impurity [1—3]. It is known that the impurity composition also determines the mechanical properties of AHC. Evidently this fact led to the correlation of AHC yield point and optical strength [1, 2]. The increase of the impurity centre concentration and dimensions was found to increase the yield point and to decrease the optical breakdown threshold.

While investigating the influence of laser radiation upon the materials, it was observed that the multiple irradiation of the same region of an NaCl crystal by a series of successive pulses of pre-threshold intensity ($\lambda = 10.6$ μm) results in an increase of the optical breakdown threshold [4, 5]. It has been hypothesized that this phenomenon is due to laser-induced change of the absorbing centre state.

To confirm this hypothesis, we have investigated NaCl crystals with Ca impurity contents from 30 to 500 ppm subjected both to laser radiation and thermal treatments. The radiation source was a single-mode CO_2 laser with a pulse of 150 ns and maximum radiation energy of 2 J. The beam was focused within the crystal volume. The intensity of the incident radiation was varied by means of neutral filters. The assumed optical breakdown threshold W represented the minimum intensity causing local destruction fixed according to the flash, accompanying the optical breakdown. To measure the

yield point we used a deformation-testing machine with a stability of $2 \cdot 10^3$ kg/mm; the deformation rate was 10^{-3} s^{-1}.

We measured experimentally the optical breakdown threshold W and the yield point τ of initial specimens and those subjected to multiple action of laser radiation of pre-threshold intensity (10—15 successive pulses with an intensity equal to 0.7—0.8 of that of the optical breakdown W_0 in the case of unit effect).*

W and τ were also measured on the specimens subjected to isothermal annealing at various temperatures with the subsequent "quenching" (rapid cooling in air).

The experimental results presented in Fig. 1 and in Table I show that repeated laser irradiation substantially influences W and τ only in the case of crystals with high

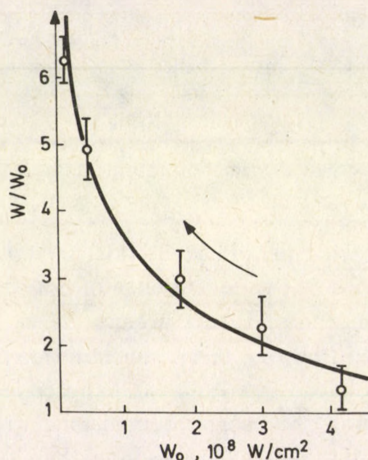


Fig. 1. Dependence of relative change of optical breakdown threshold on the initial optical strength W_0 of NaCl crystals subjected to repeated laser irradiation (the arrow points to the increase of Ca impurity content from 30 to 500 ppm)

Table I

Change of yield point of $NaCl:Ca^{++}$ crystals under laser and thermal treatments

	Yield point τ, g/mm^2	
	$C=30$ ppm	$C=500$ ppm
Control sample	80 ± 10	320 ± 25
Laser treatment	80 ± 10	230 ± 20
Thermal treatment at 400 °C	75 ± 10	200 ± 20

* When measuring τ, the whole specimen (in portions) was subjected to multiple action of laser radiation of pre-threshold intensity.

impurity concentration. The optical strength of such crystals increases (in some cases by an order of magnitude) and the yield point decreases, whereas in crystals with low impurity content these parameters remain unchanged after laser irradiation.

Thermal treatment also leads to W increase (Fig. 2) and τ decrease (Table I) in the crystals with impurities, the largest effect being observed at the temperature 400 °C.

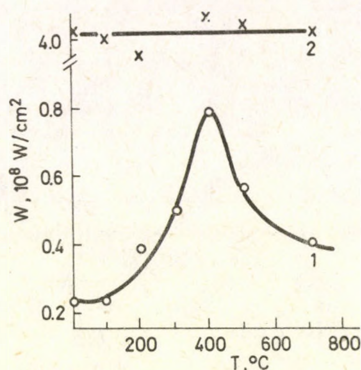


Fig. 2. Dependence of optical breakdown threshold on annealing temperature: $1-C=500$ ppm; $2-C=30$ ppm

The identical nature of the effects resulting from repeated laser irradiation and from thermal treatment of the crystal shows the same influence of both factors upon the impurity state. It is known that at high concentrations of divalent cationic impurities, the crystal annealing and rapid cooling dissolve the impurity centres—the process being the most intensive for quenching from 400 °C [7]. This causes the optical strength increase and yield point decrease observed in our experiments on crystals with impurities. In the case of low concentrations, no impurity clusters are formed and the impurity remains in the dispersed state [6, 7]; therefore, neither W nor τ changes are observed during thermal treatment.

Evidently, the same process of the impurity complex dissolution determines the phenomenon of the optical strength increase and yield point decrease of the crystals under the action of repeated laser irradiation. If the radiation intensities exceed the optical breakdown threshold, the electromagnetic radiation absorption by the impurity centre leads to thermal microexplosion; consequently, the impurity centre burns out [8]. In the case of pre-threshold intensities used in the present work, local heating of the material with subsequent rapid cooling takes place thereby contributing to the decomposition of the impurity complexes similarly to ordinary annealing followed by quenching.

The repeated action of laser radiation causes the irreversible process of defect accumulation due to the breaking of the interatomic bonds [9]. This process reduces the optical and mechanical strength of the material, finally leading to visible

destruction. However, the results of the present paper show that for NaCl crystals with a high content of Ca impurity the process of impurity complex dissolution causing the effective increase of their optical strength, plays the predominant role.

References

1. M. P. Shaskoljskaya, G. F. Dobrzhanskii, O. M. Kugaenko, M. M. Tagieva, L. M. Soifer, S. F. Ulanov, Fiz. Tverd. Tela, *23*, 1834, 1981.
2. V. V. Nabatov, L. M. Beliaev, N. N. Dimenko, R. Voska, Acta Phys. Hung., *47*, 107, 1979 (in Russian).
3. B. G. Gorshkov, Yu. K. Danilenko, A. S. Epifanov, V. A. Lobachev, A. A. Manenkov, A. V. Sidorin, JETP, *72*, 1171, 1977.
4. B. G. Gorshkov, Yu. K. Danilenko, A. A. Manenkov, A. M. Prokhorov, A. V. Sidorin, Quantum Electronics, *8*, 148, 1981.
5. M. V. Galustashvili, Soobshchenia AN GSSR, *106*, 505, 1982.
6. I. Sharkesi, A. A. Predvoditelev, Crystallography, *24*, 622, 1979.
7. M. G. Buravleva, L. M. Soifer, Ukrain. Fiz. Zhurn., *19*, 1148, 1974.
8. V. S. Golubev, V. N. Snopko, Fiz. Tverd. Tela, *19*, 293, 1977.
9. S. N. Zhurkov, S. B. Eronjko, A. Chmel, Fiz. Tverd. Tela, *22*, 2040, 1980.

Acta Physica Hungarica, 57 (3—4) pp. 313—316 (1985)

KDP Q-SWITCHES AND SECOND HARMONIC GENERATOR FOR HIGH-POWER SOLID-STATE LASERS

J. BAKOS, T. JUHÁSZ, CS. KUTI and L. V. VANNAY

Institute of Physics, Technical University of Budapest
1521 Budapest, Hungary

KDP electrooptic light modulators and optical second harmonic generators have been developed for high power solid-state laser applications. The device parameters and the results in switching and frequency doubling of Nd-glass lasers are reported.

1. Introduction

KDP (potassiumdihydrogenphosphate) crystals grown from water solution are very suitable for high power laser applications because of their good optical quality, high resistivity to laser damage and large size available. Longitudinal Z-cut KDP electro-optic modulators are preferred to other modulators in Q-switching of high power solid-state lasers because of their high transmission ratio, large aperture and simple construction [1]. Although KDP crystals are characterized by medium nonlinearoptic properties [2] they are widely used for frequency doubling of pulsed and CW Nd-YAG, Nd-glass, ruby and different dye lasers because of the advantages listed above and mainly the possibility of their room temperature phase matching.

2. Preparation of KDP samples

Optical quality KDP monocrystals were grown by automatically controlled, slowly cooling (0, 3 °C/day) of water solution of "Optipur" grade MERCK KH_2PO_4 material from 50 °C to 30 °C. The double sheet thermostat was supplied with electronically controlled, programmable alternating mixing [3]. Two prisms of monocrystals sized about 45 mm × 50 mm × 85 mm were obtained. Rectangular samples oriented optically were cut from the prisms for Q-switch and frequency doubling devices. The end faces of the samples were optically flat polished. The samples were selected by strain-optic, interferometric and light-scattering methods of optical test equipment KP-74 [4]. Mechanical and microscopic investigation of the dislocation structure was performed [5] on the crystal samples and their resistivity against laser damage was also investigated [6]. Strain free 0° Z cut crystal samples of $1000/cm^2$ dislocation density and $1\ GW/cm^2$ damage threshold, measured by ruby laser, were used for large aperture longitudinal electroopic modulators designed for Q-

switching of high power Nd-glass solid state lasers. 45° Z cut samples similarly treated were further oriented and sliced to obtain phase matched samples for optical second harmonic generation of Nd-glass radiation.

3. Investigation of longitudinal 0° Z KDP modulators

Some construction versions of 0° Z longitudinal KDP modulators were investigated. End-plate and stripe electroded modulator crystals were compared considering electric field characteristics of the electrooptic modulators as follows: the static and dynamic half wave voltages, the electric field distribution uniformity and the electric breakdown. The effect of mounting and mechanical clamping on the transmission distribution and the contrast ratio of the modulators was investigated.

20 mm × 20 mm aperture, 25 mm long 0° Z cut KDP samples supplied with copper plate electrodes with 15 mm diameter circular openings pressed by spring to the Z-faces of the sample were compared to the identically sized and oriented modulator crystals electroded with 4 mm width, 0.1 mm soft copper stripes rolled on the lateral faces at the optical windows. The results obtained for the half-wave voltages, the electric field uniformity and the contrast-ratio are summarized in Table I.

Table I

Comparison of the modulator characteristics

Electrode	Clear aperture	Half-wave voltage [kV]		Field nonum	Contr. ratio
		static	dynamic	%	
end-plate	15 mm diam.	14.1 ± 1	18.2 ± 2	16	980:1
stripe	20 mm × 20 mm	10.5 ± 1	15.3 ± 2	5	1100:1

The half-wave voltages were determined by taking the transmission-voltage characteristics of the modulator crystals placed between crossed polarizers. The transmission was measured with 1 mm diameter He—Ne laser beam passing through the modulator in the middle of the aperture. To characterize the electric field nonuniformity the difference of the maximum and minimum half-wave voltage expressed in the percentage of the maximum one, i.e. the half-wave voltage measured in the middle of the aperture, was used. The half-wave voltage distribution was obtained by scanning the He—Ne laser beam across the aperture. The stripe-electrode construction proved to be more advantageous than the end-plate construction having less half-wave voltage and more uniform driving field and, consequently, a smoother transmission distribution [7]. Only the problem of electric breakdown at lower voltages arose because of the small electrode distance. To avoid the breakdown the length of the crystal can be increased or, as it was done in our experiments, the stripe electroded modulator crystal can be embedded in silicon rubber before mounting it in the polyamide house.

4. The frequency doubling crystals

For frequency doubling of 1.06 μm laser output KDP optical second harmonic generators were constructed. 20 mm × 20 mm aperture, 25 mm long KDP samples phase matched for the *oo→e* interaction frequency doubling were sliced from the 45° Z cut prism samples selected by quantity tests mentioned. The phase matching direction, i.e. 41.2° inclination of the axis of the KDP sample to the Z axis, was optically set. The frequency doubling crystals were embedded in silicon rubber and mounted in polyamide house.

5. Q-switching and frequency doubling of the Nd-glass laser radiation

The stripe-electroded modulators have been used for controlling the Q-factor of the laser resonator (Fig. 1) operating with 20 mm diameter, 240 mm long Nd-glass rod as active element. The modulator was biased to be closed and the laser has been switched by very fast switch off the biasing electric field. As the modulator crystal was transmitted by the laser radiation twice, the modulator was closed by quarter wave voltage taken about 10 kV-s which, regarding the wavelength difference of the He—Ne and Nd-glass laser radiation, agrees with the measured static half-wave voltage. 1.06 μm wave length, 15 ns halfwidth, 10 J energy giant pulses have been reproducibly generated.

The giant pulses of the Nd-glass laser were frequency doubled by the optical second harmonic generator reported above. 500 mJ maximum energy, frequency doubled, 0.53 μm wavelength pulses could be obtained. Although 0.5 J in green is a fairly high energy, the 5% efficiency of the frequency doubling is a moderate one considering the 14% efficiency reported in [8] for KDP crystals. The possible reason for the moderate efficiency could be the large divergence and multimode operation of the Nd-glass laser constructed for the investigations. Improvements of the beam parameters are designed to repeat the frequency doubling experiments.

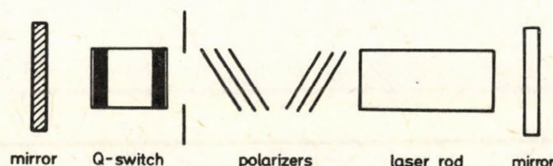


Fig. 1. Nd-glass laser switched by KDP 0° Z cut electrooptic modulator

References

1. W. Koechner, Solid-State Laser Engineering, Springer-Verlag, New York, 1976.
2. F. Zernike and J. E. Midwinter, Applied Nonlinear Optics, John Wiley and Sons, New York, 1973.
3. L. Vannay, Kristall und Technik, *14* 39 (1979).
4. L. Szabó, E. Hartmann, Cs. Kuti and L. Vannay, Quality Test of KDP Monocrystals by Optical Methods. 3rd Hungarian Conference on Crystal Growth 1983; contributed paper.
5. J. Sárközi, L. Vannay, A. Tóth and T. Keszthelyi, Periodica Polytechnica, Chem. Eng., *23*, 255, 1979.
6. J. Kovács and E. Szil, private communication: damage threshold is about 1 GW/cm^2 in focused ruby-laser beam; no damage was observed by 400 MW/cm^2 power density radiation of Nd-glass laser reported in recent work.
7. Yu. E. Kamach, R. P. Kiathin and E. N. Kozlovski, OPM No 4. 18, 1975.
8. L. Szabó, Thesis, Technical University, Budapest, 1981.

Acta Physica Hungarica, 57 (3—4) pp. 317—321 (1985)

ON THE VARIATIONS IN SURFACE DIFFUSION COEFFICIENT OF LFM CRYSTAL

M. RAK

Institute of Physics, Technical University of Łódź
93-005 Łódź, Poland

Calculations of values of surface diffusion coefficient as a function of surface concentration of growth units show that for solution growth of $HCOOLi \cdot H_2O$ crystal this coefficient practically remains constant (maximum variation is about 1%) for the variation of estimated n_s for supersaturation $1.01 \leq \beta \leq 1.1$.

1. Introduction

Crystal growth from solution is usually described by the BCF surface diffusion model [1—5]. In this theory it is assumed that the surface diffusion coefficient is independent of local solute concentration and that surface vacancies are not present. The surface supersaturation ratio is defined as [1]

$$\beta_s = \frac{n_s}{n_{s0}}, \tag{1}$$

where n_s is the number of adsorbed growth units per unit area and n_{s0} is the value of n_s at saturation.

Madsen proposed another definition of surface supersaturation ratio [6, 7]

$$\mu_2 = \mu_2^c + kT \ln \beta_s, \tag{2}$$

where μ_2 is the chemical potential of solute in solution, μ_2^c the chemical potential of the crystal, and in terms of thermodynamic considerations of a Bragg—Williams monolayer model with empty sites, he derived equations for the variations of surface supersaturation and surface diffusion coefficient with surface concentration of growth units [7]. By analogy with Eyring's theory for volume diffusion [8], he expressed the surface diffusion coefficient by [7]

$$D_s = (a^2/\tau_{sd}) \cdot (1 - \Theta), \tag{3}$$

where a is the shortest distance between growth units, τ_{sd} is the time constant for the jump of a molecule to a neighbouring empty site and Θ is the fraction of sites occupied. As $(1-\Theta)$ is function of β_s which depends on n_s, so does D_s [7], and τ_{sd} may be treated as

constant according to Bragg—Williams theory. Thus the dependence D_s on n_s is given by [7]

$$\frac{D_s}{D_{s0}}=\frac{n_0-n_s}{n_0-n_{s0}}\cdot\exp\left[\frac{c(\varphi_2-\varphi_1+2\varphi')\cdot(n_s-n_{s0})}{2kTn_0}\right], \qquad (4)$$

where D_{s0} is the value of D_s at saturation and n_0 the number of sites per unit area of interface.

For the one-dimensional problem of a series of equidistant parallel steps, using the first law of Fick, Madsen [7] established a generalized form of the BCF equation. This is differential equation in relation to n_s and it may be integrated numerically if n_s and $\frac{dn_s}{dy}$ are known at a given point on the surface.

2. Method of calculations

The generalized form of BCF equation, obtained by Madsen, was solved for $HCOOLi\cdot H_2O$ (LFM) crystal [9, 10]. The solution of this equation shows that in case of LFM crystal n_s can change from 8.40 nm^{-2} to 8.60 nm^{-2} for supersaturation $1.01\leq\beta\leq 1.10$. Therefore the surface diffusion coefficient, D_s, is calculated for values of n_s from 8.40 nm^{-2} to 8.60 nm^{-2}. The solution of the equation enables to find n_s as a function of distance y from the step centre for given supersaturation and n_s as a function of supersaturation, β, at the step centre ($y=0$). From these calculations [9] it follows that the concentration of adsorbed growth units, n_s, is the greatest at the centre of the "terrace" and decreases non-linearly as the step edge is approached. The calculation [9] of n_s at the step centre as a function of β shows that $n_s(\beta)$ curves have a maximum. For low supersaturation ($\beta<\beta_M\approx 1.02$) concentration n_s increases with increasing supersaturation, while for higher values of $\beta>\beta_M$ it decreases with increasing supersaturation. The $n_s(\beta)$ and $n_s(y)$ dependences enable to obtain D_s as a function of β and D_s as a function of y.

The values of D_s as a function of n_s are presented for $\{10\bar{1}\}$ and $\{110\}$ faces of LFM crystal, and D_s as functions of β and of y are presented for $\{10\bar{1}\}$ faces only, because the dependence for all faces of these crystal are similar. All calculations are made for a crystal grown in temperature $T=307.65$ K, and the values of the parameters of Eq. (4) are given in Table I.

3. Results

Calculations show that in the case of LFM crystals the surface diffusion coefficient practically remains constant (maximum variations about 1%) for the variation of estimated n_s for supersaturation used and that it decreases quasilinearly with increasing n_s (Fig. 1). It is a consequence of poor variations in exponential term of Eq. (4).

Because D_s decreases with increasing n_s, the dependence $D_s(\beta)$ at $y=0$ and $D_s(y)$ (Figs. 2, 3) are the inverse of those $n_s(\beta)$ at $y=0$ and $n_s(y)$, respectively. So, the surface diffusion coefficient, D_s, is the smallest at the centre of the "terrace" and increases non-linearly as the step edge is approached (Fig. 2), and curve $D_s(\beta)$ has a minimum for supersaturation $\beta=\beta_M$, for which concentration n_s has a maximum.

The poor variations in surface diffusion coefficient, D_s, with supersaturation confirm that in crystal growth models D_s may be taken as constant, particularly that in view of the present state of the theory, most of the parameters of these models can merely be estimated.

Table I

The values of parameters used in calculations

Parameter	Value of parameter for		Unit	Reference
	$\{10\bar{1}\}$	$\{110\}$		
c	4	4	–	[9, 10]
D_s	$9.39\cdot10^6$	$1.61\cdot10^6$	nm^2s^{-1}	[9]
n_0	84	84	nm^{-2}	[9, 10]
n_{s0}	8.39	8.39	nm^{-2}	[9, 10]
T	307.65	307.65	K	–
$\varphi_1-\varphi_2$	$4\,kT$	$4\,kT$	J	[9, 10]
φ'	$0.5\,kT$	$0.3\,kT$	J	[9, 10]

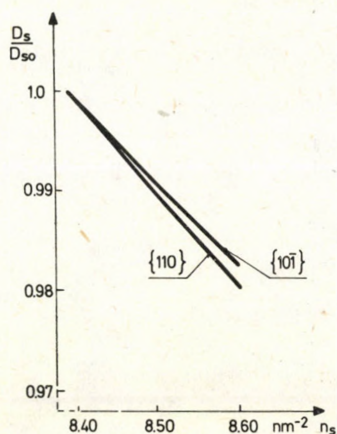


Fig. 1. Dependence of surface diffusion coefficient, D_s, on concentration, n_s, of growth units for $\{10\bar{1}\}$ and $\{110\}$ faces of LFM crystal

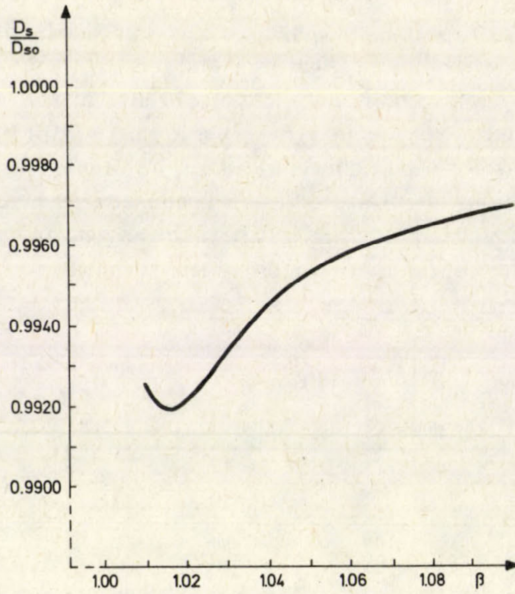


Fig. 2. Dependence of surface diffusion coefficient, D_s, on supersaturation, β, at the centre of step, for $\{10\bar{1}\}$ face of LFM crystal

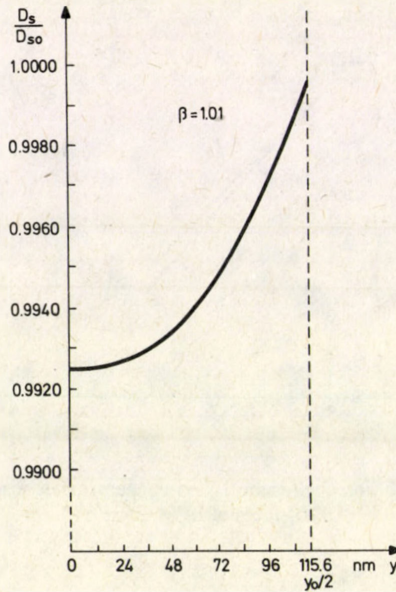


Fig. 3. Dependence of surface diffusion coefficient, D_s, on distance, y, from step centre for $\{10\bar{1}\}$ face of LFM crystal at supersaturation $\beta = 1.01$

References

1. W. K. Burton, N. Cabrera and F. C. Frank, Phil. Trans. Roy. Soc. London, *A 243*, 299, 1951.
2. P. Bennema, J. Crystal Growth, *1*, 278, 1967.
3. P. Bennema, J. Crystal Growth, *1*, 287, 1967.
4. P. Bennema, J. Crystal Growth, *5*, 29, 1969.
5. P. Bennema and G. H. Gilmer, in: Crystal Growth: An Introduction, Ed. P. Hartman, North-Holland, Amsterdam, 1973, p. 263.
6. H. E. L. Madsen, J. Crystal Growth, *39*, 250, 1977.
7. H. E. L. Madsen, J. Crystal Growth, *46*, 495, 1979.
8. H. Eyring, J. Chem. Phys., W, 283, 1936.
9. M. Rak, Thesis, Technical University, Łódź, 1980.
10. J. Karniewicz and M. Rak, in: Industrial Crystallization '81, Proceedings of the 8th Symposium on Industrial Crystallization, Budapest 1981, Ed. S. J. Jančić and E. J. de Jong, North-Holland Amsterdam 1982, p. 293.

List of symbols

a — shortest distance between growth units
c — number of horizontal nearest neighbours
D_s — the surface diffusion coefficient
D_{s0} — the values of surface diffusion coefficient at saturation
k — Boltzmann constant
n_0 — the number of sites per unit area of interface
n_s — the number of adsorbed growth units per unit area
n_{s0} — the number of adsorbed growth units per unit area at saturation
T — absolute temperature
y — the axis being perpendicular to the step edge
y_0 — the distance between two steps of an equidistant step train (the width of "terrace")
β — the supersaturation ratio
β_M — the value of supersaturation ratio for which dependence $n_s(\beta)$ has a maximum
β_s — surface supersaturation ratio
μ_2 — the chemical potential of solute in solution
μ_2^c — the chemical potential of the crystal
φ_1 — the energy of interaction of two liquid molecules on neighbouring sites
φ_2 — the energy of interaction of two solid molecules on neighbouring sites
$\frac{1}{2}(\varphi_1+\varphi_2)+\varphi'$ — the energy of interaction of a pair of liquid-solid molecules on neighbouring sites.

Acta Physica Hungarica, 57 (3—4) pp. 323—334 (1985)

PRELIMINARY STUDIES ON CRYSTAL GROWTH OF $Pb_{1-x}Cd_xS$ SOLID SOLUTIONS

P. ŞTEŢIU

Department of Physics, University of Cluj-Napoca
3400 Cluj-Napoca, Romania

Results are presented concerning the growth of $Pb_{1-x}Cd_xS$ solid solutions crystals using the Bridgman technique (BT) and the vapour-phase technique (VPT). It was found that the component concentration of VPT crystals was nearer to that of the starting mixtures than that of BT crystals. The VPT crystals showed only p-type electrical conductivity, independently of the type of starting materials. Using BT samples and X-ray powder diffraction analysis two points of the liquidus line of the X—T phase diagram of the system were determined. By means of microbeam analysis in a scanning microscope we found good average homogeneity of Cd distribution in the examined samples.

1. Introduction

Our interest in studying the PbS—CdS system is justified by the well known semiconductor properties of both components.

It is established that in this system there is only one region of solid solubility [1—4]. The phase diagram of the system is not known in every detail. In [1, 2] the position of the solidus line was determined at temperatures up to 950 °C. In [4], based on DTA studies, the solidus line was drawn between 1080 °C and 1115 °C, while in [3] the same technique was used to draw both solidus and liquidus lines between 1055 °C and 1113 °C. The so-determined lines do not fit together, the limit of solid solubility (x_M) being $x_M = 0.24$ at 950 °C [2], $x_M = 0.15$ at 1080 °C [4] and $x_M = 0.30$ at 1055 °C [3].

In two works [5, 6] laser heterojunction diodes made on $Pb_{1-x}Cd_xS$ synthetic crystals were reported. The authors mentioned that their crystals were grown using BT [5] and VPT [6] but no details were given about the growth procedure. In [3] a VPT in a closed system was used in order to grow $(Pb_{1-x}Cd_x)_{0.51}S_{0.49}$ single crystals. These crystals showed n-type electrical conductivities and only at $x > 0.08$ does CdS precipitate as a second phase.

In order to study the physical properties of $Pb_{1-x}Cd_xS$ crystals we tried to grow bulk crystals using BT and VPT. This work presents our results and discusses them on the basis of some structural investigations performed with X-ray analysis and secondary emission microscopy.

2. Experiment

2.1. Raw materials

Because we had no pure PbS we had first to synthesize it. We used for this purpose spectral pure sulphur (Johnson—Matthey) and lead purified in our laboratory by zone melting.

The synthesis of PbS was performed following a variant of Lawson's old method [7]. This variant is to be described in detail elsewhere [8].

In order to check the quality of our PbS we grew from it single crystals with BT and we measured their electrical properties (Hall and Zeebeck effects, electrical conductivity) at room temperature. The obtained crystals showed p-type electrical conductivity with carrier concentration of $p \approx 4.10^{24}\,m^{-3}$. This corresponds to an excess of sulphur of about 1.10^{-3} at % relative to the exact stoichiometry [9].

We submitted the p-type PbS to high temperature treatment in order to obtain n-type PbS (with lead in excess). For this purpose we used a two zone temperature furnace. In the high temperature region was set p-type PbS, in the low temperature region was set pure lead (both being sealed in an evacuated fused quartz ampoule). Appropriately choosing the temperatures of these regions [9] we succeeded in obtaining n-type PbS with a carrier concentration of $n \approx 4.10^{24}\,m^{-3}$.

Further we used both n and p-type PbS in an attempt to check the influence of nonstoichiometry of PbS on crystal growth of the solid solutions under discussion.

CdS was obtained by the courtesy of Prof. E. Gutsche of the IVth Physical Institute of Humboldt University (Berlin, GDR). It was there purified by multiple recrystallizations from the vapour phase. We used this polycrystalline material to grow CdS single crystals using the well known Pipper–Polich method [10]. Our crystals were lemon yellow colour and showed a dark electrical resistivity of about 10Ω m at room temperature, which in our opinion was acceptable.

2.2. The BT growth of $Pb_{1-x}Cd_xS$ crystals

For the purpose of growth we built a Bridgman furnace whose section and temperature distribution along its axis are shown schematically in Fig. 1.

The two components (PbS and CdS) were introduced in the desired quantities in purified quartz-glass ampoules, which were afterwards evacuated down to $2.6.10^{-3}$ Pa and then sealed. These ampoules were placed in purified larger ones which were also evacuated to the same pressure and then sealed. In this way we prevented the diffusion of oxygen through the walls of the inner ampoules.

The ampoules were then introduced to the upper zone of the furnace and the power supply was turned on. The steady temperatures in the two zones were chosen to be 50 °C above and 50 °C below the corresponding point on the solidus line previously determined by us 4 (Fig. 2)

The temperature regime was maintained constant with an accuracy of 0.5%.

After 10 hours of heating at the chosen temperature, the ampoule was moved to the lower temperature region at a rate of several millimetres per hour.

When the whole ampoule reached the lower temperature region of the furnace, we removed it at once, in order to freeze-in the cadmium concentration at its high

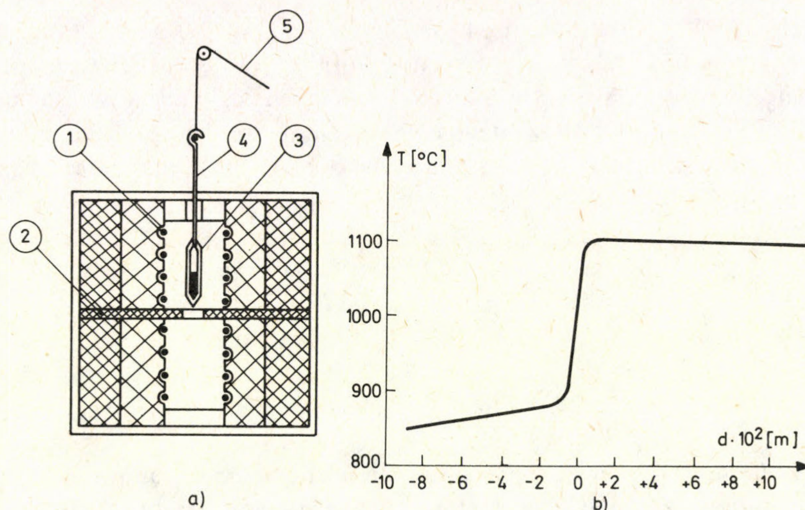


Fig. 1. Details of Bridgman apparatus. a) Longitudinal section of the furnace: 1-heating element, 2-ceramic obturator, 3-fused quartz-glass ampoule, 4-quartz stick, 5-metallic wire. b) Typical temperature distribution along the axis of the furnace. The zero point on the distance axis scale corresponds to the middle of the ceramic obturator

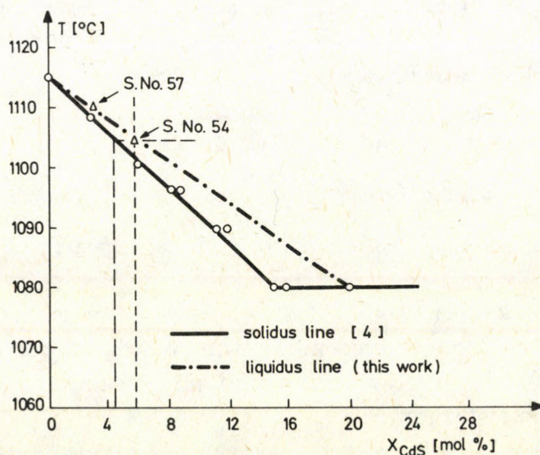


Fig. 2. Phase diagram of PbS—CdS system at higher temperatures. The dotted lines indicate how the points on the liquidus line were found using BT samples

temperature level. In this way we surely induced great mechanical stresses but our primary goal was to try to obtain crystals with large CdS concentrations.

2.3. The VPT growth of $Pb_{1-x}Cd_xS$ crystals

For the VPT approach we used the same technology to seal and evacuate the ampoules as described in the previous Section. For the furnace we used the upper part of our Bridgman furnace, the top part of which was taken off. In this way we obtained a temperature gradient of a few degrees per centimetre whereas the maximum temperature difference between the top and the bottom of the ampoules was about 20 °C. The temperature at the bottom of the ampoules was between 700 °C and 1000 °C.

We held our samples in the heated furnace for two weeks; the temperature was then merely decreased to 500 °C, after which the furnace cooled down.

2.4. X-ray diffraction analysis

In order to check the CdS concentration in the obtained samples we used the known dependence of the lattice constant versus CdS concentration [4].

For this purpose we used Debye-Scherrer cameras (XDSA) of 57.3 mm diameter and $Cu_{K\alpha}$ radiation. Measuring 622, 711, 640 and 642 reflexions and employing an extrapolation procedure [11] we achieved an accuracy of 1.10^{-4} nm in lattice constant determinations, i.e. an accuracy of about 0.002 in x determinations.

We checked the monocrystallinity of our samples by means of Laue patterns (XLA). For this purpose a slide was cut from each sample perpendicularly to its axis. This slide was first mechanically polished and then chemically cleaned in order to remove the damaged surface layers.

These slides were introduced into a Laue camera and irradiated with a wide X-ray spectrum using a W anode and a voltage of 55 kV.

2.5. Scanning electron microscopy and microbeam analysis

Scanning electron microscopy (SEM) was used in order to have information about the surface of the samples (presence of twin boundaries, precipitations of CdS, etc.). Microbeam analysis (MBA) was employed in order to have information about cadmium distribution in the samples.

The samples used in XLA and in physical measurements were used in SEM (JEOL JSM-50A) provided with an EDAX spectrometer at an acceleration voltage of 25 kV.

Our EDAX detection system had a resolution of 100 eV. Because a silver paste was used to connect electrically the samples with the microscope, the cadmium $L_{\alpha 1}$ line was superposed on the silver $L_{\beta 1}$ line in channel 32 and the cadmium $L_{\beta 1}$ line was superposed on the silver $L_{\beta 2}$ line in channel 34. Therefore, in order to avoid any possible interference with silver, we had to use for cadmium distribution studies the cadmium $L_{\beta 2}$ line (channel 36) in spite of the fact that its intensity was several times less than that of the other two mentioned lines.

On the same photos we superposed both SEM and MBA images for each sample.

3. Results and discussion

3.1. BT samples

These ingots always consisted of several rather big twins (each of them with a volume of 15—35 mm^3) and had visible irregular borders (Fig. 3).

Along their longitudinal axes the BT ingots showed that Cd concentration at the bottom or at the top of the samples was lower or higher respectively than that of the initial mixtures. This fact suggested a way to find some points on the liquidus line of the phase diagram of the system using BT samples as follows:

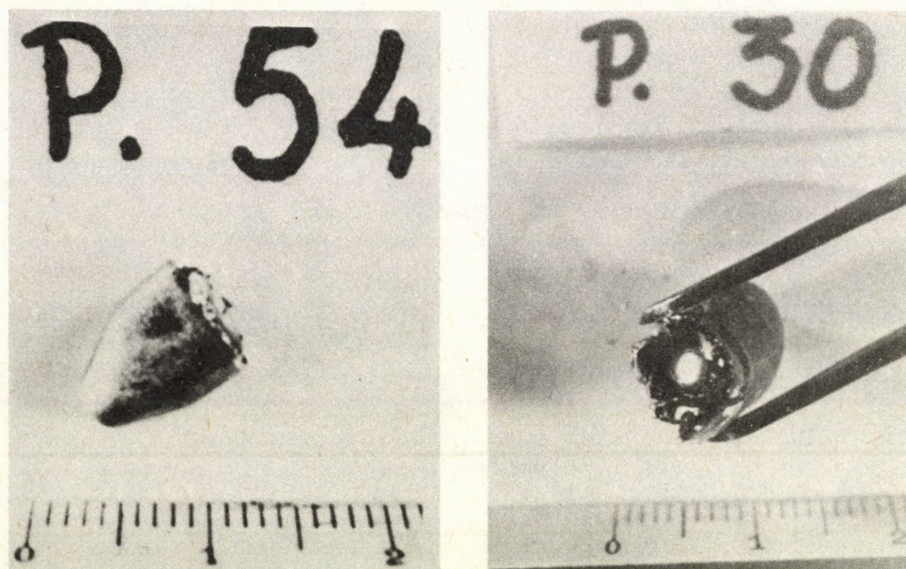


Fig. 3. Photos of two Bridgman ingots of $Pb_{1-x}Cd_xS$ solid solutions. The numbers correspond to the sample numbers from Table I. The scale is in millimetres

Let us denote by x_i the cadmium concentration of the initial mixture submitted to the growth process; by x_L and x_S the cadmium concentrations on the liquidus L and solidus S lines of the $T-x$ phase diagram corresponding to the same temperature $T=T_L=T_S$.

Examining the experimental procedure (Sect. 2.2) and refs. [3, 4] we assume:

a) For $T \geq T_L$ (but not much higher than T_L) there is complete miscibility between CdS and PbS at any x_i. This means that at $T=T_L$, $x=x_i$.

b) The BT crystallization process occurs in almost equilibrium conditions. This means that the very first crystals which appear during the cooling process of the samples have $x=x_S$.

If one measures x_B · the concentration on the very bottom-B of the ampoules, $x_B=x_S$ and one may determine $T_S(T_S=T_L)$ if the position of the solidus line of the system is known. If in addition to all this the bottom is very sharp, the amount of the liquid phase which is crystallized at $T=T_S=T_L$ is very small. As a result $x_L=x_i$ and the diffusion of Cd through the bottom is not important during the next steps of crystallization of the sample. Intersecting the line $T=T_S$ with the line $x=x_i$ one may find a point on the liquidus line.

We succeeded in separating the very bottoms from samples No. 54 and 57 and in determining with XDSA their x_B (Table I).

Now let us suppose that the phase diagram given in [3] is valid. Then for $x_i=0.0568$ (sample 54), $x_S \cong 0.002$ and for $x_i=0.03$ (sample 57), $x_S \cong 0$. However, our data show that for sample 54 $x_B=x_S=0.045$ and for sample 57 $x_B=x_S=0.024$, viz. a considerable discrepancy with those predicted by [3].

If we define the segregation coefficient of CdS in PbS as $K=x_S/x_L$, we find $K=0.79$ for sample 54 and $K=0.8$ for sample 57.

Reference [4] mentioned that in DTA experiments the $Pb_{1-x}Cd_xS$ ($x=0$ included) samples exhibited a tendency to supercooling. This behaviour might be understood if we accept the supposition [12] (based on viscosity measurements) that at

Table I

CdS concentrations in some synthetic $Pb_{1-x}Cd_xS$ crystals

Sample number	Preparation procedure (*)	Initial CdS conc. [mol%] x_i	Determined CdS conc. [mol%]		Obs.
			Top x_T	Bottom x_B	
28	VPT	5.46	5.6	5.6	p-type
58	VPT	5.6	5.66	5.66	p-type
54	BT	5.68	12.70	4.48	p-type
57	BT	3.0	3.6	2.40	n-type
30	BT	8.06	—	5.8	p-type

* VPT-vapour phase technique
BT-Bridgman technique

$T > T_L$ (but not much higher) the liquid PbS has another structure compared with f.c.c. of the solid PbS. If the liquid PbS cools, this "liquid structure" (LS) still holds below T_L. In that our liquid $Pb_{1-x}Cd_xS$ solutions exhibited thermic behaviour similar to PbS, we are prone to think that a similar LS may exist in these solutions too. From here ensue several consequences:

Firstly, the points reported in [3] as liquidus points might in fact be something else and in particular might be points on the "liquid crystallization" line of the system. This hypothesis deserves a separate study and probably XSDA at high temperatures will answer this question.

Secondly, in BT experiments the speed of cooling is much lower than that in DTA. For instance, in the present work the speed of cooling in BT experiments was 5—10 times less than that reported in [4]. Therefore, the crystallization process in BT experiments is closer to the true thermodynamic equilibrium than in DTA. Hence BT experiments deserve more confidence. In Fig. 2 we plotted the liquidus line for the solidus line known from [4].

If we had accepted [3], the eutectic point would have had its coordinates $T_E = 1055\ °C$ and $x_E \cong 0.32$. If we accept [4], $T_E = 1080\ °C$, $x_E = 0.2$ (Fig. 2). At the same time if we accept the eutectic temperatures mentioned before, we may extrapolate the solidus line reported in [2] and find $x_E \cong 0.5$ for $T_E = 1055\ °C$ and $x_E \cong 0.6$ for $T_E = 1080\ °C$, in disagreement with [3] and [4].

We consider that these great differences are due to the oxygen diffusion through ampoule walls at high temperatures because in [2] no special care was taken in this respect. We have also to express some reservations concerning the DTA work reported in [3] as the authors mentioned that at $x > 0.08$ CdS precipitates as a second phase, while the phase diagram drawn in this work has $x_M = 0.32$.

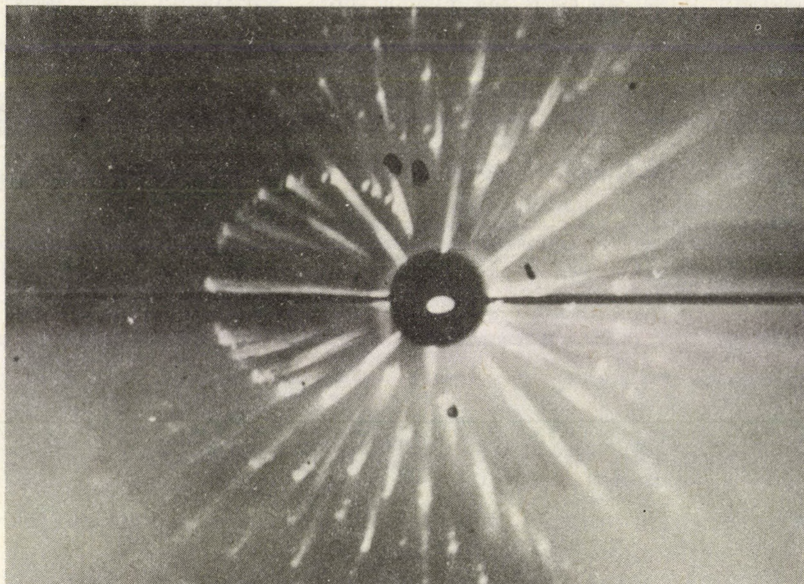
Thirdly, the solid and liquid lines on the $T-x$ phase diagram are very close to each other confirming the supposition [4]. We have to note that in the homologous PbTe–CdTe system the distance between the liquidus and solidus lines is also small [13].

Fourthly, the crystallization in BT processes comes close to equilibrium without being entirely in this regime. This means that due to the fact that LS holds below T_L and in addition to the presence of CdS it may influence the rearrangement of atoms to form the solid f.c.c. lattice, and the crystallization of $Pb_{1-x}Cd_xS$ occurs under not well controllable conditions. As a result of this a tremendous amount of crystallite forms initially and in the ingots we will have always several twins.

A Laue pattern from a slide cut from sample 57 is presented in Fig. 4a. It can be seen that big mechanical strains are present and that the slide consisted of several twins.

In Fig. 5a we present a combined photo made with SEM and MBA about a slide cut from sample No. 57.

We remark first that the background of this photo, due to SEM, is dark almost everywhere. This means that the greater part of the surface of the sample is clean and smooth.



a)

b)

Fig. 4. Laue patterns from slides cut from: a) sample No. 57 (BT), exposition 60 hours; b) sample No. 28 (VPT), exposition 20 hours. The diffraction patterns were obtained in transmission with a TUR-M apparatus W-anode, $U = 55$ kV, $I_a = 12$ mA

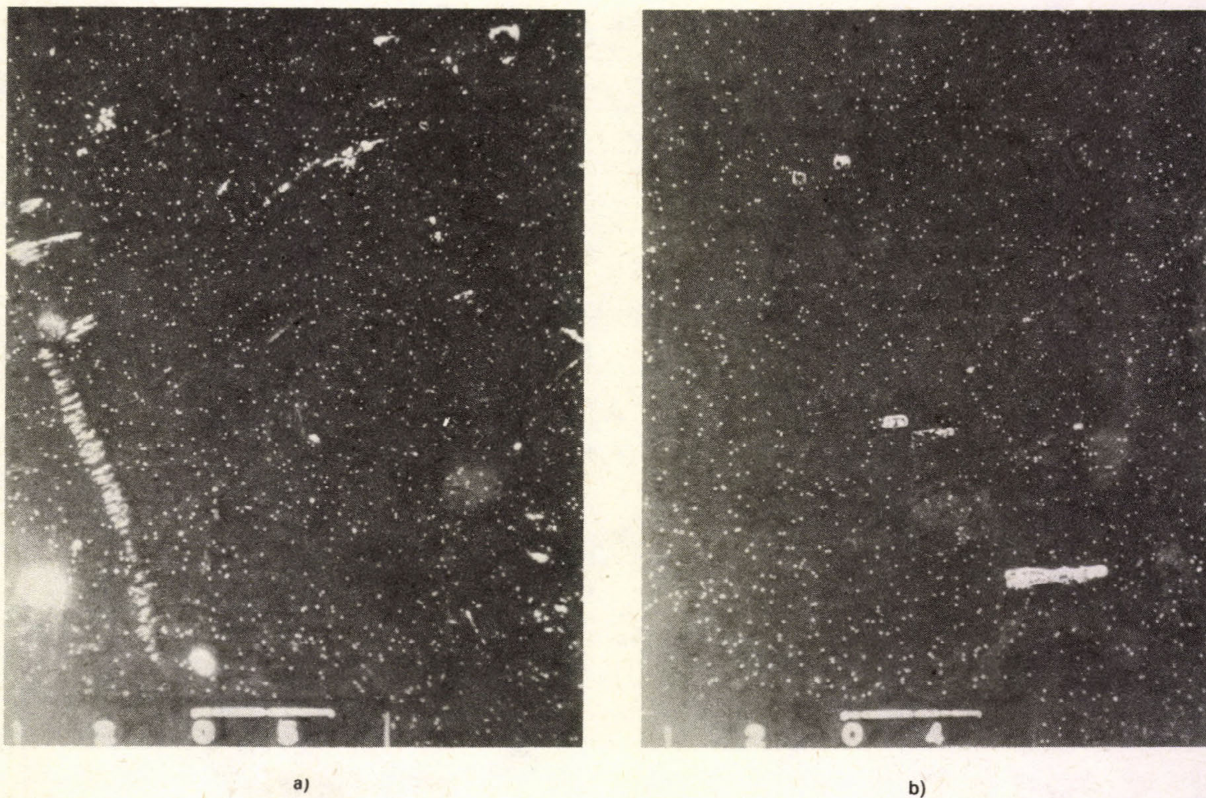


Fig. 5. Photos of Cd distribution in slides cut from: a) sample No. 57 (BT), b) sample No. 28 (VPT). Magnification 100×. White points mark cadmium distribution using $Cd_{L\beta 2}$ characteristic radiation

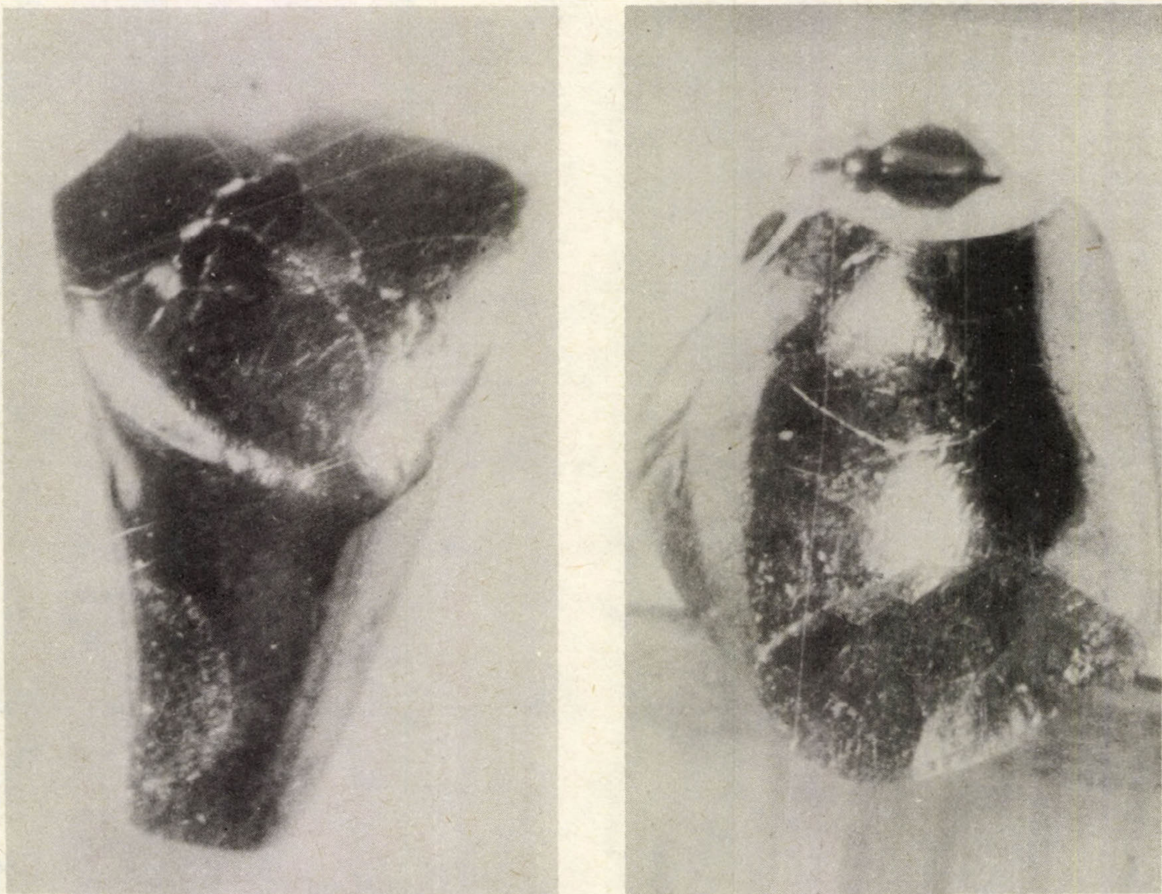


Fig. 6. Two $Pb_{1-x}Cd_xS$ crystals obtained by vapour phase technique. a) sample No. 28. b) sample No. 58. The size of crystals is several mm

The bright isles in the upper part of the figure might be some dust on the surface while the large bright line in the bottom left hand corner must be a twin border.

The superposed picture of Cd distribution obtained by MBA is presented by white small points. We see that cadmium is fairly uniformly distributed and therefore the sample is homogeneous over quite a large area.

3.2. VPT samples

Two VPT samples are shown in Fig. 6.

We can see that these samples also consist of 3–4 big twins with quite well defined but irregular-shaped borders.

Table I shows that in our VPT samples the cadmium concentration is very close to that of the initial mixtures. We think this is natural taking into account the adopted technology in VPT.

Figure 4b shows the Laue pattern obtained from a slide cut from sample No. 28. The spots are well defined. This means that there are no mechanical strains. Our attempts to interpret this Laue pattern with a view to determining the orientation of crystallographic axes relative to the surface of the slide suggest that the slide consists of several twins.

Figure 5b is a combined photo made with SEM and MBA about the same slide (sample No. 28) used to obtain Fig. 4b. We see that in the explored field there are no twin boundaries, the cadmium is fairly homogeneously distributed along the sample.

It is interesting to note that our VPT samples showed only p-type electrical conductivity, regardless of which type of starting PbS was employed. This remark opens the problem of the exact phase diagram of the PbS—CdS system, with influences of deviations from stoichiometry. This problem does not seem to have been researched, or possibly it is just unpublished so far.

4. Conclusions

1. The VPT seems promising and enables good homogeneous p-type crystals to be grown. In these crystals the CdS concentration is near to that of polycrystalline starting mixtures.

2. BT ingots showed a moderate variation of CdS concentration due to a segregation coefficient K smaller than unity—evaluated as $K = 0.8$. A small distance between the solidus and liquidus lines on the $T-x$ phase diagram of the system makes K close to unity.

3. It seems that BT may be very useful in establishing the position of the liquidus line for systems with known position of solidus line and which tend to supercool.

4. SEM and MBA showed a good average homogeneity of Cd distribution in our samples.

Acknowledgements

The author is indebted to Mrs. T. Makarenko for the English version of this work and to Mr. Doru Iurian for his essential technical support.

References

1. P. M. Behtke and P. B. Barton, U. S. Geol. Survey Prof. Pap., *424* B, 266, 1961.
2. P. M. Behtke and P. B. Barton, Amer. Mineralogist, *56*, 2034, 1971.
3. A. R. Calawa, J. A. Mroczkowski and T. C. Harman, Mass. Inst. Technol., Lincoln Lab. Rep. 1971, V. 3, p. 131, mentioned in "Poluprovodnikovie tverdie rastvori i ih primenenie" spravochnie tablitzi, Voennoe Izd. Min. Oboroni SSSR, Moskva 1982, p. 145.
4. P. Şteţiu, Phys. Stat. Solidi, (a), *15*, K 19, 1973.
5. A. R. Calawa, J. A. Mroczkowski and T. C. Harman, J. Electron. Mat., *1*, 191, 1972.
6. K. W. Nill, A. J. Strauss and F. A. Blum, Appl. Phys. Lett., *22*, 677, 1973.
7. V. D. Lawson and S. Nielssen, Preparation of single crystals, Butterworths, London, 1958.
8. P. Şteţiu — to be published.
9. J. Bloem and F. A. Kröger, Z. Phys. Chem. (Frankfurt), *7*, 1, 1956.
10. W. W. Pipper and S. J. Polich, J. Appl. Phys., *32*, 1278, 1961.
11. A. Guinier, Théorie et technique de la radiocristallographie, Dunod, Paris, 1956.
12. N. N. Glagoeva, A. N. Krastovnikov, V. M. Glazov, Izv. A. N. SSSR, Ser. Neorg. Mat., *4*, 1980, 1968.
13. A. J. Rosemberg, R. Grierson, J. C. Wooley and P. Nicolic, Trans. AIME, *230*, 342, 1964.

Acta Physica Hungarica, 57 (3—4) pp. 335—346 (1985)

BUNCHING OF MONOMOLECULAR STEPS DURING GROWTH AND EVAPORATION OF CRYSTALS

L. MALICSKÓ

Research Laboratory for Crystal Physics
Hungarian Academy of Sciences
1502 Budapest, Hungary

The growth and evaporation layers formed on KCl, KBr and sodium thiosulphate crystal faces grown or evaporated under different conditions (in solutions, melts or in high vacuum) were studied by optical and transmission electron microscopic methods. Different stages of bunching of the monoatomic steps, visualized by the surface gold decoration replica technique, were observed. The probability density functions of the heights of step bunches (growth and evaporation layers) were approximately determined. The results are presented and discussed.

1. Introduction

The atomically smooth faces of polyhedral crystals generally grow, dissolve or evaporate, by single monomolecular steps formed by random two-dimensional nucleations (islands or holes) and/or by dislocations emerging on the faces. The growth and decomposition processes show certain similarities. Under actual conditions the monomolecular steps often coalesce into step bunches forming layers of different thicknesses and profiles. Since, during growth, the bunching processes influence the homogeneity and the as-grown dislocation structure of the crystal products [1, 2], studies of step bunching are of fundamental and practical importance.

Several theoretical models describing step bunching are known [3—5]. According to these models the bunching begins with fluctuations in the distances between neighbouring steps belonging to a monomolecular step train. Such step distance fluctuations were also experimentally observed by optical and electron microscopy of monomolecular growth and evaporation step trains [6, 7]. These fluctuations are the consequences of impurity effects [8, 9] as well as of temperature fluctuations [10]. The later stages of the bunching of monoatomic steps into layers of different thicknesses were also studied with different optical microscopic methods (see, e.g. [6, 11, 12]). In these studies a reciprocal relation was empirically found between the tangential moving rate and the thickness of the layers [6, 11].

The aim of this work is to collect experimental results on the stages of the bunching of evaporation and growth steps with the help of different observation methods and to draw attention to some common statistical features among the observed processes.

2. Experiments

2.1. For the *static* investigations of the monoatomic step structures on crystal faces the surface gold decoration replica technique is a very good tool [8, 13]. With this technique the various disturbances in the motion of monoatomic steps and the different stages of the step bunching could be studied on KCl single crystals in a conventional transmission electron microscope. For the investigations macroscopic KCl single crystals grown from Pb doped aqueous solutions by temperature reducing or by the solvent evaporation technique [14, 15], and undoped KCl single crystals grown from the melt by the Kyropoulos method, were used. Both original {100} growth faces and {100} cleavage faces evaporated previously in high vacuum, were investigated. The temperature and the duration of the vacuum evaporations were varied between 670 and 770 K and between 20 and 120 minutes at pressures of about 10^{-3} Pa.

The steps developing on the faces during the evaporation may lead to disturbance. On the basis of the observable disturbances between a chosen step and the surrounding known monoatomic steps, the heights of chosen steps can be measured in $a/2$ atomic units, where a is the lattice spacing ($a = 0.628$ nm for KCl). In this way the steps of heights between 1 and about 30 atomic units could be well-studied.

For the static observation of the later stages of step bunching the structures of the growth steps on the original {100} growth faces were studied by optical microscopy. For this purpose, after growth had finished the remainders of the solution were quickly and carefully removed from the crystals with fine filter paper.

The growth steps on the growth faces of the crystals are visible by reason of dark contrast lines in the optical microscope at suitable illumination. From the measurable widths of the dark contrast lines the relative step heights could be determined. The average heights of more thin single steps and the heights of individual thick steps could be measured directly. Thus, by comparing the relative and the directly measured values, step heights from 50 (measurable limit) up to several 1000 nm could be measured. By decorating the original growth faces at 450 K the fine structure of the polyatomic growth steps could also be studied.

2.2. In optical microscopic in-situ experiments the dynamics of the step bunching was studied. The motion of the growth layers on thin side faces of small plate-like KBr and sodium thiosulphate crystals growing in supersaturated solution drops or in undercooled melt drops, respectively, was observed and filmed [11]. On the microfilms the heights of growth steps could be measured. They varied in the range between 0.5 and 25 μm.

2.3. The height data measured both on the evaporation and on the growth steps formed under given constant conditions, were statistically evaluated enabling the probability density functions of the heights of step bunches (evaporation and growth layers) to be approximately determined.

3. Results and discussion

According to the BCF theory [16] the dislocations emerging onto the crystal faces produce trains of equidistant monomolecular steps during the growth or decomposition of atomically smooth faces. In the case of the evaporation of nominally pure alkali halide crystals this fact was verified [7, 8]. For comparison Fig. 1 shows such trains of decorated equidistant monoatomic steps formed around the emergence points of a pair of screw dislocations (a) and of an edge dislocation (b) with a/2 $\langle 110 \rangle$

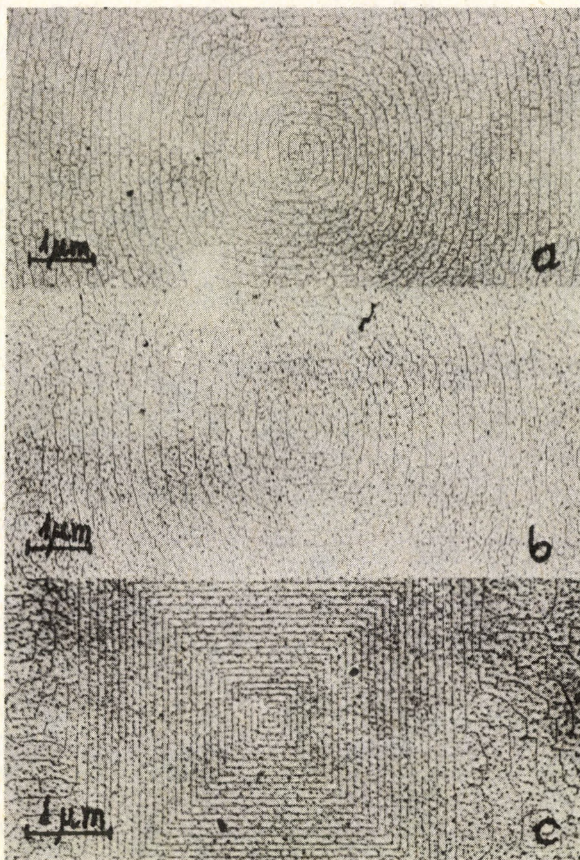


Fig. 1. TEM micrographs of trains of equidistant steps a — of monoatomic heights, formed around the emergence point of a pair of screw dislocations with $\frac{a}{2}\langle 110 \rangle$ Burgers vectors, b — of monoatomic heights, formed around the emergence point of an edge dislocation with $\frac{a}{2}\langle 110 \rangle$ Burgers vector, c — of biatomic heights, formed around the emergence point of an $a\ \langle 100 \rangle$ type screw dislocation on one of the $\{100\}$ cleavage faces of a KCl crystal after thermal treatment and subsequent decoration in high vacuum

type Burgers vectors and a train of equidistant biatomic steps formed around an a ⟨100⟩ type screw dislocation (c) on the cleavage faces of KCl crystals after suitable thermal treatments in high vacuum. The explanation of such evaporation step configurations is well known [8].

With similar step systems (step trains) well measurable statistical fluctuations were found in the step distances [7]. These fluctuations occur due to temperature and/or impurity concentration fluctuations on the evaporating faces of the crystals.

The local fluctuations of the temperature and the impurity concentration may lead to different local disturbances in the motion of mono- and biatomic steps. For instance, Fig. 2 demonstrates surface ranges in which the motion of the atomic steps is disturbed. The main direction of the step motion during the evaporation is marked by the arrow v_t. Figure 2a shows small irregularities in the contours of monoatomic steps (see the arrow); they are caused by single foreign atoms or very small clusters of nonvisible foreign atoms [17]. In Fig. 2d, e and f dark granulates and spots of different sizes (marked with arrows) causing local disturbances in the motions of more monoatomic steps, are seen. They are already able to lead to the bunching of the steps in the surrounding areas (Fig. 2f). The arrows in Fig. 2b and c show small areas where the monoatomic steps have already been bunched into step sections of heights of 2 or 3 atomic units due to nonvisible fluctuations.

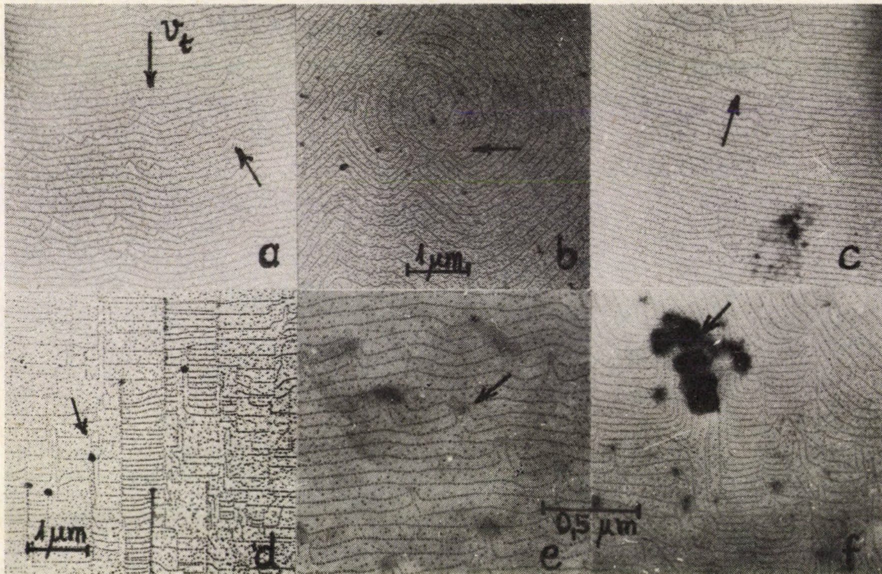


Fig. 2. TEM micrographs showing different disturbances in the motion of monoatomic steps on {100} faces of KCl crystals: — small disturbances caused mainly by single nonvisible foreign atoms (a) or by enrichment of foreign atoms in small areas (dark spots in e), — greater disturbances leading to local bunching of the monoatomic steps, caused by nonvisible fluctuations (b, c) or by small foreign particles (dark granulates in d and f), as marked by the arrows. The arrow v_t marks the main direction of the step motion

Besides the three simple types of evaporation step systems seen in Fig. 1 evaporation systems of so-called roughly quadratic steps can also be found on the faces. In Fig. 3 some stages of the development of such step systems are demonstrated (a—c).

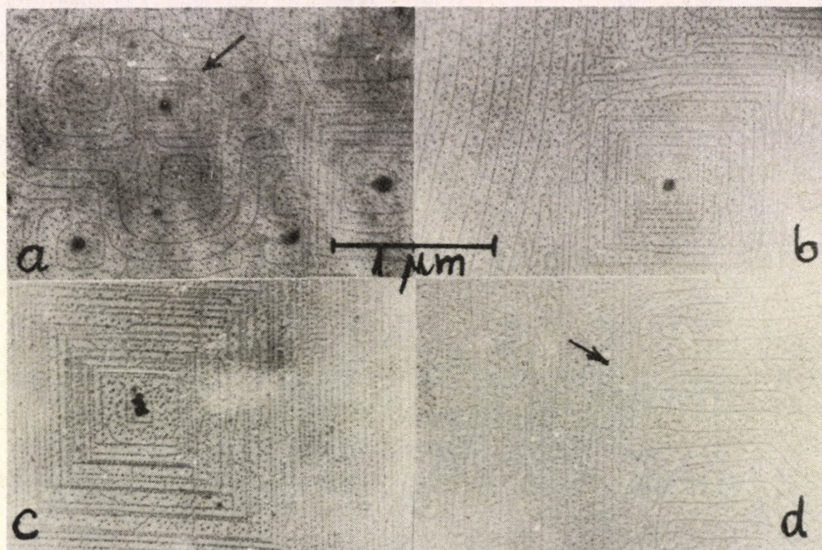


Fig. 3. TEM micrographs showing developed stages of roughly quadratic step systems (a, b and c) and atomic steps each bunching into straight steps of a quadratic step system (d)

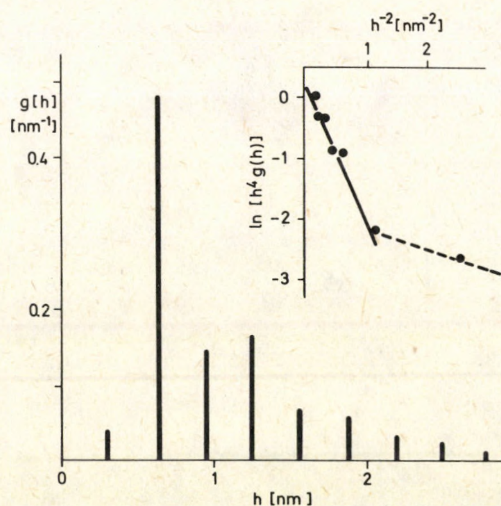


Fig. 4. Histogram of (discrete) heights of straight steps belonging to a well developed quadratic step system on {100} face of a KCl crystal

These systems are formed around step sources of step emission activities greater than the step emission activities of the single dislocations, e.g. around strongly poisoned dislocations [8, 18]. Great activities result in great average step densities around these step sources. Due to the great step densities and the mentioned possible disturbances, the bunching of mono- or biatomic steps into steps of greater heights occurs already in the beginning stages of the evaporation (Fig. 3a). The bunching becomes more and more marked in later stages (Fig. 3b and c).

On the basis of the number of surrounding atomic steps each coalescing into straight steps (see the arrow in Fig. 3d) the heights of steps belonging to quadratic systems can be measured. Judging by the observations the step heights in these systems vary randomly. Figure 4 shows a typical histogram of the heights of straight steps belonging to a well developed quadratic system.

By calculating an average monoatomic step density $\bar{d}$ for the surface areas of each quadratic step system, an approximately proportional relation could be found between the average heights $\bar{h}$ of the real steps and these average monoatomic step densities $\bar{d}$ (Fig. 5). The radii of the circles are proportional to the number of data yielding the measurement points. This relation means that the probability of step bunching increases with the rising step density. From Fig. 5 it can be concluded that above a step density of about 2.10^5 cm^{-1} the tendency to bunch becomes considerable. This density value corresponds to an average step distance of about 50 nm. This means that the

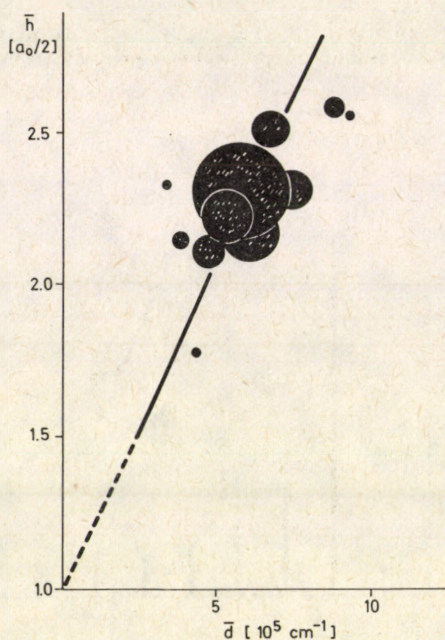


Fig. 5. Relation between average step height $\bar{h}$ and average step densities $\bar{d}$ measured on roughly quadratic step systems developed during different stages of evaporation

considerable overlapping of the surface diffusion fields belonging to each individual atomic step, is average at this step distance. Consequently, this step distance may be approximately equal to the mean surface diffusion path of the adatoms. This value of 50 nm, estimated in this way for the mean surface diffusion path in the case of KCl, is commensurable with the value of about 200 nm found for the free evaporation of NaCl at 673 K [19].

The KCl single crystals grown from solutions are mainly bordered by {100} and {111} faces as demonstrated by the macrograph in Fig. 6a. On the original {100} growth faces of the crystals nearly concentric step systems (Fig. 6b) caused by dominating growth centres are observable in an optical microscope [18]. Because of the activity of the growth centres, the {100} type growth faces are fully covered by such layers. Besides the rough steps (Fig. 6c) the bunches of finer steps of mono- or more atomic heights can also be revealed by decoration in the TEM (Fig. 6d).

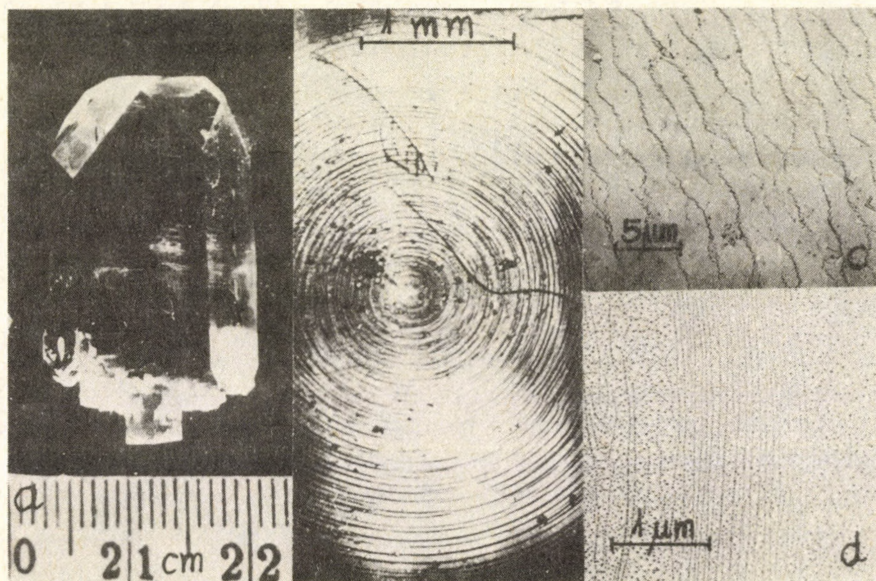


Fig. 6. Rough and fine structures of growth layers on {100} growth faces of KCl crystals:
a — macrograph of KCl single crystal bordered by {100} and {111} faces
b — optical micrograph of growth layers of polyatomic thicknesses on {100} growth face
c — TEM micrograph of fine growth steps of polyatomic heights
d — TEM micrograph of a fine growth layer bunched from decorated atomic steps.

The optical microscopic observations also show the heights of such growth steps to vary randomly between 50 and several 1000 nm. Figure 7 shows a typical histogram of the heights of growth steps measured by optical microscopy on one of the {100} faces of a solution grown KCl crystal. In contrast to the previous case (Fig. 4) no maximum appears. It means that the maximum should lie in the range below 400 nm.

The behaviour of the later stages of the step bunches was studied in in-situ optical microscopic experiments. During the growth of small plate-like KBr crystals within supersaturated solution drops fine moving layers were observable on the thin {100}

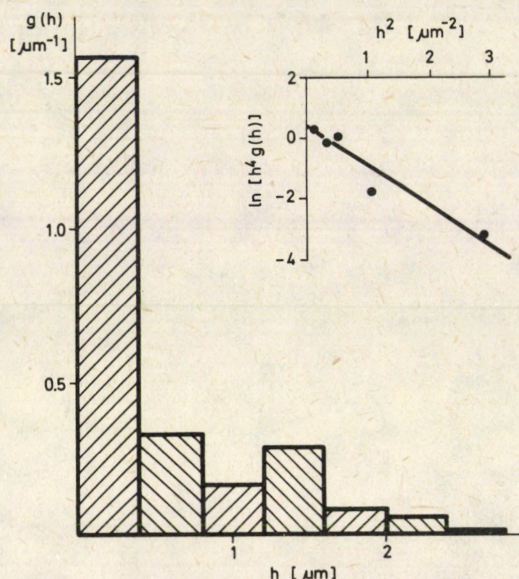


Fig. 7. Histogram of heights of growth steps measured by optical microscopy on {100} face of a solution grown KCl crystal

type side faces. This is demonstrated by two successive photos of a ciné film in Fig. 8a and b. The shifting between the small perpendicular lines drawn-in marks the motion of one of the visible layers. v_n and v_t represent the normal growth rate of the side faces and the tangential rate of the growth layers, respectively. In the case of different individual crystals, v_n and v_t varied between 0.3 and 4 μm/s or between 0.3 and 50 μm/s, respectively.

The same moving layers were observed on the thin {110} type side faces of plate-like sodium thiosulphate crystals grown within undercooled melt drops (Fig. 9a and b). In these cases, however, the normal and tangential rates amounted to approximately 100 times greater than those in the case of KBr crystals. In Figs 8 and 9 the bright parallel lines within the crystals mark the layered internal defect structure caused by the layer growth. In both latter cases the corner ranges of the small crystals were the sources of the observable growth layers as described by Kossel [20].

The thicknesses of the growth layers observed either on the KBr or on the sodium thiosulphate crystals, varied randomly. Figure 10 shows a typical histogram of the layer thicknesses measured on small KBr crystals grown at the same average normal rate of 1.5 μ/s in supersaturated solution drops. A similar histogram is given in

Fig. 8. Optical micrographs showing growth layers moving (see the marks) along the thin side face of a small plate-like KBr crystal grown within a supersaturated solution drop

Fig. 11 which presents layer thicknesses measured on a sodium thiosulphate crystal grown at an average normal rate of 150 μm/s in an undercooled melt drop.

Having found the thicknesses of the evaporation and growth layers formed in different stages of the bunching of monomolecular steps to be probability variables, the possible common description of their statistical behaviour was attempted. Between the tangential moving rate v_t and the thicknesses h of the growth layers the following empirical relation was found [6, 11]:

$$v_t = \frac{A}{h} + B \approx \frac{A}{h}, \tag{1}$$

where A and B are constants and the approximation is valid if h is small enough. It was also shown that the statistical fluctuations of the tangential moving rates of the growth

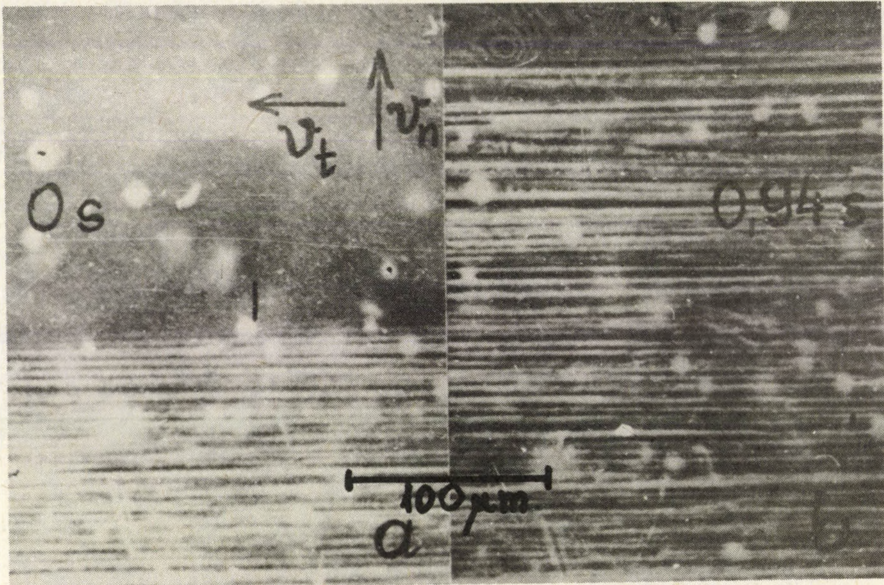


Fig. 9. Optical micrographs showing growth layers moving along the thin side face of a plate-like sodium thiosulphate crystal grown within an undercooled melt drop

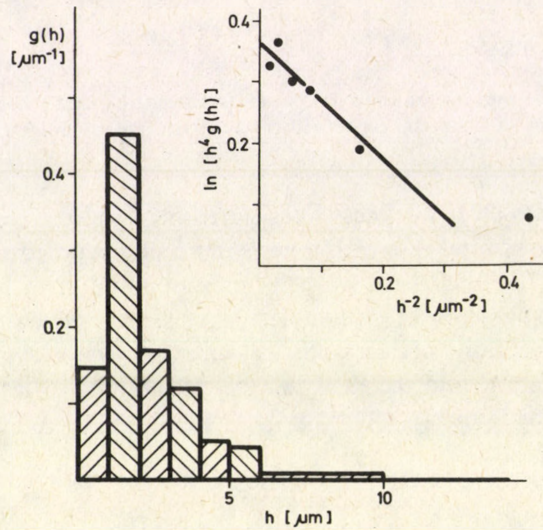


Fig. 10. Histogram of thicknesses of growth layers measured on KBr crystals grown at the same average normal rate in supersaturated solution drops

layers can approximately be described by a Maxwellian type density function [21]:

$$f(v_t)=a\cdot v_t^2 \exp\{-bv_t^2\}, \tag{2}$$

where a and b are constants. According to the known rules for the probability variables standing in functional relations Eq. (2) can be transformed into the density function of the layer thicknesses using the approximate form of Eq. (1), viz.

$$g(h) = \frac{\alpha}{h^4}\exp\left\{\frac{\beta}{h^2}\right\}, \tag{3}$$

where $\alpha=aA^3$ and $\beta=bA^2$ are constants. In accordance with Eq. (3) the insets of the histograms in Figs 4, 7, 10 and 11 show the term of $\ln[h^4\cdot g(h)]$ versus $1/h^2$. The

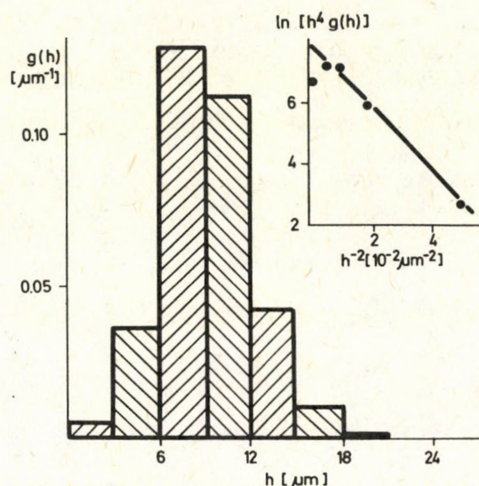


Fig. 11. Histogram of thicknesses of growth layers measured on a sodium thiosulphate crystal grown in undercooled melt drop

straight lines fitted to the measurements points verify that Eq. (3) approximately describes the common probability density function of the thicknesses of layers being formed under very different experimental conditions. The results suggest that the bunching of the height steps from atomic up to μm scales occurs according to common basic principles.

Acknowledgements

The author thanks Mrs. M. Kiss and Mrs. S. Tóth for their technical assistance; acknowledgement is made to the State Office for Technical Development (OMFB) for supporting this work.

References

1. J. Friedel, Dislocations, Pergamon Press, Oxford, 1964, pp. 203—208.
2. F. R. N. Nabarro, Theory of Crystal Dislocations, Clarendon Press, Oxford, 1967, pp. 327—339.
3. R. L. Parker, in Solid State Physics Ser. Vol. 25. Acad. Press, New York, 1970, pp. 151—299.
4. P. Bennema and G. H. Gilmer, in Crystal Growth: An Introduction, Ed. P. Hartman, North-Holland, Amsterdam, 1973, pp. 263—327.
5. A. A. Chernov, in Crystal Growth and Characterization, Ed. J. B. Mullin, North-Holland, Amsterdam, 1975, pp. 33—52.
6. G. G. Lemmlejn, E. D. Dukova and A. A. Chernov, Kristallografiya, *2*, 428, 1957.
7. K. W. Keller, Dissertation, Halle, GDR 1968.
8. H. Bethge, Phys. Stat. Sol., *2*, 3 and 755, 1962.
9. L. Malicskó, Periodica Polytechnica, *16*, 141, 1972.
10. H. Bethge, H. Höche, D. Katzer, K. W. Keller, P. Bennema and B. van der Hoek, J. Cryst. Growth, *48*, 9, 1980.
11. L. Malicskó, Phys. Stat. Sol., *3*, 1547, 1963.
12. F. H. M. Mischgofsky, Dissertation, Delft, 1977.
13. G. A. Bassett, Phil. Mag., *3*, 1042, 1958.
14. L. Malicskó and L. Jeszenszky, J. Cryst. Growth, *7*, 13, 1970.
15. L. Malicskó and L. Jeszenszky, J. Cryst. Growth, *15*, 243, 1972.
16. W. Burton, N. Cabrera, F. Frank, Phil. Trans., *A243*, 299, 1951.
17. H. Stenzel, Diplomarbeit, Halle, GDR, 1962.
18. L. Malicskó, Acta Phys. Hung., *53*, 59, 1982.
19. H. Bethge, K. W. Keller, Zs. Naturf., *15a*, 271, 1960.
20. W. Kossel, in Quantentheorie und Chemie, Ed. H. Falkenhagen, Leipziger Vorträge, 1928, pp. 1—46.
21. L. Malicskó, Krist. u. Technik, *5*, K15, 1970.

Acta Physica Hungarica, 57 (3—4) pp. 347—354 (1985)

MACRO- AND MICROANALYTICAL INVESTIGATIONS OF IMPURITIES DURING PRODUCTION OF TeO_2 SINGLE CRYSTALS

L. MALICSKÓ, I. CRAVERO and L. ÖTSI

Research Laboratory for Crystal Physics
Hungarian Academy of Sciences
1502 Budapest, Hungary

On samples chosen from the main production steps of tellurium dioxide single crystals macroscopic and microscopic analyses were carried out using flame atomic absorption spectroscopy and energy dispersive X-ray spectroscopy connected to a scanning electron microscope. The propagation of the impurities from the raw material into the single crystals grown by the Czochralski technique is followed and the results are discussed.

1. Introduction

Because of their excellent acousto-optical properties, TeO_2 single crystals are of practical importance; practical applications require single crystals of high quality. According to earlier works, Czochralski-grown TeO_2 single crystals contain visually imperfect ranges in which gas and solid inclusions are randomly formed during the growth in the presence of various impurities [1—5]. In a recent work the enrichment of different impurity elements and their close connection with the gas bubbles within the visually imperfect ranges of TeO_2 single crystals were directly shown by optical and scanning electron microscopy combined with electron beam X-ray microanalysis [6].

The aim of the present investigations was to obtain information about the origin and the passing-on of the undesirable impurities during the main production steps of TeO_2 single crystals, using macroscopic and microscopic analyses.

2. Experiments

For the investigations the following samples were chosen from the main steps of the TeO_2 single crystal production in our laboratory:

— Cleavage pieces of 5N pure (Aluterv-FKI, Hungary) cast metal tellurium (M) containing single crystalline grains of up to a few cm^3 in size.

— Fragments from the upper part of a zone-refined tellurium ingot (ZM) assumed to be enriched with certain impurities. The vertical zone-refining was carried out in a conical quartz crucible, a little above the melting point of the tellurium (725 K).

— Crystalline reaction products (P) precipitated within the solution during the oxidation of metallic tellurium with nitric acid. After suitable heat treatment these reaction products serve as *starting material* for growing the tellurium dioxide single crystals.
— Pieces from the visually imperfect parts of TeO_2 single crystals (C) grown by the Czochralski technique [4, 7].
— Remaining impurity phases (R) extracted from imperfect TeO_2 crystals by treatment in conc. HCl [4].

In samples of *macroscopic* quantities the average concentrations of K, Mg, Fe and Ni were determined by flame atomic absorption spectroscopy (FAAS) using a SPECTROMOM 190 A spectrophotometer working with an air-acetylene flame. Because of the high sensitivity of this method the specimen materials, after dissolution, were directly measured without any previous enrichment or separation procedure. The disturbing effect of the tellurium as the majority component could be avoided by the addition of disodium hydrogen citrate. The composition of the basic water solution used for the calibration and for the measurements of the samples in the spectrophotometer was: 0.1 M Te + 1.0 M HCl + 10^{-3} M disodium hydrogen citrate. The composition of the calibration and the sample solutions was the same and they also contained the elements to be measured in appropriate concentrations. The errors of the measurements depended on the quality of the measured elements but lay below $\pm 10\%$. The detection limits were 10 μM/M for K, Mg and Fe and 20 μM/M for Ni, respectively.

In order to obtain further information about the distribution of the impurities in *microscopic* dimensions, the samples were investigated by imaging and analysing modes in a JSM-35 C scanning electron microscope (SEM) equipped with a KEVEX-7000 energy dispersive X-ray spectrometer (EDS). In the samples defect-free and imperfect places were imaged, analysed and compared. The impurity concentrations were formally computed in atom percentages (at%) related to the Te majority component using a usual standardless analysis program (SLA). These concentration data, however, are suitable only for qualitative evaluations because of the sample structure differences.

3. Results and discussion

The results of the FAAS and EDS analyses are separately listed (Table I) according to the analysing methods and the sorts (St) of samples (M, ZM, P, C and R). The first column contains the symbols of the impurity elements found. The concentrations are given in μM/M units for FAAS and in at% units for EDS data.

According to the FAAS analyses (column M—F in Table I) the 5N Te raw material in average stands for its specification. This material consists of large single crystalline grains. The secondary electron (SE) image of two neighbouring large columnar grains (A and B) is shown in Fig. 1. Defects could be observed (see arrows)

Table I

Foreign elements found by flame atomic absorption spectroscopy (F) and by energy dispersive X-ray spectrometry (E) in specimens chosen from different main steps connected with the production of TeO_2 single crystals, and the element concentrations in μM/M Te (F) or in at% units (E). The sorts (St): M — metallic Te, ZM — metallic Te enriched in impurities by zone-refining, P — crystals precipitated in nitric acid solution, C — TeO_2 single crystals grown by the Czochralski method and R — rests after dissolving the imperfect crystalline materials

St	M		ZM		P		C		R	
El	F	E	F	E	F	E	F	E	F	E
Mg	10	0.3	100	—	17	3.1	85	14.2	20	—
K	10	—	<10	4.3	<10	—	<10	10.6	10	1.2
Fe	<10	—	250	0.4	4200	—	3000	1.4	90	1.9
Ni	<20	0.3	<20	—	85	—	70	—	50	1.7
Al		—		—		2.5		77.5		3.0
Si		1.0		—		5.2		48.2		71.0
S		0.8		—		2.2		17.2		6.3
Cl		1.0		0.7		2.4		6.0		9.7
Pt		—		—		—		9.2		2.2

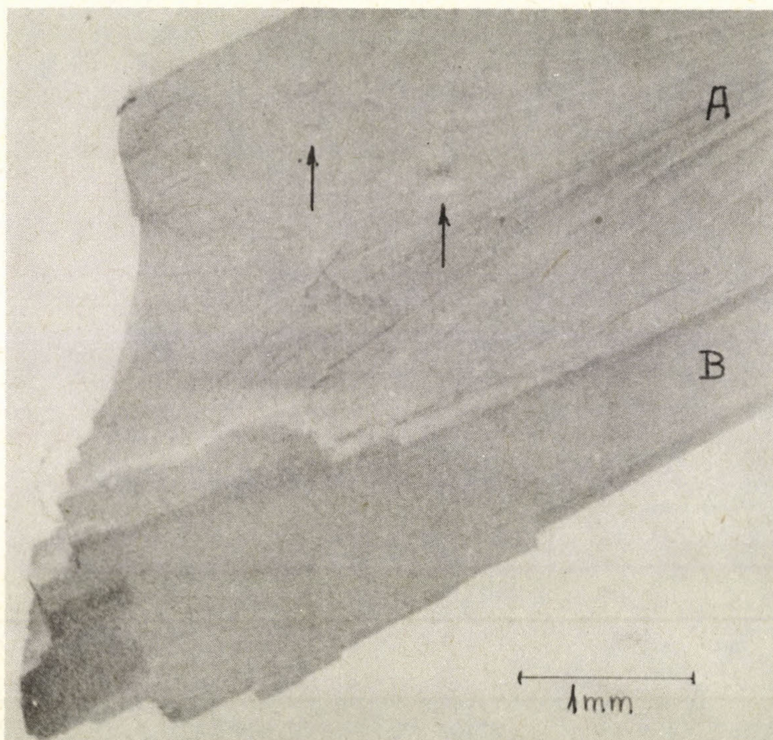
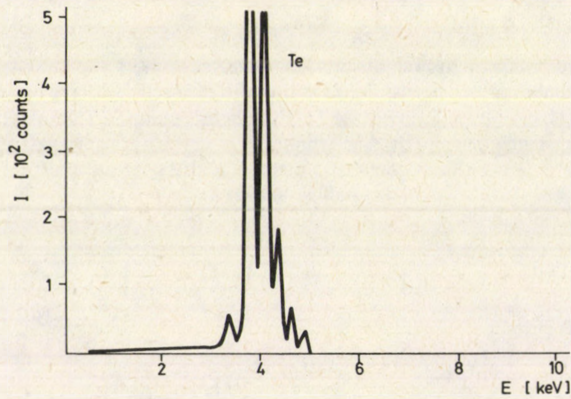
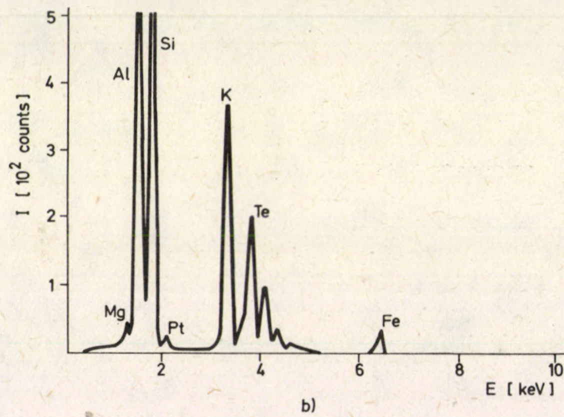


Fig. 1. Secondary electron image of a cleaved piece from metallic Te raw material. The arrows mark defects in the inside of one of the two large single crystalline grains (A and B)





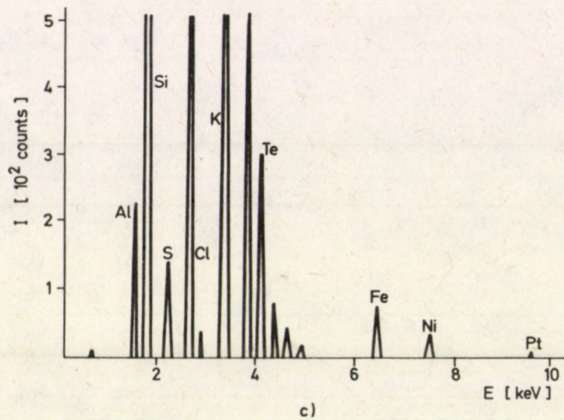


Fig. 2. Energy dispersive X-ray spectra. Made on

a — defect-free range of a Czochralski-grown TeO_2 single crystal, showing only the characteristic lines of Te. This spectrum is typical for defect-free ranges of tellurium grains, solution- and melt-grown intermediates, and the single crystalline product;

b — imperfect range of a TeO_2 single crystal, showing characteristic lines of some impurity elements in addition to Te. The spectrum is typical for imperfect ranges of metallic Te, solution- and melt-grown intermediates, and the end product;

c — piece of dissolution remainder showing the characteristic peaks of impurity elements

inside the single grains. In the defect-free places of the grains only the characteristic X-ray peaks of Te can be detected with EDS as is demonstrated by a background corrected X-ray spectrum in Fig. 2a. However, in the defect areas, marked in Fig. 1, and along grain boundaries the characteristic peaks of certain impurity elements also appear together with the Te X-ray lines. The impurity elements detected in M-samples are listed in column M—E of Table I. From the results obtained with FAAS and EDS it can be concluded that the impurities in the raw material are mainly present as smaller or larger inclusions formed during the casting process within the grains and along the grain boundaries.

The fractured surface of the polycrystalline upper part of a *zone-refined Te rod* is shown in Fig. 3. Here again in the cleavage faces of larger columnar single crystalline grains (see arrows), tiny crystallites can be seen. In these samples the average concentrations of Mg and Fe increase related to the raw metal (column ZM—F in Table I), which means that the distribution coefficients of Mg and Fe are smaller than that between the liquid and the solid phases [8]. The defect-free parts of larger grains

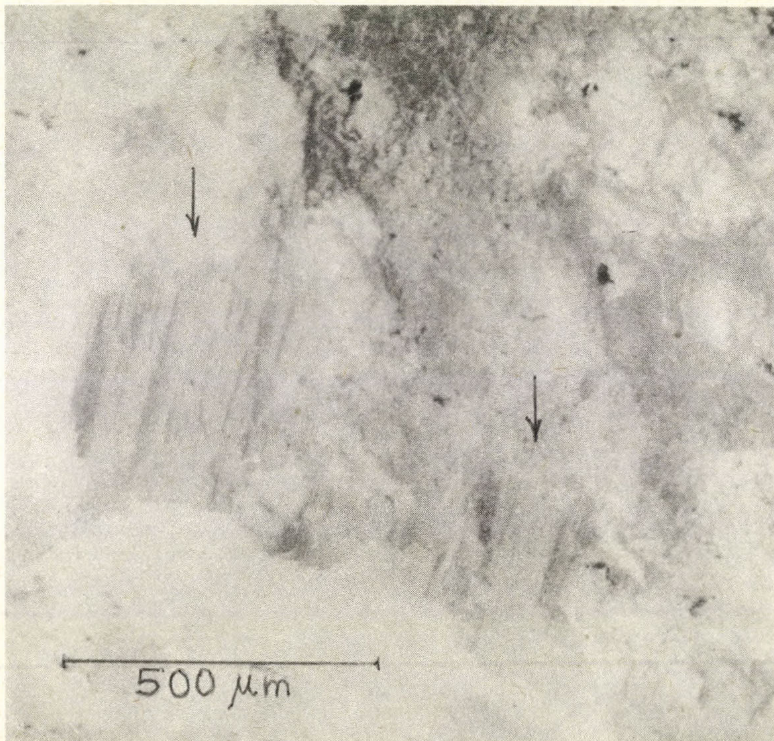
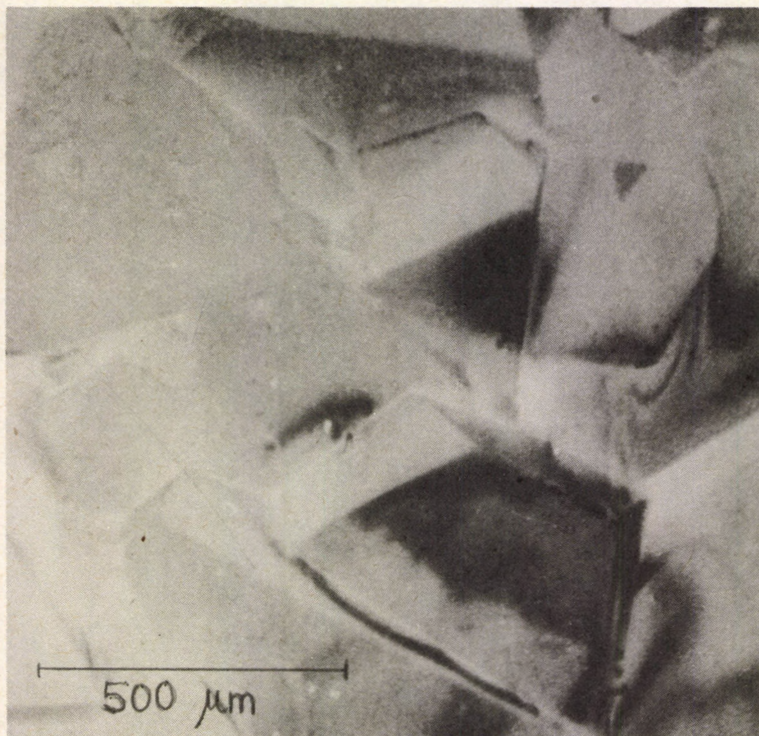
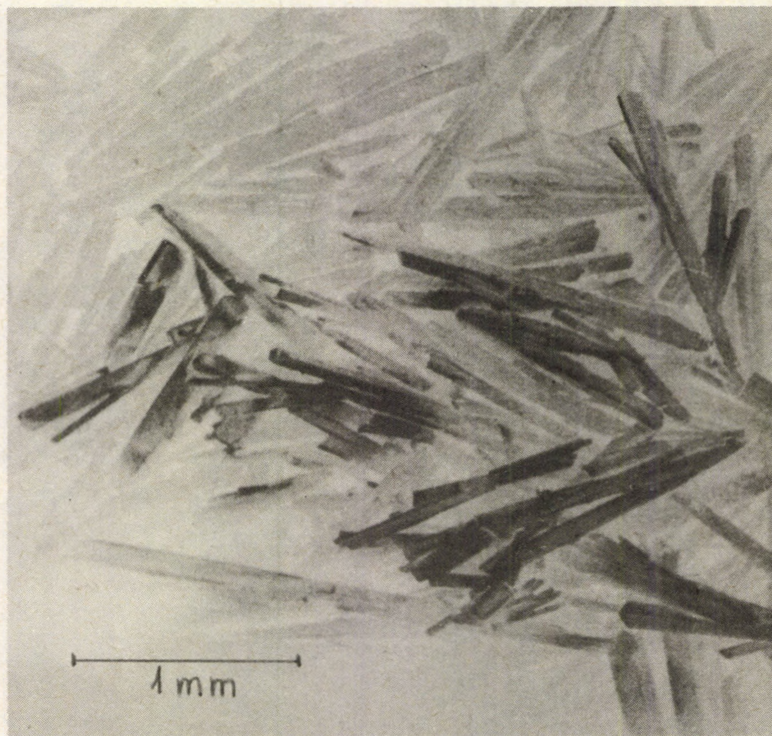


Fig. 3. Fractured surface of upper part of zone-refined tellurium rod, imaged by secondary electrons



a)



b)

Fig. 4. Secondary electron images of
a — block-like TeO_2 crystallites, b — whisker-like basic tellurium nitrates precipitated from nitric acid solution saturated with Te

are practically pure and, of course, the finely grained debris contains mainly the impurities found by EDS (column ZM—E in Table I).

When the metallic tellurium is oxidized with nitric acid, *crystalline reaction products* (P) containing tellurium dioxide precipitate within the solution. The crystals that appear have two typical forms. From about 45% nitric acid solution saturated with Te small block-like TeO_2 crystallites precipitate at about 390 K. The SE image of such crystallites is shown in Fig. 4a. On slowly cooling the nitric acid solution whisker crystals of basic tellurium nitrate appear below 390 K (see the SE image in Fig. 4b).

In the solution-grown crystals (P) the increased concentrations of Ni and mostly of Fe impurity elements are conspicuous (column P—F in Table I). According to the EDS analyses they, in spite of their high average concentrations, were not found to be enriched in any defects contrary to other impurity elements (see column P—E in Table I). Comparison of the EDS data in columns M, ZM and P indicates that the higher concentration values of Al, Si, S and Cl may mark a contaminating effect of the solution procedure.

From the data in column C—F of Table I, it can be seen that the impurity elements found in the previous steps also appear with high average concentrations in the *visually imperfect parts* of the *Czochralski-grown single crystals* (C). In the case of the single microscopic defects observed by SE imaging, some impurity elements (column C—E of Table I) could generally be detected by the microanalyses described earlier in detail [6]. A typical EDS spectrum is shown in Fig. 2b. The suspiciously high concentration of the impurities found in the bulk and locally at the defects, clearly show the important but undesirable role of these impurity elements in the deterioration of the quality of the single crystals. It should be mentioned that in single crystals grown from the melt the platinum crucible material also appears as an additional impurity element [4—6].

The results of the control macroscopic and microscopic analyses of the *dissolution remainders* (R), listed in column R of Table I and demonstrated by an EDS spectrum in Fig. 2c, support again that almost all of the impurity elements mentioned before accompany the technological steps of single crystal production.

Summarizing, it can be concluded that for the incorporation of the impurities into tellurium dioxide the solution phase procedure seems to be the most delicate step.

Acknowledgements

The authors are indebted to the Directorate of the Institute for Solid State Physics and Electron Microscopy of the Academy of Sciences of GDR, Halle/Sa, for enabling the SEM investigations, and to Dr. Th. Krajewski for his valuable help and advice.

This work was sponsored by the State Office for Technical Development (OMFB) in Hungary.

References

1. S. Miyazawa and H. Iwasaki, Japan J. Appl. Phys., *9*, 441, 1970.
2. M. Cerclet, Mater. Res. Bull., *7*, 721, 1972.
3. J. G. Grabmaier, R. D. Plättner and M. Schieber, J. Cryst. Growth, *20*, 82, 1973.
4. I. Földvári, K. Raksányi, R. Voszka, E. Hartmann and Á. Péter, J. Cryst. Growth, *52*, 561, 1981.
5. I. Földvári, R. Voszka and Á. Péter, J. Cryst. Growth, *59*, 651, 1982.
6. L. Malicskó and Th. Krajewski, J. Cryst. Growth, *60*, 195, 1982.
7. F. Schmidt and R. Voszka, Cryst. Res. and Technol., *16*, K127, 1981.
8. F. A. Kröger, The Chemistry of Imperfect Crystals, North-Holland, Amsterdam, 1964, pp. 3—77.

TITLES OF OTHER PAPERS PRESENTED AT HCCG—3:

Invited papers

Metalorganic vapour phase epitaxy of indiumphosphide and related ternaries
K. W. Benz
Physikalisches Institut, Kristallabor, Universität Stuttgart
D-7000 Stuttgart 80, Germany FR

The study of the crystal growth processes by in-situ X-ray topography and optical microscopy
A. A. Chernov, Yu. G. Kuznetsov, I. L. Smol'skii, V. F. Parvov and V. N. Rozhanskii
Institute of Crystallography, USSR Academy of Sciences
Moscow, USSR

Epitaxial growth and physical properties of indiumphosphide
M. H. Pilkuhn
Physikalisches Institut, Universität Stuttgart
D-7000 Stuttgart 80, Germany FR

The role of hydrothermal synthesis in materials preparation
A. Rabenau
Max-Planck-Institut für Festkörperforschung
7000 Stuttgart-80, Germany FR

Relaxed surfaces are rougher
Nai-Ben Ming and Franz Rosenberger
Department of Physics, University of Utah
Salt Lake City, Utah 84112, USA

Interfacial transport in crystal growth, a parametric comparison of convective effects
Franz Rosenberger* and Georg Müller**
* Department of Physics, University of Utah
Salt Lake City, Utah 84112, USA
** Institut für Werkstoffwissenschaften VI, Universität Erlangen-Nürnberg
D-8520 Erlangen, Germany FR

Growth of thin films and molecular beam epitaxy
Ryuzo Ueda
Department of Applied Physics, School of Science and Engineering, Waseda University, 3—4—1 Ohkubo, Shinjuku-ku, Tokyo, 160 Japan

Contributed papers

Microprocessor system for automatic Czochralski puller
L. Farkas, J. Paitz, G. Szemlaki, G. Zimmer, P. Szőke
Central Research Institute for Physics
H-1525 Budapest, P.O.Box 49. Hungary

Theoretical investigation of phantom controlled crystal growth
A. Mecseki*, B. Mező*, R. Voszka*, A. Galántai** and B. Szabó***
* Research Laboratory for Crystal Physics of the Hungarian Academy of Sciences
H-1112 Budapest, Budaörsi út 45. Hungary
** Mathematical Institute of the University for Agricultural Sciences
*** Information Centre of the Ministry of Health

Crystal growth by intensive processes at mass crystallization
Tibor Blickle, Zsuzsa Halász-Péterfi and Béla Lakatos
Research Institute for Technical Chemistry of the Hungarian Academy of Sciences,
Veszprém, Hungary

The growth of sapphire crystals by the variable shaping technique
V. A. Borodin, T. A. Steriopolo, V. A. Tatarchenko
Institute of Solid State Physics, USSR Academy of Sciences
142432 Chernogolovka, Moscow district, USSR

Production of sapphire tubes of high structural perfection for high pressure sodium lamps
V. A. Borodin, T. A. Steriopolo, V. A. Tatarchenko, L. J. Chernyshova, T. N. Yalovets
Institute of Solid State Physics, USSR Academy of Sciences
142432 Chernogolovka, Moscow district, USSR

The growth of sapphire single crystals; their properties and utilization
B. Perner, J. Kvapil, J. Kvapil, K. Blažek, B. Mánek
Monokrystaly Turnov ČSSR

Periodic and aperiodic thermal oscillation in a model of Czochralski growth
L. Gosztonyi, Z. Juhász and G. Szabó
Central Research Institute for Physics
H-1525 Budapest, P.O.B. 49. Hungary

The dissolution forms of gyromagnetic single crystal spheres
E. Beregi*, E. Hartmann**, J. Lábár***, E. Sterk* and F. Tanos*
* Research Institute for Telecommunication, Gábor Áron út 65. H-1026 Budapest, Hungary
** Research Laboratory for Crystal Physics, Budaörsi út 45, H-1112 Budapest, Hungary
*** Research Institute for Technical Physics, Fóti út 56, H-1047 Budapest, Hungary

The study of polished surfaces of Hg_2Cl_2 crystals — Application of plasma etching method
J. Žemlička and J. Kočka
Institute of Physics, Czechoslovak Academy of Sciences
Na Slovance 2, 180 40 Praha 8, ČSSR

Study of growth bands in potassium bichromate crystals by chemical etching
M. Szurgot, K. Sangwal and J. Karniewicz
Institute of Physics, Technical University of Łódz,
Wólczańska 219, 93—005 Łódz, Poland

Determination of the crystallization parameters of aluminium alloys in a technological casting process
É. Hidvégi, A. Lendvai
ALUTERV—FKI
Budapest, Hungary

Incorporation of OH^- ions in pure and doped $LiNbO_3$ crystals
L. Kovács, V. Szalay, S. Erdei, R. Capelletti*
Research Laboratory for Crystal Physics, Hungarian Academy of Sciences
H-1502 Budapest, P.O.B. 132, Hungary
* Institute of Physics, University of Parma, Parma, Italy

A simple method to determine the real composition of $LiNbO_3$
I. Földvári, K. Polgár, I. Cravero, R. Balasanyan*
Research Laboratory for Crystal Physics of the Hungarian Academy of Sciences
H-1502 Budapest, P.O.B. 132, Hungary
* Institute for Physical Research of the Armenian Academy of Sciences, Ashtarak

Dislocation study of TeO_2 crystals
J. Castaing*, E. Fries*, J. Jasznky** and Á. Péter**
* Laboratoire de Physique des Matériaux, CNRS, Meudon, France
** Research Laboratory for Crystal Physics of the Hungarian Academy of Sciences
H-1502 Budapest, P.O.B. 112, Hungary

Characterization of growth striations in $Gd_3Ga_5O_{12}$
M. Balaskó, J. Paitz, A. Süveges, É. Zsoldos
Central Research Institute for Physics
H-1525 Budapest, P.O.Box 49. Hungary

Crystal growth and spectroscopic testing of $ZnWO_4$
I. Földvári, I. Cravero, F. Schmidt, A. Watterich
Research Laboratory for Crystal Physics, Hungarian Academy of Sciences
H-1502 Budapest, P.O.B. 132, Hungary

Quality test of KDP monocrystals by optical methods
L. Szabó, E. Hartmann, Cs. Kuti, L. Vannay
Institute of Physics, Department of Experimental Physics, Technical University
1521 Budapest, Hungary

Hardness and fracture strength of garnet single crystals
M. Pardavi-Horváth, L. Gosztonyi, J. Paitz
Central Research Institute for Physics
H-1525 Budapest, Hungary

Thermoluminescent properties of gadolinium gallium garnet containing Ca^{2+} impurity
M. Pardavi-Horváth, M. Osvay*, I. Fellegvári**, J. Paitz, L. Gosztonyi**
Central Research Institute for Physics
H-1525 Budapest, Hungary
* Institute of Isotopes, Budapest, Hungary
** Hungarian Optical Works, Budapest, Hungary

Spectroscopic properties of Ca-doped $Gd_3Ga_5O_{12}$
I. Földvári, M. Pardavi-Horváth*, L. Gosztonyi*, J. Paitz*
Research Laboratory for Crystal Growth, H-1502 Budapest, Hungary
* Central Research Institute for Physics, H-1525 Budapest, Hungary

Investigation of microstructure of GaYIG single crystal faces, as-grown, polished and etched
E. K. Pál**, E. Beregi*, E. Sterk*, M. Farkas-Jahnke**, J. Lábár**
* Research Institute for Telecommunication, Budapest, Hungary
** Research Institute for Technical Physics of the Hungarian Academy of Sciences, Budapest, Hungary

Structure and properties of Ni – P thin films
G. Keresztury*, Á. Cziráki**
* Research Institute for Telecommunication Budapest, Hungary
** Institute for Solid State Physics, Eötvös University, Budapest, Hungary

Study of distribution of dopants in yttrium aluminium garnets crystals
K. Blažek*, J. Kvapil*, B. Perner*, Z. Hendrich*, B. Štverák** and J. Kopejto**
* UVVVM Turnov, ČSSR
** UVVVR Praha, ČSSR

On the growth of $\alpha - LiJO_3$ crystals
A. Yu. Klimova
Institute of Crystallography, USSR Academy of Sciences
Moscow, USSR

Growth of ordered crystals from disordered solid solution
L. N. Larikov
IMPh, Kiev, USSR

Effect of iron and silicon on the structure of Al and AlMn alloys
P. Bárczy, J. Sólyom,
Technical University for the Heavy Industry
Miskolc, Hungary

Manuscript received by Akadémiai Kiadó: 30 March 1984
Manuscript received by the Printers: 10 April 1984
Date of publication: 10 August 1985

PRINTED IN HUNGARY
Akadémiai Kiadó és Nyomda, Budapest

NOTES TO CONTRIBUTORS

I. PAPERS will be considered for publication in Acta Physica Hungarica only if they have not previously been published or submitted for publication elsewhere. They may be written in English, French, German or Russian.

Papers should be submitted to

Prof. I. Kovács, Editor
Department of Atomic Physics, Technical University
1521 Budapest, Budafoki út 8, Hungary

Papers may be either articles with abstracts or short communications. Both should be as concise as possible, articles in general not exceeding 25 typed pages, short communications 8 typed pages.

II. MANUSCRIPTS

1. Papers should be submitted in three copies.
2. The text of papers must be of high stylistic standard, requiring minor corrections only.
3. Manuscripts should be typed in double spacing on good quality paper, with generous margins.
4. The name of the author(s) and of the institutes where the work was carried out should appear on the first page of the manuscript.
5. Particular care should be taken with mathematical expressions. The following should be clearly distinguished, e.g. by underlining in different colours: special founts (italics, script, bold type, Greek, Gothic, etc.); capital and small letters; subscripts and superscripts, e.g. x^2, x_3; small *l* and I; zero and capital O; in expressions written by hand: *e* and *l*, *n* and *u*, *v* and v, etc.

A List of Symbols on a separate sheet should be attached to each paper.

6. References should be numbered serially and listed at the end of the paper in the following form: J. Ise and W. D. Fretter, Phys. Rev., *76,* 933, 1949.

For books, please give the initials and family name of the author(s), title, name of publisher, place and year of publication, e.g.: J. C. Slater, Quantum Theory of Atomic Structures, I. McGraw-Hill Book Company, Inc., New York, 1960.

References should be given in the text in the following forms: Heisenberg [5] or [5].

7. Captions to illustrations should be listed on a separate sheet, not inserted in the text.
8. In papers submitted to Acta Physica all measures should be expressed in SI units.

III. ILLUSTRATIONS AND TABLES

1. Each paper should be accompanied by three sets of illustrations, one of which must be ready for the blockmaker. The other sets attached to the copies of the manuscript may be rough drawings in pencil or photocopies.
2. Illustrations must not be inserted in the text.
3. All illustrations should be identified in blue pencil by the author's name, abbreviated title of the paper and figure number.
4. Tables should be typed on separate pages and have captions describing their content. Clear wording of column heads is advisable. Tables should be numbered in Roman numerals (I, II, III, etc.).

IV. RETURN OF MATERIAL

Owing to high postage costs, the Editorial Office cannot undertake to return *all* material not accepted for any reason for publication. Of papers to be revised (for not being in conformity with the above Notes or other reasons) only *one* copy will be returned. Material rejected for lack of space or on account of the Referees' opinion will not be returned to authors outside Europe.

Periodicals of the Hungarian Academy of Sciences are obtainable at the following addresses:

AUSTRALIA
C.B.D. LIBRARY AND SUBSCRIPTION SERVICE
Box 4886, G.P.O., *Sydney N.S.W. 2001*
COSMOS BOOKSHOP, 145 Ackland Street
St. Kilda (Melbourne), Victoria 3182

AUSTRIA
GLOBUS, Höchstädtplatz 3, *1206 Wien XX*

BELGIUM
OFFICE INTERNATIONAL DE LIBRAIRIE
30 Avenue Marnix, *1050 Bruxelles*
LIBRAIRIE DU MONDE ENTIER
162 rue du Midi, *1000 Bruxelles*

BULGARIA
HEMUS, Bulvar Ruszki 6, *Sofia*

CANADA
PANNONIA BOOKS, P.O. Box 1017
Postal Station "**B**", *Toronto, Ontario M5T 2T8*

CHINA
CNPICOR, Periodical Department, P.O. Box 50
Peking

CZECHOSLOVAKIA
MAD'ARSKÁ KULTURA, Národní třida 22
115 66 Praha
PNS DOVOZ TISKU, Vinohradská 46, *Praha 2*
PNS DOVOZ TLAČE, *Bratislava 2*

DENMARK
EJNAR MUNKSGAARD, Norregade 6
1165 Copenhagen K

FEDERAL REPUBLIC OF GERMANY
KUNST UND WISSEN ERICH BIEBER
Postfach 46, *7000 Stuttgart 1*

FINLAND
AKATEEMINEN KIRJAKAUPPA, P.O. Box 128 *SF- 00101 Helsinki 10*

FRANCE
DAWSON-FRANCE S. A., B. P. 40, *91121 Palaiseau*
EUROPÉRIODIQUES S. A., 31 Avenue de Versailles, *78170 La Celle St. Cloud*
OFFICE INTERNATIONAL DE DOCUMENTATION ET LIBRAIRIE, 48 rue Gay-Lussac
75240 Paris Cedex 05

GERMAN DEMOCRATIC REPUBLIC
HAUS DER UNGARISCHEN KULTUR
Karl Liebknecht-Straße 9, *DDR-102 Berlin*
DEUTSCHE POST ZEITUNGSVERTRIEBSAMT Straße der Pariser Kommüne 3–4, *DDR-104 Berlin*

GREAT BRITAIN
BLACKWELL'S PERIODICALS DIVISION
Hythe Bridge Street, *Oxford OXI 2ET*
BUMPUS, HALDANE AND MAXWELL LTD.
Cowper Works, *Olney, Bucks MK46 4BN*
COLLET'S HOLDINGS LTD., Denington Estate *Wellingborough, Northants NN8 2QT*
WM. DAWSON AND SONS LTD., Cannon House *Folkstone, Kent CT19 5EE*
H. K. LEWIS AND CO., 136 Gower Street
London WCIE 6BS

GREECE
KOSTARAKIS BROTHERS INTERNATIONAL BOOKSELLERS, 2 Hippokratous Street, *Athens-143*

HOLLAND
MEULENHOFF-BRUNA B. V., Beulingstraat 2, *Amsterdam*
MARTINUS NIJHOFF B.V.
Lange Voorhout 9–11, *Den Haag*
SWETS SUBSCRIPTION SERVICE
347b Heereweg, *Lisse*

INDIA
ALLIED PUBLISHING PRIVATE LTD., 13/14
Asaf Ali Road, *New Delhi 110001*
150 B-6 Mount Road, *Madras 600002*
INTERNATIONAL BOOK HOUSE PVT. LTD.
Madame Cama Road, *Bombay 400039*
THE STATE TRADING CORPORATION OF INDIA LTD., Books Import Division, Chandralok 36 Janpath, *New Delhi 110001*

ITALY
INTERSCIENTIA, Via Mazzé 28, *10149 Torino*
LIBRERIA COMMISSIONARIA SANSONI, Via Lamarmora 45, *50121 Firenze*
SANTO VANASIA, Via M. Macchi 58
20124 Milano
D. E. A., Via Lima 28, *00198 Roma*

JAPAN
KINOKUNIYA BOOK-STORE CO. LTD.
17-7 Shinjuku 3 chome, Shinjuku-ku, *Tokyo 160-91*
MARUZEN COMPANY LTD., Book Department, P.O. Box 5050 Tokyo International, *Tokyo 100-31*
NAUKA LTD. IMPORT DEPARTMENT
2-30-19 Minami Ikebukuro, Toshima-ku, *Tokyo 171*

KOREA
CHULPANMUL, *Phenjan*

NORWAY
TANUM-TIDSKRIFT-SENTRALEN A.S., Karl Johansgatan 41–43, *1000 Oslo*

POLAND
WĘGIERSKI INSTYTUT KULTURY, Marszalkowska 80, *00-517 Warszawa*
CKP-I W, ul. Towarowa 28, *00-958 Warszawa*

ROUMANIA
D. E. P., *Bucuresti*
ILEXIM, Calea Grivitei 64–66, *Bucuresti*

SOVIET UNION
SOJUZPECHAT — IMPORT, *Moscow*
and the post offices in each town
MEZHDUNARODNAYA KNIGA, *Moscow G-200*

SPAIN
DIAZ DE SANTOS, Lagasca 95, *Madrid 6*

SWEDEN
GUMPERTS UNIVERSITETSBOKHANDEL AB
Box 346, *401 25 Göteborg 1*

SWITZERLAND
KARGER LIBRI AG, Petersgraben 31, *4011 Basel*

USA
EBSCO SUBSCRIPTION SERVICES
P.O. Box 1943, *Birmingham, Alabama 35201*
F. W. FAXON COMPANY, INC.
15 Southwest Park, *Westwood Mass. 02090*
READ-MORE PUBLICATIONS, INC.
140 Cedar Street, *New York, N. Y. 10006*

YUGOSLAVIA
JUGOSLOVENSKA KNJIGA, Terazije 27, *Beograd*
FORUM, Vojvode Mišića 1, *21000 Novi Sad*

Index: 26.022