RAMAN SPECTROSCOPY OF DOUBLE ACTIVATED $YNbO_4:Eu^{3+},Tb^{3+}$ AND $YTaO_4:Eu^{3+},Tb^{3+}$

M. Nazarov

Department of Materials Science and Engineering, Gwangju Institute of Science and Technology, 1, Oryong-dong, Buk-gu, 500-712, Gwangju, Republic of Korea
Institute of Electronic Engineering and Industrial Technologies, Academy of Sciences of Moldova, 3/3, Academiei str., MD-2028, Chisinau, Republic of Moldova
(Received 12 August 2009)

Abstract

Three new peaks at 503, 524, and 589 cm^{-1} in Raman spectroscopy of double activated $YNbO_4:Eu^{3+},Tb^{3+}$ phosphor and four new peaks at 446, 484, 542, and 592 cm^{-1} in $YTaO_4:Eu^{3+},Tb^{3+}$ were observed and identified. The nature of theses peaks is discussed.

1. Introduction

The blue light emission from yttrium niobate $YNbO_4$ phosphors (at about 400 nm), or UV emission (~ 330 nm) from yttrium tantalate $YTaO_4$ is associated with NbO_4^{3-} or TaO_4^{3-} groups from the host crystalline lattice [1]. Such luminescent emission could be shifted toward longer wavelengths when rare earth ions such Eu^{3+} or Tb^{3+} replace partially the yttrium ions in the host crystalline lattice. In this case, Eu^{3+} and Tb^{3+} emission centers are created and can generate the corresponding red and green luminescence, respectively. The general luminescence of Eu^{3+} or Tb^{3+} in niobates has been reported previously [1-5]. Some recent publications with separated activation by Eu^{3+} or Tb^{3+} yttrium niobate systems [6-8] and yttrium tantalate systems [9-12] appeared in the last years. However, to the best of our knowledge, no work has been reported dealing with the doubly activators (Eu^{3+} and Tb^{3+}) for $YNbO_4$ and $YTaO_4$ phosphors. To better understand the luminescence mechanism and crystallographic structure of these phosphors, the Raman spectroscopy seems very useful and actual. The Raman spectroscopy for investigation non-activated $YNbO_4$ and $YTaO_4$ was first applied by Blasse [13]. Till now the vibrational spectra of $YNbO_4$ and $YTaO_4$ double activated by Eu^{3+} and Tb^{3+} have not been reported. These compounds have a slight distortion of the structure in comparison with non-activated materials, and it seems interesting to investigate the influence of activators on the Raman spectra.

2. Experimental details

2.1. Sample preparation

Yttrium tantalate and niobate phosphors activated by the rare earth elements, such as Eu^{3+} and Tb^{3+}, were prepared by solid state reaction method from homogeneous mixture consisting of Y_2O_3 (99.9%), Ta_2O_5 (Optipur) and Nb_2O_5 (99%). The oxide precursors for the host lattice, Eu_2O_3 and/or Tb_4O_7, are used in the activator system and Na_2SO_4 (99%) as flux [14]. The mixtures were homogenized with a ball mill, in acetone medium, and dried at 70°C. The phosphor samples were baked at 1200°C for 4 h and slowly cooled to room temperature. Finally, the samples were water washed, dried, and then sieved.

Two types of host lattices (M and M') with double activation by rare earth elements Eu^{3+} and Tb^{3+} were prepared and investigated. The chemical compositions of prepared phosphor samples are listed in Table 1.

Table 1. Composition of yttrium tantalite/niobate phosphors with double activation by Eu^{3+} and Tb^{3+}.

N	Phosphor type	Crystallographic structure	Host lattice (mol%)		Re-activator (mol%)	
			$YTaO_4$	$YNbO_4$	Eu^{3+}	Tb^{3+}
1.	$YTaO_4$:Eu,Tb	M'-modified monoclinic fergusonite	100	-	0	0
					2.5	2.5
					5	0
					0	5
2.	$YNbO_4$:Eu,Tb	M- Monoclinic fergusonite crystal	-	100	0	0
					2.5	2.5
					5	0
					0	5

In M'-$YTaO_4$ structure, tantalum atoms are in a distorted octahedral coordination with six Ta-O bonds whereas in M-$YTaO_4$ tantalum atoms are in tetrahedral coordination [15, 16]. TaO_6 in a fergusonite structure is often described as a twisted triangular prism [17]. These distorted TaO_6 units share edges with one another to form strings. Yttrium atoms are surrounded by 8-coordinated oxygen atoms forming a distorted cube. The total number of atoms inside of the M'-$YTaO_4$ structure is: 2 yttrium atoms, 2 tantalum, and 8 oxygen atoms [18].

The niobium atoms of the fergusonite (distorted sheelite) M-$YNbO_4$ structure can present the same octahedral coordination with six Nb-O bonds, such as M'-$YTaO_4$ with the coordination number of niobium atoms (4 + 2) or tetrahedral coordination with 4 oxygen atoms. The tetrahedrally coordinated NbO_4 structure is not a typical structure for niobium oxide because the Nb^{5+} atom is too large to fit into an oxygen-anion tetrahedron. Only a few rare-earth $ANbO_4$ (A = Y, Yb, Sm, and La) compounds have been found to possess tetrahedral coordination [13, 19, 20]. The transition from the octahedral fergusonite to the tetrahedral scheelite depends on temperature. An increase in temperature results in bond breaking between the niobium atom and the two further oxygen atoms, and the Nb atom coordination becomes tetrahedral.

In the present study, $YNbO_4$ was used as the niobium oxide reference compound containing a tetrahedrally coordinated NbO_4 structure.

3. Results and discussion

X-ray diffraction spectrum of $YNbO_4:Eu^{3+},Tb^{3+}$ (Fig. 1a) shows the evidence of the monoclinic M-$YNbO_4$ fergusonite crystal structure and $YTaO_4:Eu^{3+},Tb^{3+}$(Fig. 1b) shows M'-$YTaO_4$ crystal structure. In Fig. 1, the black points are the measured data and the solid lines are the calculated fitting of Rietveld analysis, which was developed by Rietveld for structure profile refinement of X-ray powder diffraction data. The measured peak positions are in a good agreement with Rietveld fitting and data given in JCPDS.

3.1. $YNbO_4:Eu^{3+},Tb^{3+}$

Raman spectroscopy is very sensitive to the structure and bond order of metal oxides, especially in the region of metal-oxygen stretching modes, because many of the Raman frequen-

cies depend on the bond order in the structure. A higher metal-oxygen bond order, corresponding to a shorter bond distance, shifts the Raman bond to higher wavenumbers. In this study, we investigate both non-activated host lattice as well as double activated by Eu^{3+} and Tb^{3+}.

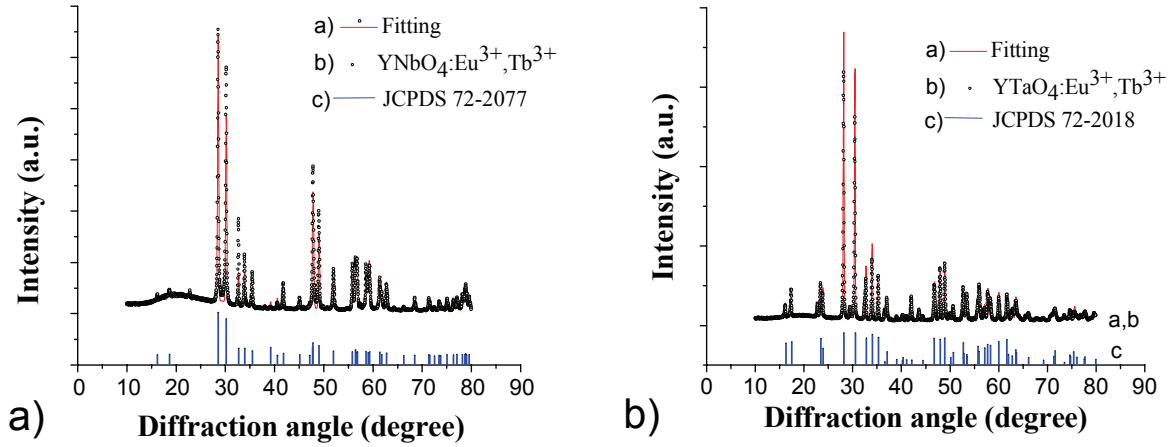


Fig. 1. XRD patterns of $YNbO_4:Eu^{3+},Tb^{3+}$(a), and $YTaO_4:Eu^{3+},Tb^{3+}$(b) and its fitting. Expected peak positions from JCPDS card are indicated.

Raman spectra were measured in the spectral range 100-1000 cm^{-1} with the inVia Reflex Raman Microscope, using a He-Ne laser (632.8 nm). This spectral range usually contains bands due to metal-oxide bonds and/or lattice vibrations. In the Raman spectra, more associated with lattice vibrations and symmetrical bonds, small amount of activator can be observed in additional peaks. We observed ν_1, ν_2, ν_3, and ν_4 vibration modes from the host lattice $YNbO_4$ (spectrum 1 in Fig. 2), which correspond to a regular NbO_4 tetrahedron with no interactions and distortions at 815 (ν_1), 650 (ν_3), 420 (ν_4), and 340 cm^{-1} (ν_2) predicted by Blasse [13]. Raman bands appearing at 815 and 340 cm^{-1} are due to Nb-O symmetric modes of the NbO_4 tetrahedral structure, and Raman bands appearing at 650 and 420 cm^{-1} are due to Nb-O antisymmetric modes of the NbO_4 tetrahedral structure. The phonon energies below 300 cm^{-1} are assigned to external vibrations. One can see four new peaks around 446 and 503 (small peaks, which we will not take into consideration in future), 524 and 589 cm^{-1} in the spectrum 2 from the double activated $YNbO_4$ by 2.5 mol % Eu^{3+} and 2.5 mol % Tb^{3+} (Fig. 2(2)). The detailed analysis shows that no additional peak is observed when yttrium niobate is doped only with terbium and the Raman spectra of $YNbO_4:Tb^{3+}$ are identical with spectra $YNbO_4$ (Fig. 2(1)). We assumed that these four mew peaks are related to europium activation. Raman spectrum 3 (Fig. 2) from $YNbO_4:(Eu^{3+}$, 5% mol) confirms our assumption. Moreover, the intensity of these peaks is proportional to europium concentration, while the intensity of vibration modes from the host lattice $YNbO_4$ at 815 cm^{-1} (ν_1) does not change. According to Blasse [13] (ν_1) is related to symmetrical stretching vibrations of NbO_4 near 815 cm^{-1}. The nature of this phenomenon is discussed.

A new series of $YNbO_4:Eu^{3+}$ samples with Eu concentration up to 35 mol % were synthesized with different fluxes (H_3BO_3 and LiCl) and Raman spectra were measured. Fragments of these spectra (with LiCl flux) are shown in Fig. 3. The linear increase of peak intensity at 589 cm^{-1} with Eu concentration is evident from Fig. 3.

To estimate the Eu concentration in yttrium niobate host lattice, the calculation of relative intensity of two peaks at 589 cm^{-1} and 815 cm^{-1} were performed. The ratio of Raman intensities I_{589}/I_{815} versus europium concentration is presented in Fig. 4.

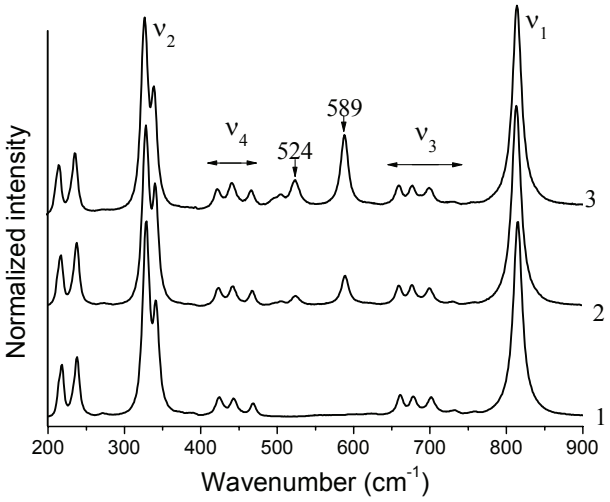


Fig. 2. Raman spectra of the $YNbO_4$ (1), $YNbO_4:Eu^{3+},Tb^{3+}$ (2), and $YNbO_4:Eu^{3+}$ (3).

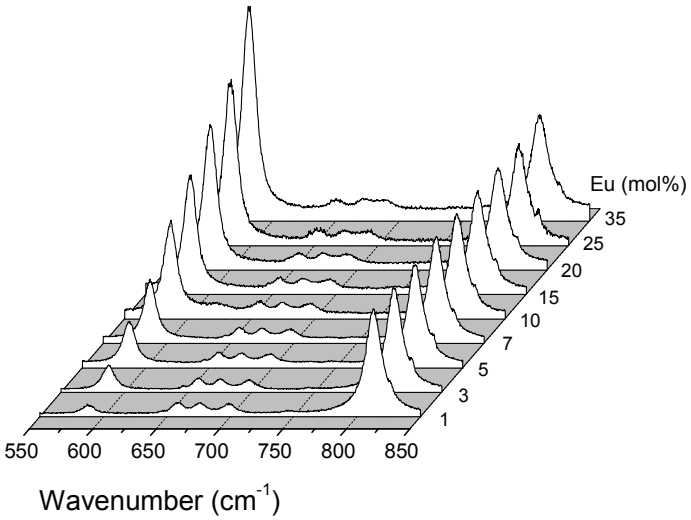


Fig. 3. Fragments of Raman spectra of the $YNbO_4:Eu^{3+}$ with different Eu concentration.

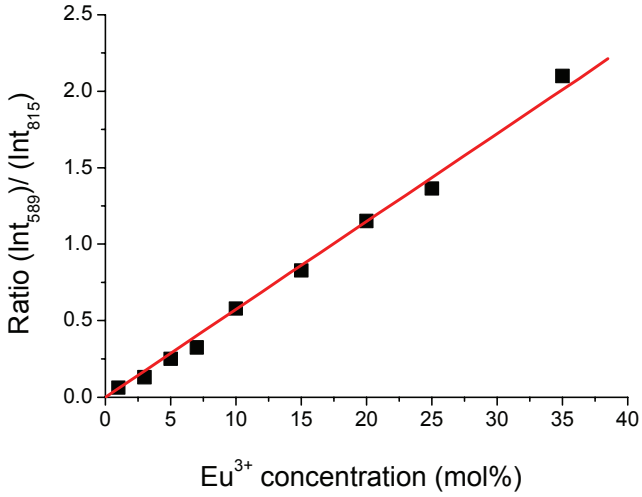


Fig. 4. Ratio of Raman intensities I_{589}/I_{815} vs. europium concentration.

Experimental data can be reproduced by linear equation Y = kX, where Y is the ratio of Raman intensities I_{589}/I_{815} and X is the europium concentration. From experimental measurement of Raman spectra and comparison of the intensities of Eu (589 cm^{-1}) and the host lattice peak (815 cm^{-1}), it is easy to estimate the activator concentration in any synthesized $YNbO_4:Eu^{3+}$ sample.

To be sure in quantitative data, a new series of $YNbO_4:Eu^{3+}$ sample with Li_2SO_4 flux was prepared [21] and investigated with Raman spectroscopy. The results are shown in Fig. 5.

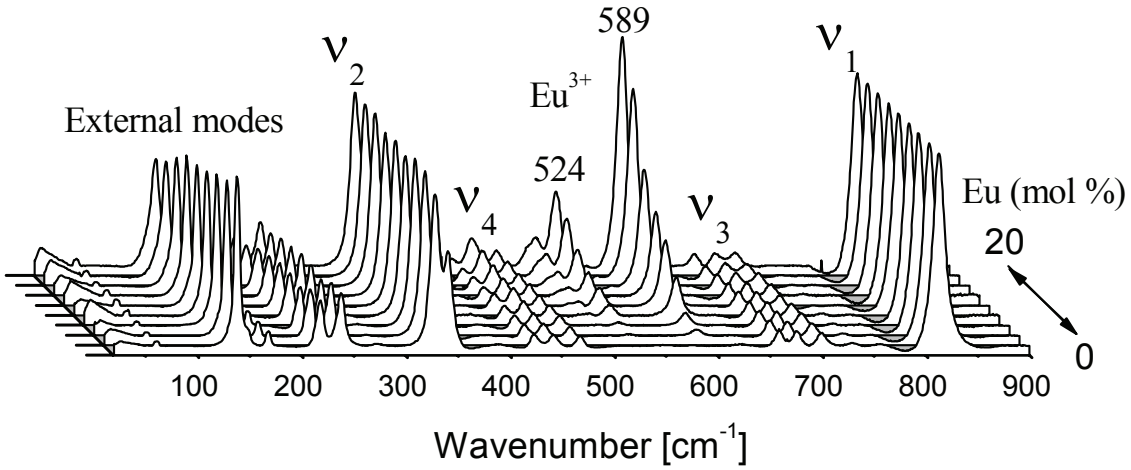


Fig. 5. Series of $YNbO_4:Eu^{3+}$ Raman spectra with different Eu^{3+} concentration.

All observed vibration modes ν_1, ν_2, ν_3, and ν_4 associate with the host lattice $YNbO_4$, [13] and these modes, as well as external modes below 300 cm^{-1}, do not depend on the Eu concentration. Only two main peaks at 524 cm^{-1} and 589 cm^{-1} show strong dependence on the Eu concentration. Both peaks have a linear dependence that is clear seen from Fig. 6, but the slope of these lines is different due to the different influence of Eu (the slope is equal to zero for all other peaks, when there are no Eu influence).

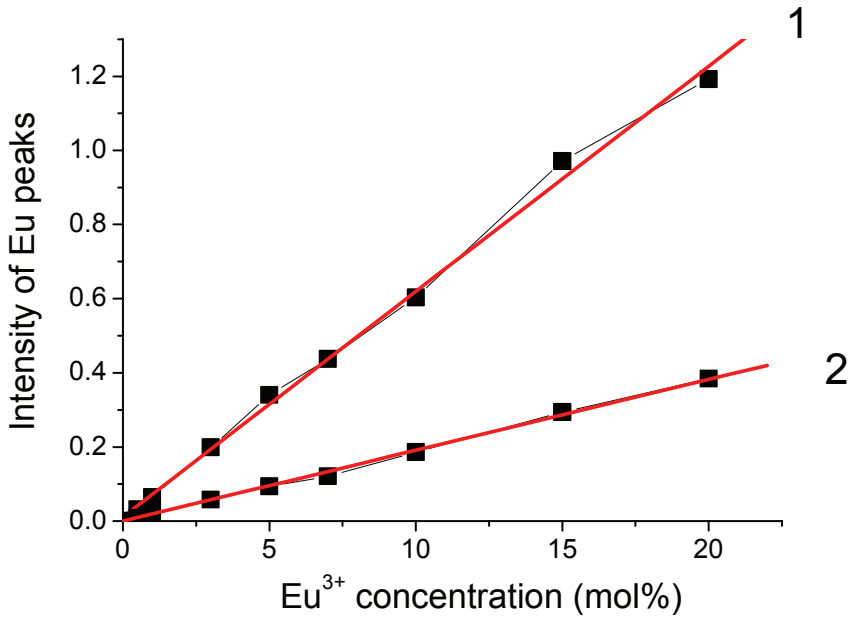


Fig. 6. The ratio of Raman intensities I_{589}/I_{815} (1) and I_{524}/I_{815} vs. europium concentration.

If we compare the same ratio I_{589}/I_{815} from different series prepared with various fluxes (LiCl, Fig. 4 and Li_2SO_4, Fig. 6), we can see that the linear tendency is very similar (Fig. 7) and the intensities are a little higher at the samples prepared with Li_2SO_4 flux.

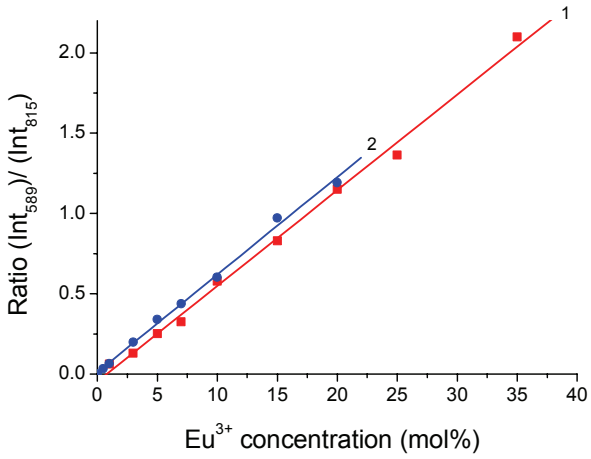


Fig. 7. The ratio of Raman intensities I_{589}/I_{815} vs. europium concentration for two different $YNbO_4:Eu^{3+}$ sample series: (1) series prepared with LiCl as a flux, and (2) with Li_2SO_4.

This phenomenon can be explained by different activator incorporation degree (AID) [22] in synthesis process. The term AID is used to define the ability of europium to efficiently substitute yttrium from the host lattice. Activator substitution in general is a function of concentration, annealing temperature, flux, and other synthesis parameters. We measured the real Eu concentration in powders synthesized with different fluxes under the other synthesis conditions being the same. Quantitative non-destructive elemental analyses were performed with EPMA machine (SX-100, the electron acceleration 15 kV, beam current 10 nA, and the diameter of electronic beam was 1 μm.). The data are summarized in Table 2.

Table 2. Synthesis and real Eu mol% concentration with different fluxes.

Eu mol% concentration (synthesis)	Flux LiCl		Flux Na_2SO_4		Flux Li_2SO_4	
	Real Eu%	AID (%)	Real Eu%	AID (%)	Real Eu%	AID (%)
0.5					0.45	90
1					0.9	90
2.5			2.12	85		
3	2.44	81			2.64	88
5	3.75	75	4.25	85	4.6	92
7	5.58	79			6.3	90
10					8.9	89
15	12.51	83			13.5	90
20	16.11	81			18.2	91
35	27.95	80				
Average AID (%) →		80		85		90

We can conclude from Table 2 that the most important parameter is not the concentration of activator proposed during the synthesis, but the real concentration in prepared samples. The best results can be achieved with Li_2SO_4 as a flux when the AID is about 90%. If we take into account AID, the Raman spectra results (Fig. 7) versus real Eu concentration become identical.

3.2. $YTaO_4:Eu^{3+},Tb^{3+}$

The Raman spectrum of non-activated M'-$YTaO_4$ in comparison with M-$YNbO_4$ is shown in Fig. 8.

We can see from this figure that the Raman spectra differ significantly. M-$YNbO_4$ and M'-$YTaO_4$ compounds has a unique Raman spectrum that is related to the symmetry and bond order of its structure. Only symmetric modes ν_1 in Raman spectra of M-$YNbO_4$ coincide with M'-$YTaO_4$ and peaking both at 818 cm^{-1}. The main differences are due to different Nb-O antisymmetric modes of the NbO_4 tetrahedral structure (ν_3 and ν_4 vibration modes) and Ta-O antisymmetric modes in distorted octahedral TaO_4 structure. Moreover, Ta-O symmetric mode in Raman spectra ν_2 is shifted to the short wavenumber area at about 285 cm^{-1}, while Nb-O symmetric mode ν_2 is at 340 cm^{-1}. The similar tendency was observed by Brixner between M' and M structure in $YTaO_4$ [15].

It is interesting to check the influence of activators on vibrational spectra in tantalate system. The results of activation M'-$YTaO_4$ by Tb^{3+} and Eu^{3+} are presented in Fig. 9.

Experimental data show that Tb^{3+} activation does not change the vibrational spectra, while activation by Eu^{3+} results in appearing four new peaks at 446, 484, 542, and 592 cm^{-1} (Fig. 9). The intensity of these peaks strongly depends on Eu^{3+} concentration (Fig. 10).

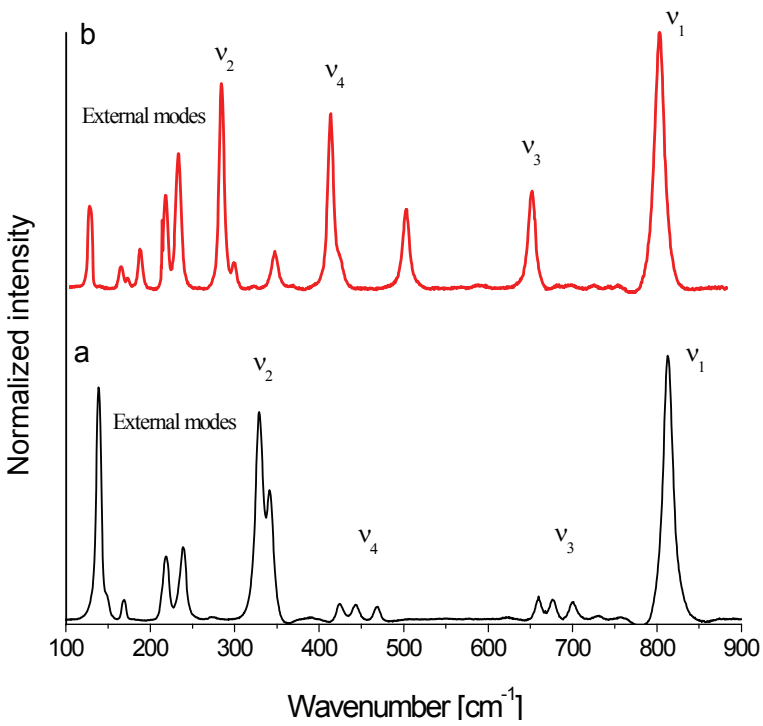


Fig. 8. Raman spectra of the M-$YNbO_4$ (a), and M'-$YTaO_4$ (b).

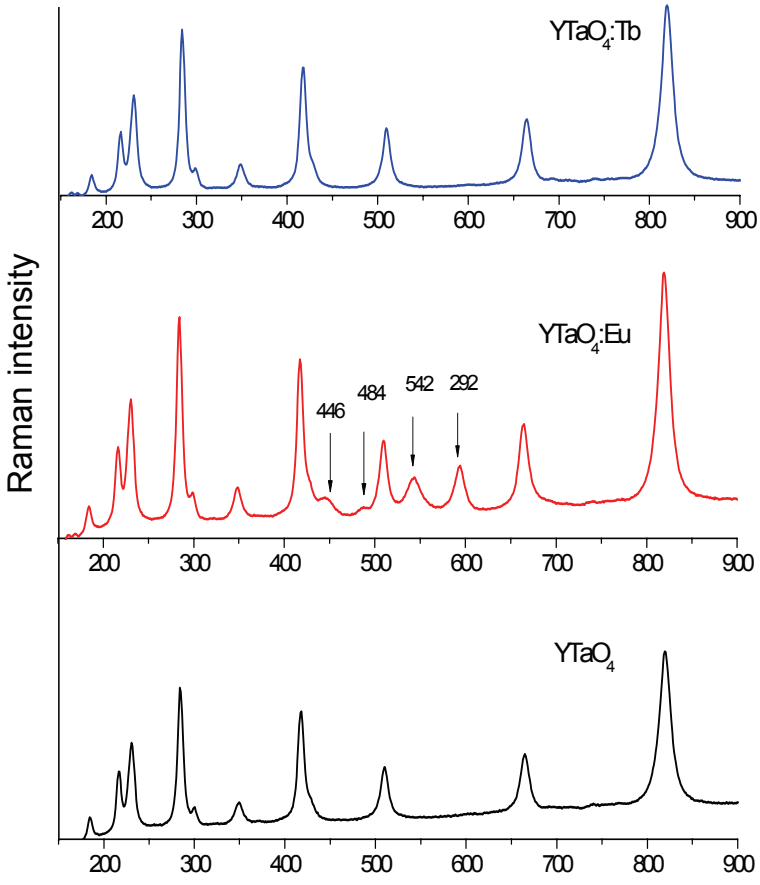


Fig. 9. Raman spectra of the M'-$YTaO_4$, M'-$YTaO_4:Eu^{3+}$ (5 mol %) and M'-$YTaO_4:Tb^{3+}$ (5 mol %).

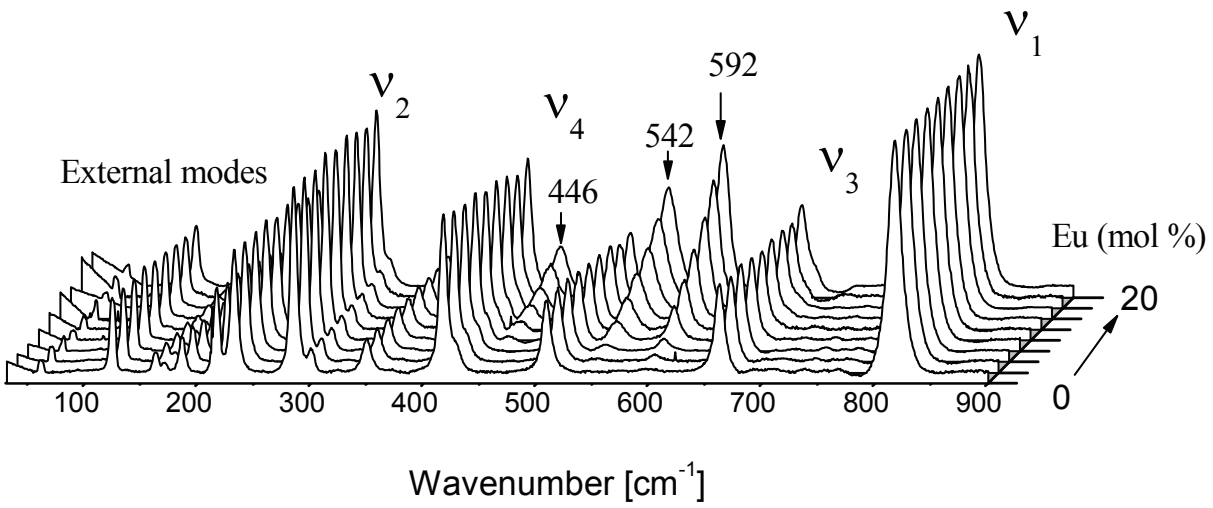


Fig. 10. Series of M'-$YTaO_4$:Eu^{3+} Raman spectra with different Eu^{3+} concentration.

Both peaks at 484 and 446 cm^{-1} (small peak at 484 cm^{-1} is not seen in this picture) have a weak dependence on Eu and can be described by linear function that is clear seen from Fig. 11. Two other peaks at 542 and 592 cm^{-1} show linear dependence only at low Eu concentrations (less than 10 mol %) and can be described by exponential functions.

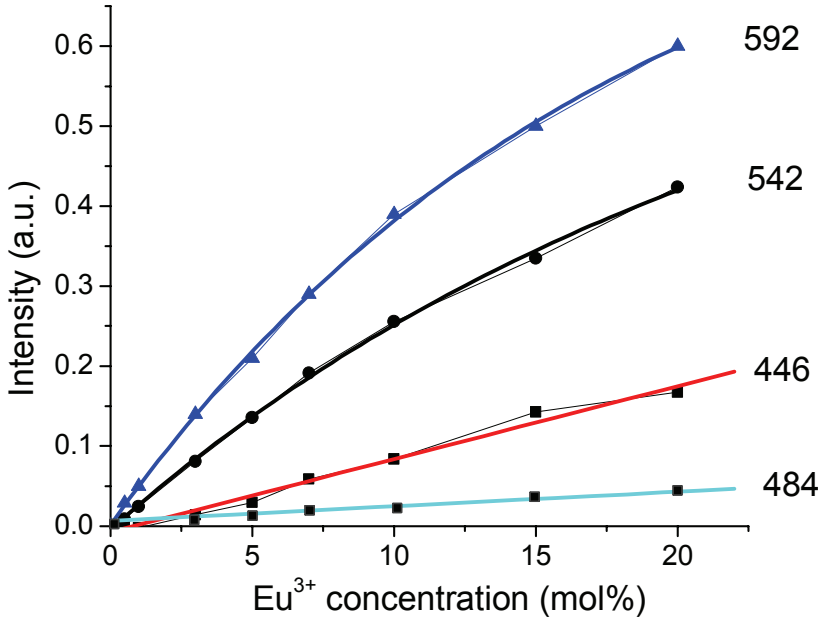


Fig. 11. Intensities of Raman peaks at 446, 484, 542, and 592 cm^{-1} (normalized on intensity of ν_1 at 818 cm^{-1}) vs. europium concentration.

3.3. General discussion

To better understand the nature of additional peaks in niobate and tantalate systems, we used three different lasers with excitation wavelengths 532, 632, and 785 nm (Fig. 12).

At room temperature, He-Ne laser (632 nm) can excite Eu^{3+} ions to $^5D_0$ energy level. The luminescence from $^5D_0$ energy level to $^7F_4$ - $^7F_3$ band gets to 650-660 nm wavelength re-

gion and overlaps with spontaneous Raman spectrum of $YNbO_4$ host crystal. We can suppose that the registered spectrum at 632 nm excitation is a mixture of Raman spectrum and luminescence lines of Eu^{3+}. It is most probable, and we have to take into account this explanation. When we excite the host lattice by 532 nm, these peaks are shifted from 500-600 cm^{-1} to 925-1075 cm^{-1} (this area is not shown in Fig. 12). At the same time, new additional luminescence transitions $^5D_1$-$^7F_1$ became possible; they overlap the vibrational peaks in Raman spectra in the range 700-800 cm^{-1}. Only weak energy excitation (785 nm) cannot excite Eu^{3+} ions to $^5D_0$ energy level and cannot be a cause of luminescence.

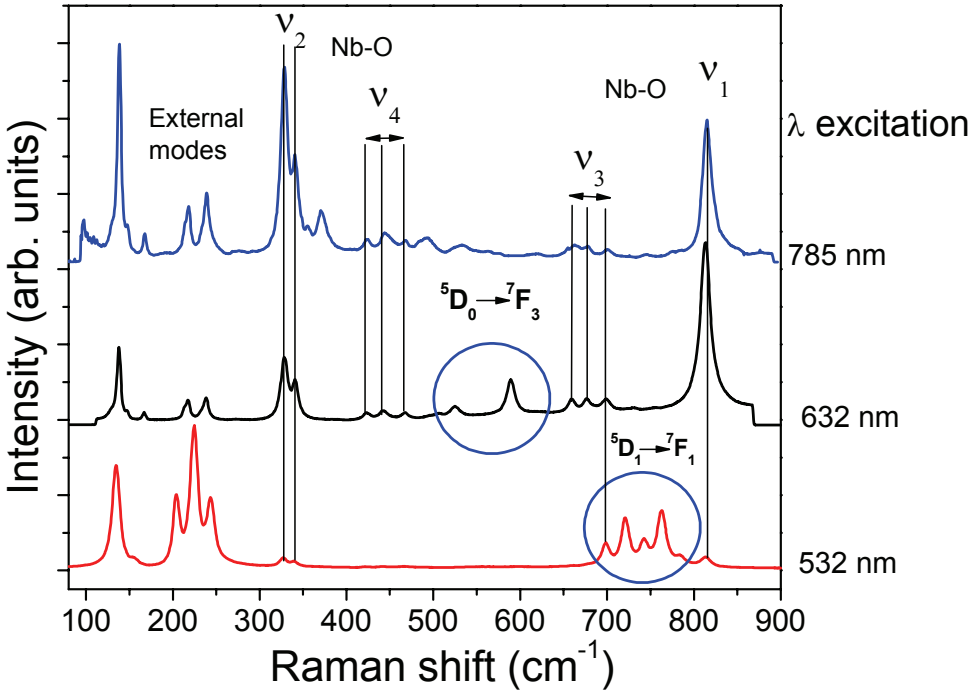


Fig. 12. Raman spectra of the M-$YNbO_4$:Eu^{3+} under different laser excitation: 785, 632, and 532 nm.

We compared our data with Blasse [13] who studied non-activated M-$YNbO_4$ and M-$YTaO_4$ compounds by Raman spectroscopy and predicted the vibrational modes of a regular NbO_4 tetrahedron with no interactions and distortions to occur at 816 (ν_1), 650 (ν_3), 420 (ν_4), and 340 cm^{-1} (ν_2). The real experimental data are a little different from the theory and shown in Table 3.

For the same host lattices (M-$YNbO_4$), Blasse data [13] and our experimental results are very similar, but for different structures (M' and M-$YTaO_4$) the results differ. It is no wonder, because the tantalum atoms in these structures are in a distorted octahedral and tetrahedral coordination, respectively. Eu activated phosphors show three (niobate) and four (tantalate) additional luminescence peaks. Intensities of these peaks depend on Eu concentration.

The Raman band assignments are based on the corresponding Nb-O bond order and known structure. After Blasse [13] and Rooksby et al. [19] the $ANbO_4$ (A = Y, Yb, Sm, and La) compounds were further studied by Kinzhibalo et al. [25] and Trunov et al. [26], who determined that these compounds possess a fergusonite structure at room temperature. In the fergusonite structure, the coordination number of the Nb atom contains four oxygen atoms with distances of 1.83-1.93 Å and two further oxygen atoms with distances of 2.40-2.52 Å as a result of the connection between two adjacent tetrahedra.

Table 3. Raman spectra of the $YTaO_4$ and $YNbO_4$ groups in fergusonite. M is monoclinic fergusonite structure, M' is modified monoclinic structure, (R) means Raman-active; s, strong; m, medium; w, weak; sh, shoulder.

Free NbO_4 ion	NbO_4 ion in $YNbO_4$	Experimental (values in cm^{-1})						Assignment
		$YNbO_4$		$YNbO_4$:Eu	$YTaO_4$		$YTaO_4$:Eu	
		M	M	M	M	M'	M'	
(T_d)	(C_2)	Ref [13]	Our	Our	Ref [13]	Our	Our	
ν_1:A_1(R) 860	A	832(s)	815(s)	815(s)	825(s)	818(s)	818(s)	Symmetric Nb-O stretching vibrations [13, 23]
		-	-	783(sh)	-	-	-	
ν_3:T_2(R) 650	A+2B	715(w)	758(w)	763(w)	720(w)	-	-	Antisymmetric Nb-O stretching vibrations [13]
		695(w)	730(w)	743(w)	705(m)	-	-	
		675(w)	700(w)	720(w)	670(m)	662(m)	662(m)	
		650(w)	701(w)	700(w)	655(w)	-	-	
		-	678(w)	676(w)	-	-	-	
		560(w)	660(w)	660(w)	-			
				589(*)			592(*)	(*) Eu^{3+} peaks (This work)
				524(*)			542(*)	
				503(*)			484(*)	
				446(*)			446(*)	
ν_4:T_2(R) 420	A+2B	480(w)	469(w)	468(w)	480(w)	508(m)	508(m)	Nb-O antisymmetric modes of the NbO_4 tetrahedral structure [13].
		455(w)	445(w)	443(w)	450(w)	-	-	
		435(w)	427(w)	424(w)	-	417(s)	417(s)	
		-	403(w)	403(w)	-	-	-	
		385(w)	370(w)	372(w)	375(w)	-	-	
		-	356(w)	355(w)	-	349(w)	349(w)	
ν_2:E(R)	2A	350(s)	340(s)	340(s)	345(s)	300(w)	300(w)	Symmetric stretching vibrations of the groups, ν_s (NbO_4) [13, 23]
		340(s)	328(s)	328(s)	320(s)	285(s)	285(s)	
External modes		245(m)	238(m)	238(m)		233(m)	233(m)	Lattice vibrations [24]
		225(m)	218(m)	218(m)		217(m)	217(m)	
		170(w)	168(w)	168(w)	215(s)	185(w)	185(w)	
		-	147(w)	147(w)	120(m)	162(w)	162(w)	
		-	138(s)	138(s)		124(m)	124(m)	

The multiple Raman bands appearing in the 400-800 cm^{-1} region for the $YNbO_4$ compound reported by Blasse [13] are probably due to the simultaneous presence of NbO_4 tetrahedra and NbO_6 octahedra in the structure.

4. Conclusions

Four new peaks at 446, 503, 524, and 589 cm^{-1} in Raman spectroscopy of double activated $YNbO_4:Eu^{3+},Tb^{3+}$ and four new peaks at 446, 484, 542, and 592 cm^{-1} in $YTaO_4:Eu^{3+},Tb^{3+}$ were observed and identified as luminescence peaks. The Raman method can be recommended as an accurate and quantitative express analysis of europium activated yttrium niobate and tantalate phosphors. Raman and luminescence experiments show that the best results can be achieved in synthesis of high efficient phosphors using Li_2SO_4 as a flux.

5. Acknowledgements

This work was supported by the Korea Science and Engineering Foundation (KOSEF) through NRL (Program No. M10400000045-04J0000-04510) and NCRC grant (No. R15-2008-006-00000-0) funded by Ministry of Education, Science, and Technology in Korea. This work was also supported by Photonics 2020 research project through a grant provided by the Gwangju Institute of Science and Technology in 2008.

References

[1] G. Blasse and A. Bril, J. Electrochem Soc., 115, 1067, (1968).
[2] G. Blasse and B.C. Grabmeyer, Luminescent Materials, Springer Verlag, Berlin, Heidelberg, 232 p., 1994.
[3] L.H. Brixner, Materials Chemistry and Physics, 16, 253, (1987).
[4] G. Blasse and A. Bril, Journal of Luminescence, 3, 109, (1970).
[5] A. Bril, W.L.Wanmaker, and R.E. Schuil, in Intern. Symp. on Luminescence, Munich, 310, 1965.
[6] X. Xiao and B. Yan, Journal of Non-Crystalline Solids, 351, 3634, (2005).
[7] X. Xiao and B. Yan, Materials Letters, 61, 1649, (2007).
[8] X. Xiao and B. Yan, Applied Physics A: Materials Science and Processing, 88, 333, (2007).
[9] T. Takayama, T. Katsumata, S. Komuro, and T. Morikawa, Journal of Crystal Growth, 275, e2013, (2005).
[10] B. Li, Z. Gu, J. Lin, and M.Z. Su, Materials Research Bulletin, 35, 1921, (2000).
[11] M.H. Hwang and Y.J. Kim, Ceramics International, 34, 1117, (2008).
[12] M. Nazarov, B. Tsukerblat, C.C. Byeon, I. Arellano, E.J. Popovici, and D.Y. Noh, Applied Optics, 48, 17, (2009).
[13] G. Blasse, Journal of Solid State Chemistry, 7, 169, (1973).
[14] M. Nazarov, E.J. Popovici, I. Arellano, and D.Y. Noh, Mold. J. Phys. Sci., 7, 4, 433, (2008).
[15] L.H. Brixner and H.Y. Chen, J. Electrochem. Soc., 130, 2435, (1983).
[16] G. Blasse and L.H. Brixner, J. Solid State Chem., 82, 115, (1989).
[17] V.K. Trunov, L.N. Kinzhibalo, V.A. Efremov, and V.G. Krongauz, DAN USSR, 260, 103, (1981).
[18] S.L. Issler and C.C. Torardi, Journal of Alloys and Compounds, 229, 54, (1995).
[19] H.P. Rooksby and E.A.D. White, Acta. Crystallogr., 16, 888, (1963).
[20] S. Yoshida, Y. Nishimura, T. Tanaka, H. Kanai, and T. Funabiki, Catal. Today, 8, 67, (1990).
[21] E.J. Popovici, M. Nazarov, L. Muresan, Do Young Noh, E. Bica, M. Morar, I. Arellano, and E. Indrea, in Proc. of the 15th Inter. Conf. on Luminescence and Optical Spectroscopy of Condensed Matter (ICL'08), Lyon, 219, (2008).
[22] M.V. Nazarov, B.S. Tsukerblat, E.J. Popovici, and D.Y. Jeon, Solid State Communications, 133, 203, (2005).
[23] K.I. Petrov, N.V. Gundobin, V.V. Kravchenko, and S.S. Plotkin, J. Inorg. Chem., [in Russian], 18, 928, (1973).
[24] J.M. Jehng and I.E. Wachs, Chem. Mater., 3, 100, (1991).
[25] L.A. Kinzhibalo, V.K. Trunov, A.A. Evdokimov, and V.G. Krongauz, Kristallografia, 27, 43, (1982).
[26] V.K. Trunov, V.A. Efremov, Y.A. Velikopodnyi, and I.M. Averina, Soviet Physics: Crystallography, 26, 35, (1981).

COLLECTIVE ELEMENTARY EXCITATIONS OF TWO-DIMENSIONAL MAGNETOEXCITONS IN THE BOSE-EINSTEIN CONDENSATION STATE WITH WAVE VECTOR $k = 0$

S.A. Moskalenko[1], M.A. Liberman[2], E.V. Dumanov[1,*], S.S. Russu[1], and I. Sanduleac[1]

[1]*Institute of Applied Physics, Academy of Sciences of Moldova, 5, Academiei str., MD-2028, Chisinau, Republic of Moldova*
[2]*Department of Physics, Uppsala University, Box 530, SE-751 21, Uppsala, Sweden*
[*]*E-mail: dum@phys.asm.md*
(Received 18 August 2009)

Abstract

The collective elementary excitations of the two-dimensional magnetoexcitons in a state of Bose-Einstein condensation (BEC) with wave vector $\vec{k} = 0$ were investigated in terms of the Bogoliubov theory of quasiaverages. The starting Hamiltonian of the electrons and holes lying on the lowest Landau levels (LLLs) contains the supplementary interactions due to the virtual quantum transitions of the particles to the excited Landau levels (ELLs) and return back. As a result, the interaction between the magnetoexcitons with $\vec{k} = 0$ does not vanish and their BEC becomes stable as regards the collapse. The energy spectrum of the collective elementary excitations consists of two exciton-type branches (energy and quasienergy branches) each of them with energy gap and roton-type section, the gapless optical plasmon branch, and the acoustical plasmon branch, which reveals the absolute instability in the range of small wave vectors.

1. Introduction

Properties of atoms and excitons are dramatically changed in strong magnetic fields, such that the distance between Landau levels $\hbar\omega_c$, exceeds the corresponding Rydberg energies R_y and the magnetic length $l = \sqrt{\hbar c / eH}$ is small compared to their Bohr radii [1, 2]. Even more interesting phenomena are exhibited in the case of two-dimensional (2D) electron systems due to the quenching of the kinetic energy at high magnetic fields, with the representative example being integer and fractional Quantum Hall effects [3-5]. The discovery of the FQHE [6-8] changed fundamentally the established concepts about charged elementary excitations in solids [5]. The notion of the incompressible quantum liquid (IQL) was introduced in [7] as a homogeneous phase with the quantized densities $\nu = p/q$, where p is an integer and $q \neq 1$ is odd having charged elementary excitations with a fractional charge $e^* = \pm e/q$. These quasiparticles were named anyons. A classification for free anyons and their hierarchy were studied in [9, 10]. An alternative concept to hierarchical scheme was proposed in [11], where the notion of composite fermions (CF) was introduced. The CF consists of the electron bound to an even number of flux quanta. Within the frame of this concept, the FQHE of electrons can be physically understood as a manifestation of the IQHE of CFs [11]. The statistics of anyons was determined in [10, 12]. It was established that the wave function of the system changes by a complex phase factor $\exp[i\pi\alpha]$, when the quasiparticles are interchanged. For bosons $\alpha = 0$, for fermions $\alpha = 1$, and for anyons with $e^* = -e/3$, their statistical charge is $\alpha = -1/3$. As was shown

in [13], there were no soft branches of neutral excitations in IQL. The energy gap Δ for formation of a quasielectron-quasihole pair has the scale of Coulomb energy $E_Q = e^2/\varepsilon l$, where ε is the dielectric constant of the background. However, delta was found to be small $\Delta \simeq 0.1E_Q$. The lowest branch was called magnetoroton [13]; it can be modelled as a quasiexciton [5]. As was mentioned in [5], the traditional methods and concepts based either on the neglecting of the electron-electron interaction or on self-consistent approximation are inapplicable to IQL. In a strong magnetic field, the binding energy of an exciton increases from R_y to I_l.

There are two small parameters of the theory. One of them determines how strong the magnetic field strength H is, and it verifies whether the starting supposition of a strong magnetic field is fulfilled. This parameter is expressed by the ratio $I_l/\hbar\omega_c < 1$. Here I_l is the magnetoexciton ionization potential, ω_c is the cyclotron frequency $eH/\mu c$ calculated with the reduced mass μ and the magnetic length l. Another small parameter has a completely different origin and is related with the concentration of the electron-holes(e-h) pairs. In our case, it can be expressed as a product of the filling factor $\nu = \mathrm{v}^2$ and of another factor $(1-\mathrm{v}^2)$ which reflects the Pauli exclusion principle and the phase-space filing (PSF) effect. This compound parameter $\mathrm{v}^2(1-\mathrm{v}^2)$ in the case of Bose-Einstein condensed excitons can take the form $u^2\mathrm{v}^2$, where u and v are the Bogoliubov transformation coefficients and $u^2 = (1-\mathrm{v}^2)$. The two small parameters will be used below. However, in the case of FQHE, the filling factor $\nu = \mathrm{v}^2$ basically determines the underlying physics and it can not be changed arbitrarily. Instead of the perturbation theory on the filling factor ν, the exact numerical diagonalization for a few number of particles $N \le 10$ proved to be the most powerful tool in studies of such systems [5]. The spherical geometry for these calculations was proposed [10, 14], considering a few number of particles on the surface of a sphere with the radius $R = \sqrt{S}l$, so as the density of the particles on the sphere to be equal with the filling factor of 2DEG. The magnetic monopole in the center of the sphere creates a magnetic flux through the sphere $2S\phi_0$, which is multiple to the flux quantum $\phi_0 = 2\pi\hbar c/e$. The angular momentum L of a quantum state on the sphere and the quasimomentum k of the FQHE state on the plane obey the relation $L = Rk$. Spherical model is characterized by continuous rotational group, which is analogous to the continuous translational symmetry in the plane.

The properties of the symmetric 2D electron-hole (e-h) system (i.e., $h = 0$), with equal concentrations for both components, with coincident matrix elements of Coulomb electron-electron, hole-hole, and electron-hole interactions in a strong perpendicular magnetic field also attracted much attention during last two decades [15-22]. A hidden symmetry and the multiplicative states were discussed in many papers [19, 23, 24]. The collective states such as the Bose-Einstein condensation (BEC) of two-dimensional magnetoexcitons and the formation of metallic-type electron-hole liquid (EHL) were investigated in [15-22]. The search for Bose-Einstein condensates has become a milestone in the condensed matter physics [25]. The remarkable properties of super fluids and superconductors are intimately related to the existence of a bosonic condensate of composite particles consisting of an even number of fermions. In highly excited semiconductors, the role of such composite bosons is taken on by excitons, which are bound states of electrons and holes. Furthermore, the excitonic system has been viewed as a keystone system for exploration of the BEC phenomena, since it allows to control particle densities and interactions *in situ*. Promising candidates for experimental realization of such system are semiconductor quantum wells (QWs) [26], which have a number of advantages compared to the bulk systems. The coherent pairing of electrons and holes occupying

only the lowest Landau levels (LLLs) was studied using the Keldysh-Kozlov-Kopaev method and the generalized random-phase approximation [20, 27]. The BEC of magnetoexcitons takes place in a single exciton state with wave vector *k*, supposing that the high density of electrons in the conduction band and of holes in the valence band was created in a single QW structure with size quantization much greater than the Landau quantization. In the case $k \neq 0$, a new metastable dielectric liquid phase formed by Bose-Einstein condensed magnetoexcitons was revealed [20, 21]. The importance of the excited Landau levels (ELLs) and their influence on the ground states of the systems was first noticed by the authors of papers [16-19]. The influence of the ELLs of electrons and holes was discussed in detail in papers [21, 22]. The indirect attraction between electrons (e-e), between holes (h-h), and between electrons and holes (e-h) due to the virtual simultaneous quantum transitions of the interacting charges from LLLs to the ELLs is a result of their Coulomb scattering. The first step of the scattering and the return back to the initial states were described in the second order of the perturbation theory.

Das Sarma and Madhukar [28] have investigated theoretically the longitudinal collective modes of spatially separated two-component two-dimensional plasma in solids using the generalized random phase approximation. It can be implemented in semiconductors heterojunctions and superlattices. The two-layer structure with two-component plasma is discussed below. It has long been known that two-component plasma has two branches of its longitudinal oscillations. The higher frequency branch is named optical plasmon (OP). Here the two carrier densities of the same signs oscillate in-phase and their density fluctuation operators $\hat{\rho}_{e,1}(\vec{Q})$ and $\hat{\rho}_{e,2}(\vec{Q})$ form an in-phase superposition

$$\hat{\rho}_{OP}(\vec{Q}) = \hat{\rho}_{e,1}(\vec{Q}) + \hat{\rho}_{e,2}(\vec{Q}).$$

In the case of opposite signs of electron and hole charges, they oscillate out-of-phase and their charge density fluctuation operators $\hat{\rho}_e(\vec{Q})$ and $\hat{\rho}_h(\vec{Q})$ combine in out-of-phase manner

$$\hat{\rho}_{OP}(\vec{Q}) = \hat{\rho}_e(\vec{Q}) - \hat{\rho}_h(-\vec{Q}).$$

The lower frequency branch is named acoustical plasmon (AP). Now the carriers of different signs oscillate in-phase, whereas the carriers of the same signs oscillate out-of-phase. Their charge density fluctuation operators combine in the form

$$\hat{\rho}_{AP}(\vec{Q}) = \hat{\rho}_{e,1}(\vec{Q}) - \hat{\rho}_{e,2}(\vec{Q});\ \ \hat{\rho}_{AP}(\vec{Q}) = \hat{\rho}_e(\vec{Q}) + \hat{\rho}_h(-\vec{Q}).$$

The optical and acoustical branches of two-component electron plasma have the dispersion relations in the long wavelength region as follows

$$\omega_{OP}(q) \sim \sqrt{q};\ \ \omega_{AP}(q) \sim q;\ \ q \to 0.$$

The plasmon oscillations in one-component system on the monolayer in a strong perpendicular magnetic field were studied by Girvin, MacDonald, and Platzman [13], who proposed the magnetoroton theory of collective excitations in the conditions of the fractional quantum Hall effect (FQHE). The FQHE occurs in low-disorder, high-mobility samples with partially filled Landau levels with filling factor of the form $\nu = 1/m$, where m is an integer, for which there is no single-particle gap. In this case, the excitation is a collective effect arising from many-body correlations due to the Coulomb interaction. Considerable progress has recently been achieved toward understanding the nature of the many-body ground state well described by Laughlin variational wave function [7]. The theory of the collective excitation spectrum proposed by [13] is closely analogous to the Feynman's theory of superfluid helium [29]. The main Feynman's arguments lead to the conclusions that, on general grounds, the low lying excitations of any system will include density waves. As regards the 2D system, the perpendicular magnetic field quenches the single particle continuum of kinetic energy leaving a series of dis-

crete highly degenerate Landau levels spaced in energy at intervals $\hbar\omega_c$. In the case of filled Landau level $\nu = 1$ because of Pauli exclusion principle, the lowest excitation is necessarily the cyclotron mode in which particles are excited into the next Landau level. In the case of FQHE, the LLL is fractionally filled. The Pauli principle no longer excludes low-energy intra-Landau-level excitations. For the FQHE case, the low-lying excitations, rather than the high-energy inter-Landau-level cyclotron modes, are of the primary importance [13]. The spectrum has a relatively large excitation gap at zero wave vector $kl = 0$; in addition, it exhibits a deep magneto-roton minimum at $kl \sim 1$ quite analogous to the roton minimum in helium. The magneto-roton minimum becomes deeper and deeper with decreasing filling factor ν in the row 1/3, 1/5, 1/7; it is the precursor to the gap collapse associated with the Wigner crystallization, which occurs at $\nu = 1/7$. For largest wave vectors, the low lying mode crosses over from being a density wave to becoming a quasiparticle excitation [13]. The Wigner crystal transition occurs slightly before the roton mode goes completely soft. The magnitude of the primitive reciprocal lattice vector for the crystal lies close to the position of the magneto-roton minimum. The authors of [13] suggested also the possibility of pairing of two rotons of opposite momenta leading to the bound two-roton state with small total momentum, as it is known to occur in helium. In difference from the case of fractional filling factor, the excitations from a filled Landau level in the 2DEG were studied by Kallin and Halperin [30].

Fertig [31] investigated the excitation spectrum of two-layer and three-layer electron systems. In a particular case, the two-layer system in a strong perpendicular magnetic field with filling factor $\nu = 1/2$ of the LLL in the conduction band of each layer was considered. Inter-layer separation z was introduced. The spontaneous coherence of two-component two-dimensional (2D) electron gas was introduced.

Fertig has determined the energy spectrum of the elementary excitations within the frame of this ground state. In the case of z = 0, the lowest-lying excitations of the system are the higher energy excitons.

Because of the neutral nature of the $\vec{k} = 0$ excitons, the dispersion relation of these excitations is given in a good approximation by $\hbar\omega(k) = E_{ex}(k) - E_{ex}(0)$, where $E_{ex}(k)$ is the energy of exciton with wave vector $\vec{k}$. This result was first obtained by Paquet, Rice, and Ueda [19] using a random phase approximation (RPA). In the case z = 0, the dispersion relation $\omega(k)$ vanishes as k^2 for $k \to 0$, as one expects for Goldstone modes. As was shown by Fertig [31], for z > 0, $\omega(k)$ behaves as an acoustical mode $\omega(k) \sim k$ in the range of small k, whereas in the limit $k \to \infty$ $\omega(k)$ tends to the ionization potential $\Delta(z)$.

In the region of intermediate values of k, when $kl \sim 1$, the dispersion relation develops the dips as z is increased. At certain critical value of $z=z_{er}$, the modes in the vicinity of the minima become equal to zero and are named soft modes. Their appearance testifies that the two-layer system undergoes a phase transition to the Wigner crystal state.

The similar results concerning the linear and quadratic dependences of the dispersion relations in the range of small wave vectors q were obtained by Kuramoto and Horie [32], who studied the coherent pairing of electrons and holes spatially separated by the insulator barrier in the structure of the type coupled double quantum wells (CDQW).

2. Hamiltonian of the supplementary interaction

The Hamiltonian of the Coulomb interaction of the electrons and holes within the frame LLLs has the form

$$\hat{H} = \frac{1}{2}\sum_{\vec{Q}} W_{\vec{Q}} \left[\hat{\rho}(\vec{Q})\hat{\rho}(-\vec{Q}) - \hat{N}_e - \hat{N}_h \right] - \mu_e \hat{N}_e - \mu_h \hat{N}_h + \hat{H}_{\text{suppl}}, \tag{1}$$

where $W_{\vec{Q}}$ is the Fourier transform of the Coulomb interaction within the frame of LLLs, $\hat{N}_e$ and $\hat{N}_h$ are the operators of the numbers of electrons and holes on the LLLs. They are determined below. $\hat{H}_{\text{suppl}}$ is the supplementary indirect attractive interaction between the particle lying on the LLLs in view of their virtual transitions on the ELLs and their return back [22]

$$H_{\text{suppl}} = -\frac{1}{2}\sum_{p,q,s} \phi_{e-e}(p,q;s) a_p^\dagger a_q^\dagger a_{q+s} a_{p-s} - \\ -\frac{1}{2}\sum_{p,q,s} \phi_{h-h}(p,q;s) b_p^\dagger b_q^\dagger b_{q+s} b_{p-s} - \sum_{p,q,s} \phi_{e-h}(p,q;s) a_p^\dagger b_q^\dagger b_{q+s} a_{p-s}. \tag{2}$$

Here the creation and annihilation operators $a_p^\dagger, a_p$ for electrons and $b_q^\dagger, b_q$ for holes were introduced. The matrix elements of indirect interaction $\phi_{i-j}(p,q,z)$ are described by the common expressions [22]

$$\phi_{i-j}(p,q,s) = \sum_{n,m} \frac{\phi_{i-j}(p,q,z;n,m)}{n\hbar\omega_{ci} + m\hbar\omega_{cj}}. \tag{3}$$

In the case of electron-electron and hole-hole interaction, expression (3) has the form [22]

$$\phi_{i-i}(p,q,z;n,m) \cong \sum_{t,\kappa,\sigma} W_{t,k} W_{z-t,\sigma} \exp\left(i\kappa(p-q-t)l^2\right) \times \\ \times \exp\left(i\sigma(p-q-t-z)l^2\right)(t+i\kappa)^{n+m}(t-z+i\sigma)^{n+m}, \tag{4}$$

but in the case of electron-hole interaction, it is

$$\phi_{e-h}(p,q,z;n,m) \cong \sum_{t,\kappa,\sigma} W_{t,k} W_{z-t,\sigma} \exp\left(i(\kappa+\sigma)(p+q)l^2\right) \times \\ \times (t+i\kappa)^n (t-i\kappa)^m (t-z+i\sigma)^n (t-z-i\sigma)^m, \tag{5}$$

where

$$W_{s,\kappa} = \frac{2\pi e^2}{\varepsilon_0 S\sqrt{s^2+k^2}} \exp\left[-\frac{\left(s^2+k^2\right)l^2}{2}\right], \qquad W_{s,k} = W_{-s,-k} = W_{-s,k} = W_{s,-k}. \tag{6}$$

Hamiltonian (2) has a Hermitian conjugate form, if the requirements are fulfilled

$$\phi_{i-j}^*(p-s,q+s;-s) = \phi_{i-j}(p,q;s), \qquad i,j = e,h. \tag{7}$$

Hermiticity requirement (7) can be deduced, for example, in the case of electron-electron interaction as follows

$$\phi_{i-i}^*(p-z,q+z;-z;n,m) \cong \sum_{t,\kappa,\sigma} W_{t,\kappa} W_{-z-t,\sigma} \exp(-i\kappa(p-q-2z-t)l^2) \times \\ \times \exp(-i\sigma(p-q-t-z)l^2)(t-i\kappa)^{n+m}(t+z-i\sigma)^{n+m}. \tag{8}$$

Introducing the new summation variables

$$t = t'-z, \qquad \kappa = -\sigma', \qquad \sigma = -\kappa' \tag{9}$$

and taking into account the properties (5) we will obtain exactly expression (4), what proves the affirmation. In the same way, we can write

$$\phi_{e-h}^*(p-z,q+z;-z;n,m) \cong \sum_{t,\kappa,\sigma} W_{t,\kappa} W_{t+z,\sigma} \exp(-i(\sigma+\kappa)(p+q)l^2) \times \\ \times (t-i\kappa)^n (t+i\kappa)^m (t+z-i\sigma)^n (t+z+i\sigma)^m, \tag{10}$$

which after substitution (9) coincides with expression (5). There are two other properties of the coefficients $\phi_{i-j}(p,q;s)$, namely, their reality and parity, that is,

$$\phi^*_{i-j}(p,q;s)=\phi_{i-j}(p,q;s);\ \phi_{i-j}(-p,-q;-s)=\phi_{i-j}(p,q;s). \tag{11}$$

They can be proved as was demonstrated above using the substitution $\sigma=-\sigma'$ and $\kappa=-\kappa'$, when the reality is considered and the substitution $t=-t'$, $\sigma=-\sigma'$, and $\kappa=-\kappa'$ when the parity is discussed.

Side by side with the properties demonstrated above, there is another property related with the translational symmetry of the system in one in-plane direction, which does exist in the Landau gauge description. As a result, the coefficients $\phi_{i-j}(p,q;s)$ do not depend separately on the variables p and q but in their linear combination as follows

$$\begin{aligned}\phi_{i\text{-}i}(p,q;s)&=\tilde{\phi}_{i\text{-}i}(s,\kappa); \qquad \kappa=p-q-s;\\ \phi_{e-h}(p,q;s)&=\tilde{\phi}_{e-h}(s,\sigma); \qquad \sigma=p+q.\end{aligned} \tag{12}$$

They have the properties

$$\tilde{\phi}^*_{i-j}(-s,\sigma)=\tilde{\phi}_{i-j}(s,\sigma);\ \tilde{\phi}^*_{i-j}(s,\sigma)=\tilde{\phi}_{i-j}(s,\sigma);\ \tilde{\phi}_{i-j}(-s,-\sigma)=\tilde{\phi}_{i-j}(s,\sigma). \tag{13}$$

Their Fourier transforms are

$$\psi_{i-j}(s,\sigma)=\sum_{\kappa}\tilde{\phi}_{i-j}(s,\kappa)\exp(i\kappa\sigma l^2). \tag{14}$$

Their symmetry properties follow directly from the previous ones

$$\begin{aligned}\psi^*_{i-j}(s,\sigma)&=\psi_{i-j}(-s,-\sigma); \quad \text{hermiticity}\\ \psi^*_{i-j}(s,\sigma)&=\psi_{i-j}(s,-\sigma); \quad \text{reality}\\ \psi_{i-j}(-s,-\sigma)&=\psi_{i-j}(s,\sigma). \quad \text{parity}.\end{aligned} \tag{15}$$

They lead to the conclusion

$$\psi^*_{i-j}(s,\sigma)=\psi_{i-j}(s,\sigma). \tag{16}$$

These properties will be used below during the transformation of Hamiltonian (2) written in terms of the single particle operators $a^\dagger_p, a_p, b^\dagger_p, b_p$ to the form expressed through the two-particle operators of the electron and hole densities $\hat{\rho}_e(\vec{Q})$ and $\hat{\rho}_h(\vec{Q})$ of the type

$$\hat{\rho}_e(\vec{Q})=\sum_t e^{iQ_y tl^2}a^\dagger_{t-\frac{Q_x}{2}}a_{t+\frac{Q_x}{2}};\ \hat{\rho}_h(\vec{Q})=\sum_t e^{iQ_y tl^2}b^\dagger_{t+\frac{Q_x}{2}}b_{t-\frac{Q_x}{2}}. \tag{17}$$

The relations between two sets of operators are

$$\begin{aligned}a^\dagger_{p-\frac{s}{2}}a_{p+\frac{s}{2}}&=\frac{1}{N}\sum_\kappa\hat{\rho}_e(s,\kappa)\exp(-i\kappa pl^2);\\ a^\dagger_p a_{p-s}&=\frac{1}{N}\sum_\kappa\hat{\rho}_e(-s,\kappa)\exp\left(-i\kappa pl^2+\frac{is\kappa}{2}l^2\right);\\ a^\dagger_q a_{q+s}&=\frac{1}{N}\sum_\kappa\hat{\rho}_e(s,\kappa)\exp\left(-i\kappa ql^2-\frac{is\kappa}{2}l^2\right),\end{aligned} \tag{18}$$

where $N=S/2\pi l^2$, S is the layer surface area and l is the magnetic length. Here the $\delta-$Kronecker symbol was used

$$\frac{1}{N}\sum_p\exp(ip(\sigma-\kappa)l^2)=\delta_{kr}(\sigma,\kappa). \tag{19}$$

Taking into account that

$$\sum_{p,q,s}\phi_{e-e}(p,q;s)a_p^{\dagger}a_{p-s}a_q^{\dagger}a_{q+s}=\frac{1}{N}\sum_{s,\sigma}\psi_{e-e}(s,\sigma)\hat{\rho}_e(-s,-\sigma)\hat{\rho}_e(s,\sigma),$$
$$\sum_{s}\phi_{e-e}(p,p-s;s)=\sum_{s}\tilde{\phi}_{e-e}(s,0)=B_{i-i}, \tag{20}$$
$$\sum_{p}a_p^{\dagger}a_p=\hat{N}_e;\quad \sum_{p}b_p^{\dagger}b_p=\hat{N}_h;\qquad \hat{N}=\hat{N}_e+\hat{N}_h$$

and the similar expressions for the hole-hole interaction, we can write

$$\frac{1}{2}\sum_{p,q,s}\phi_{e-e}(p,q;s)a_p^{\dagger}a_q^{\dagger}a_{q+s}a_{p-s}+\frac{1}{2}\sum_{p,q,s}\phi_{h-h}(p,q;s)b_p^{\dagger}b_q^{\dagger}b_{q+s}b_{p-s}=$$
$$=-\frac{1}{2}B_{i-i}\hat{N}+\frac{1}{2N}\sum_{s,\sigma}\psi_{i-i}(s,\sigma)\Big[\hat{\rho}_e(-s,-\sigma)\hat{\rho}_e(s,\sigma)+\hat{\rho}_h(-s,-\sigma)\hat{\rho}_h(s,\sigma)\Big]. \tag{21}$$

The supplementary electron-hole interaction can be transformed as follows

$$\sum_{p,q,s}\phi_{e-h}(p,q;s)a_p^{\dagger}b_q^{\dagger}b_{q+s}a_{p-s}=\frac{1}{N}\sum_{s,\sigma}\psi_{e-h}(s,\sigma)\hat{\rho}_e(-s,-\sigma)\hat{\rho}_h(-s,-\sigma). \tag{22}$$

The Hamiltonian of supplementary indirect attractive interaction (2) has the form

$$H_{\text{suppl}}=\frac{1}{2}B_{i-i}\hat{N}-\frac{1}{2N}\sum_{s,\sigma}\psi_{i-i}(s,\sigma)\Big[\hat{\rho}_e(-s,-\sigma)\hat{\rho}_e(s,\sigma)+\hat{\rho}_h(-s,-\sigma)\hat{\rho}_h(s,\sigma)\Big]-$$
$$-\frac{1}{N}\sum_{s,\sigma}\psi_{e-h}(s,\sigma)\hat{\rho}_e(-s,-\sigma)\hat{\rho}_h(-s,-\sigma). \tag{23}$$

Instead of density operators for electrons and holes, we can introduce their in-phase and in opposite-phase linear combinations

$$\hat{\rho}(\vec{Q})=\hat{\rho}_e(\vec{Q})-\hat{\rho}_h(-\vec{Q});\ \ \hat{D}(\vec{Q})=\hat{\rho}_e(\vec{Q})+\hat{\rho}_h(-\vec{Q});$$
$$\hat{\rho}_e(\vec{Q})=\frac{1}{2}\Big[\hat{\rho}(\vec{Q})+\hat{D}(\vec{Q})\Big];\ \ \hat{\rho}_h(\vec{Q})=\frac{1}{2}\Big[\hat{D}(-\vec{Q})-\hat{\rho}(-\vec{Q})\Big]. \tag{24}$$

They lead to the following relations

$$\hat{\rho}_e(-\vec{Q})\hat{\rho}_e(\vec{Q})+\hat{\rho}_h(-\vec{Q})\hat{\rho}_h(\vec{Q})=\frac{1}{2}\Big[\hat{\rho}(-\vec{Q})\hat{\rho}(\vec{Q})+\hat{D}(-\vec{Q})\hat{D}(\vec{Q})\Big];$$
$$\sum_{Q}\psi_{e-h}(Q)\Big[\hat{\rho}(-\vec{Q})\hat{D}(\vec{Q})-\hat{D}(-\vec{Q})\hat{\rho}(\vec{Q})\Big]=\sum_{Q}\psi_{e-h}(Q)\Big[\hat{\rho}(-\vec{Q})\hat{D}(\vec{Q})-\hat{D}(\vec{Q})\hat{\rho}(-\vec{Q})\Big]=0$$

and to the final expression

$$H_{\text{suppl}}=\frac{1}{2}B_{i-i}\hat{N}-\frac{1}{4N}\sum_{Q}V(Q)\hat{\rho}(\vec{Q})\hat{\rho}(-\vec{Q})-\frac{1}{4N}\sum_{Q}U(Q)\hat{D}(\vec{Q})\hat{D}(-\vec{Q}), \tag{25}$$

where

$$U(Q)=\psi_{i-i}(Q)+\psi_{e-h}(Q);$$
$$V(Q)=\psi_{i-i}(Q)-\psi_{e-h}(Q). \tag{26}$$

The estimations show that

$$U(0)=2A_{i-i};\ \ V(0)=0;\ \ \frac{1}{N}\sum_{\vec{Q}}U(\vec{Q})=B_{i-i}+\Delta(0).$$

It means that one can suppose the dependences

$$U(\vec{Q})\cong U(0)e^{-\frac{Q^2l^2}{2}};\quad V(\vec{Q})\cong V(0)=0. \tag{27}$$

3. Bose-Einstein Condensation of magnetoexcitons in two alternative descriptions

BEC of 2D magnetoexcitons was considered in [20, 21] within the frame of Keldysh-Kozlov-Kopaev method [27], when the influence of the ELLs was neglected. The main results of this description will be mentioned below.

The creation $d^{\dagger}(\vec{P})$ and annihilation $d(\vec{P})$ operators of the 2D magnetoexciton have the form

$$d^{\dagger}(\vec{P}) = \frac{1}{\sqrt{N}} \sum_{t} e^{-iP_y t l^2} a^{\dagger}_{t+\frac{P_x}{2}} b^{\dagger}_{-t+\frac{P_x}{2}};$$
$$d(\vec{P}) = \frac{1}{\sqrt{N}} \sum_{t} e^{iP_y t l^2} b_{-t+\frac{P_x}{2}} a_{t+\frac{P_x}{2}}. \tag{28}$$

The energy of the two-dimensional magnetoexciton $E_{ex}(P)$ depends on the two-dimensional wave vector $\vec{P}$ and forms a band with the dependence

$$E_{ex}(\vec{P}) = -I_{ex}(\vec{P}) = -I_l + E(\vec{P});$$
$$I_{ex}(\vec{P}) = I_l e^{-\frac{P^2 l^2}{4}} I_0\left(\frac{P^2 l^2}{4}\right); \qquad I_l = \frac{e^2}{\varepsilon_0 l}\sqrt{\frac{\pi}{2}};\ \sum_{\vec{Q}} W_{\vec{Q}} = I_l. \tag{29}$$

The ionization potential $I_{ex}(P)$ is expressed through the modified Bessel function $I_0(z)$, which has the limiting expressions

$$\underset{z\to 0}{I_0(z)} = 1 + \frac{z^2}{4} + \ldots;\ \underset{z\to\infty}{I_0(z)} = \frac{e^z}{\sqrt{2\pi z}}. \tag{30}$$

It means that the function $E(P)$ can be approximated as follows

$$\underset{P\to 0}{E(\vec{P})} = \frac{\hbar^2 P^2}{2M}; \qquad M = M(0) = 2\sqrt{\frac{2}{\pi}}\frac{\hbar^2\varepsilon_0}{e^2 l}; \qquad \underset{P\to\infty}{E(P)} = I_l\left(1 - \frac{\sqrt{2/\pi}}{Pl}\right); \qquad l^2 = \frac{\hbar c}{eH}. \tag{31}$$

To introduce the phenomenon of BEC of excitons, the gauge symmetry of the initial Hamiltonian was broken by means of the unitary transformation $\hat{D}(\sqrt{N_{ex}})$ following the Keldysh-Kozlov-Kopaev method [27]. We can shortly remember the main outlines of the Keldysh-Kozlov-Kopaev method [27], [33] as it was done in papers [20, 21]. The unitary transformation $\hat{D}(\sqrt{N_{ex}})$ was determined by formula (8) [20]. Here N_{ex} is the number of condensed excitons. It transforms the operators a_p, b_p to other ones α_p, β_p, as is shown in formulas (13) and (14) [20], and gives rise to the BCS-type wave function $\left|\psi_g(\bar{k})\right\rangle$ of the new coherent macroscopic state represented by expression (10) [20]. These results are summarized below

$$\hat{D}(\sqrt{N_{ex}}) = \exp[\sqrt{N_{ex}}(d^{\dagger}(\bar{k}) - d(\bar{k}))];\ \left|\psi_g(\bar{k})\right\rangle = \hat{D}(\sqrt{N_{ex}})\left|0\right\rangle;$$
$$\alpha_p = \hat{D}a_p\hat{D}^{\dagger} = ua_p - \mathrm{v}(p - \frac{k_x}{2})b^{\dagger}_{k_x - p};\ \beta_p = \hat{D}b_p\hat{D}^{\dagger} = ub_p + \mathrm{v}(\frac{k_x}{2} - p)a^{\dagger}_{k_x - p}; \tag{32}$$
$$a_p = u\alpha_p + \mathrm{v}(p - \frac{k_x}{2})\beta^{\dagger}_{k_x - p};\ b_p = u\beta_p - \mathrm{v}(\frac{k_x}{2} - p)\alpha^{\dagger}_{k_x - p}.$$

$$a_p\left|0\right\rangle = b_p\left|0\right\rangle = 0; \qquad \alpha_p\left|\psi_g(\bar{k})\right\rangle = \beta_p\left|\psi_g(\bar{k})\right\rangle = 0;$$

$$u = \cos g;\ \mathrm{v} = \sin g;\ \mathrm{v}(t) = \mathrm{v}e^{-ik_y t l^2};\ g = \sqrt{2\pi l^2 n_{ex}};\ n_{ex} = \frac{N_{ex}}{S} = \frac{\mathrm{v}^2}{2\pi l^2};\ g = \mathrm{v};\ \mathrm{v}=\mathrm{Sin}v. \tag{33}$$

The developed theory [20, 21] holds true in the limit $v^2 \approx Sin^2 v$, what means the restriction $v^2 < 1$. Within the frame of this approach, the collective elementary excitations can be studied constructing the Green's functions on the base of operators α_p, β_p and dealing with the transformed cumbersome Hamiltonian $\hat{\tilde{\mathcal{H}}} = D(\sqrt{N_{ex}})\hat{H}D^{\dagger}(\sqrt{N_{ex}})$.

We propose another way, which is supplementary but completely equivalent to the previous one and is based on the idea suggested by Bogoliubov in his theory of quasiaverages [34]. Considering the case of a 3D ideal Bose gas with the Hamiltonian

$$H = \sum_{\vec{p}} \left(\frac{\hbar^2 p^2}{2m} - \mu \right) a^{\dagger}_{\vec{p}} a_{\vec{p}} , \tag{34}$$

where a_p^{+}, a_p are the Bose operators and μ is the chemical potential, Bogoliubov added the term

$$-\eta\sqrt{V}(a_0 e^{i\varphi} + a_0 e^{-i\varphi}) \tag{35}$$

breaking the gauge symmetry and proposed to consider the BEC on the state with $p = 0$ within the frame of the Hamiltonian

$$\hat{\mathcal{H}} = \sum_{p} \left(\frac{\hbar^2 p^2}{2m} - \mu \right) a^{\dagger}_p a_p - \eta\sqrt{V}(a^{\dagger}_0 e^{i\varphi} + a_0 e^{-i\varphi}), \tag{36}$$

where

$$\eta = -\mu\sqrt{\frac{N_0}{V}} = -\mu\sqrt{n_0}; \qquad -\frac{\eta}{\mu} = \sqrt{n_0}. \tag{37}$$

We will name the Hamiltonian of type (36) the Hamiltonian of the theory of quasiaverages. It is written in terms of the operators a_p^{+}, a_p of the initial Hamiltonian (34).

Our intention is to apply this idea to the case of BEC of interacting 2D magnetoexcitons and to deduce explicitly the Hamiltonian of type (36) with the finite parameters μ and η but with the relation of type (37). We intend to formulate the new Hamiltonian with broken symmetry in terms of the operators a_p, b_p avoiding the obligatory crossing to the operators α_p, β_p (32) at least at some stages of the investigation where the representation in the a_p, b_p operators remains preferential.

It is obvious that the two representations are completely equivalent and complement each other. We will follow quasiaverage variant (36) instead of u, v variant (32, 33), because it opens some new possibilities, which have not been studied up till now, to the best of our knowledge. For example, the Hamiltonian of type (36) is simpler than the Hamiltonian $\hat{\tilde{\mathcal{H}}} = D(\sqrt{N_{ex}})\hat{H}D^{\dagger}(\sqrt{N_{ex}})$ in the α_p, β_p representation, and the deduction of the equation of motion for operators (35) and for the many-particle Green's functions constructed on their base is also much simpler. We will take this advantage at some stages of investigation. On the contrary, for the calculations of the average values of different operators on the base of the ground coherent macroscopic state (27) or using the coherent excited states, as we have done in papers [20, 21], the most convenient way is to use the α_p, β_p representation. We will use, in a wide manner, the two representations. The new variant in the style of the theory of quasiaverages can be implemented rewriting the transformed Hamiltonian $D(\sqrt{N_{ex}})\hat{H}D^{\dagger}(\sqrt{N_{ex}})$ in the a_p, b_p representation as follows below. To demonstrate it, we will represent the unitary transformation

$$\hat{D}(\sqrt{N_{ex}}) = e^{\hat{X}} = \sum_{n=0}^{\infty} \frac{\hat{X}^n}{n!};\ D^{\dagger}(\sqrt{N_{ex}}) = e^{-\hat{X}}, \tag{38}$$

where

$$\hat{X} = \sqrt{N_{ex}}(e^{i\varphi} d^{\dagger}(K) - e^{-i\varphi} d(K));\ \hat{X}^{\dagger} = -\hat{X}. \tag{39}$$

The creation and annihilation operators $d^{+}(k)$, $d(k)$ are written in the Landau gauge when the electrons and holes forming the magnetoexcitons are situated on their LLLs. This variant was considered firstly without taking into account of the ELLs, as one can see in [20]. The BEC of 2D magnetoexcitons was considered on the single-exciton state characterized by two-dimensional wave vector $\vec{k}$. Expanding in series the unitary operators $D(\sqrt{N_{ex}}), D^{\dagger}(\sqrt{N_{ex}})$, we find the transformed operator $\hat{\tilde{\mathcal{H}}}$ in the form

$$\hat{\tilde{\mathcal{H}}} = e^{\hat{X}} \hat{H} e^{-\hat{X}} = \hat{H} + \frac{1}{1!}\left[\hat{X}, \hat{H}\right] + \frac{1}{2!}\left[\hat{X}, \left[\hat{X}, \hat{H}\right]\right] + \frac{1}{3!}\left[\hat{X}, \left[\hat{X}, \left[\hat{X}, \hat{H}\right]\right]\right] + ... = \hat{\mathcal{H}} + \hat{\mathcal{H}}'. \tag{40}$$

Here the Hamiltonian $\hat{\mathcal{H}}$ contains the main contributions of the first two terms in the series expansion (40), whereas the operator $\hat{\mathcal{H}}'$ gathers the all remaining terms.

As one can see looking at formulas (39), the operator $\hat{X}$ is proportional to the square root of the exciton concentration $\sqrt{N_{ex}}$, which is proportional to the filling number ν. One can see that the contributions arising from the first commutator $\left[\hat{X}, \hat{H}\right]$ are proportional to ν, the contributions arising from the second commutator $\left[\hat{X}, \left[\hat{X}, \hat{H}\right]\right]$ are proportional to ν^2 and so on. Following the Bogoliubov's theory of quasiaverages, the linear terms of the type $\left(d^{+}(k)e^{i\varphi} + e^{-i\varphi} d(k)\right)\nu$ arising from the first commutator $\left[\hat{X}, \hat{H}\right]$ were included into $\hat{\mathcal{H}}$.

The Hamiltonian $\hat{\mathcal{H}}$ with the broken gauge symmetry describing the BEC of 2D magnetoexcitons on the state with wave vector $k \neq 0$ being written in the style of the Bogoliubov's theory of quasiaverages has the form

$$\hat{\mathcal{H}} = \hat{H} + \sqrt{N_{ex}}(\bar{\mu} - E(\vec{K}))(e^{i\varphi} d^{\dagger}(\vec{K}) + e^{-i\varphi} d(\vec{K})). \tag{41}$$

For the case of an ideal 2D Bose gas, we rewrite the coefficient $-\eta\sqrt{V}$ in Hamiltonian (36) in the form $-\eta\sqrt{N}$ and, comparing it with deduced expression (41), we find

$$\eta = \left(E(k) - \bar{\mu}\right)\nu. \tag{42}$$

Relation (42) coincides exactly with relation (20) of the Bogoliubov's theory of quasiaverages. In the case of ideal Bose gas, η and $\left(E(k) - \bar{\mu}\right)$ both tend to zero, whereas the filling number is real and different from zero. In the case of interacting exciton gas, both the parameter η and $\left(E(k) - \bar{\mu}\right)$ are different from zero.

The chemical potential μ was determined in the HFB approximation in [21, 22]. In the first of them, only the simplest case of first ELLs was discussed; in the second one, a more general case representing the influence of the all ELLs was described. We shall mention the last results. They were obtained making the (uv) transformation (32) from the initial operators a_p, b_p to new operators α_p, β_p in the starting Hamiltonian H (1). After its normal ordering within the frame of the operators $\alpha_p^{\dagger}, \alpha_p, \beta_p^{\dagger}, \beta_p$, the transformed Hamiltonian $DHD^{\dagger}$ will contain a constant part playing the role of ground state energy, a quadratic Hamiltonian H_2

containing the diagonal terms of the type $\alpha_p^\dagger, \alpha_p$ and $\beta_p^\dagger, \beta_p$ as well as the nondiagonal terms of the type $\alpha_p^\dagger \beta_{k_x-p}^\dagger$ and $\beta_{k_x-p}\alpha_p$, and a quartic Hamiltonian H', which is neglected in the HFB approximation.

The quadratic Hamiltonian H_2 was represented by formula (32) in [22], which is repeated below

$$H_2 = \sum_p [E(\mathbf{k}, v^2, \mu) + (B-2A)v^2(1-2v^2) + 2v^2(1-v^2)\Delta(k)](\alpha_p^+\alpha_p + \beta_p^+\beta_p) + \\ + \sum_p [uv(\frac{k_x}{2} - p)\beta_{k_x-p}\alpha_p + uv(p - \frac{k_x}{2})\alpha_p^+\beta_{k_x-p}^+]\left\{-\psi(\mathbf{k}, v^2, \mu) + 2v^2(B-2A+\Delta(k)) - \Delta(k)\right\}. \quad (43)$$

Here the notations of [20] were used

$$E(\mathbf{k}, v^2, \mu) = 2v^2u^2 I_{ex}(k) + I_l(v^4 - v^2u^2) - \frac{\mu}{2}(u^2 - v^2), \\ \psi(\mathbf{k}, v^2, \mu) = 2v^2 I_l + I_{ex}(k)(1-2v^2) + \mu. \quad (44)$$

The coefficients $\Delta(k)$, A_{i-i} and B_{i-i} were deduced in [22]. They are

$$A_{i-i} = \frac{I_l^2}{\pi\hbar\omega_c} S; \quad S \approx 0.481;$$

$$B_{i-i} = \frac{2I_l^2}{\pi\hbar\omega_c} T; \quad T \approx 0.216;$$

$$\Delta(0) = \frac{2I_l^2}{\pi\hbar\omega_c} 0.344.$$

Putting to zero the last bracket in equation (43), i.e., compensating the dangerous diagrams describing the spontaneous creation and annihilation of quasielectron-quasihole pairs in the new vacuum state (32), we determine the chemical potential μ of the system in the HFB approximation

$$\mu^{HFB} = -\tilde{I}_{ex}(k) + 2v^2(B - 2A + \tilde{I}_{ex}(k) - I_l) = -\tilde{I}_{ex}(k) + 2v^2(B - 2A + \Delta(k) - E(k)). \quad (45)$$

Here the renormalized ionization potential of magnetoexcitons $\tilde{I}_{ex}(k)$ containing the correction due to influence of all ELLs was introduced

$$\tilde{I}_{ex}(k) = I_{ex}(k) + \Delta(k); \qquad I_{ex}(k) = I_l - E(k); \qquad E_{ex}(k) = -I_{ex}(k). \quad (46)$$

Upon introduction of the value μ^{HFB} in the remainder part of the first line of (43), the Hamiltonian H_2 will take the form

$$H_2 = \sum_p \frac{\tilde{I}_{ex}(k)}{2}(\alpha_p^+\alpha_p + \beta_p^+\beta_p). \quad (47)$$

This Hamiltonian describes the single-particle elementary excitation extracting it from a single-exciton state with wave vector k of the condensed magnetoexcitons. To extract from the condensate one pair of new quasiparticles, the energy cost $\tilde{I}_{ex}(k)$ is equivalent to unbinding energy. For this reason, the excitation energy for one quasiparticle is $\tilde{I}_{ex}(k)/2$. Notice that the chemical potential μ^{HFB} in the point $v^2 = 0$ coincides with the position of the renormalized magnetoexciton energy band on the energy scale $\tilde{E}_{ex}(k) = -\tilde{I}_{ex}(k)$, while in the point $v^2 = 1$ it amounts to the value $-I_l + B - 2A$ and does not depend on k. The concentration corrections to

μ^{HFB} are determined by the term $2v^2(B-2A+\Delta(k)-E(k))$. The term $-E(k)$ appears within the frame of the LLLs and was obtained in [20, 21]. It determines the instability of the ground state within the HFBA when the corrections due to ELL are neglected. The term $B-2A$ appears in both phases, not only in the case of BEC of magnetoexciton but also in the case of EHL. The term $-2A$ is related with the average Hartree terms of the supplementary e–e, h–h, and e–h interactions; the term B, with the average exchange terms of the supplementary e–e and h–h interactions. The term $2v^2\Delta(k)$ is related to e–h interaction and Bogoliubov u-v transformation and is named the Bogoliubov self-energy term.

Below, we shall construct the equations of motion for the operators of creation $d^{\dagger}(P)$ and annihilation $d(P)$ of magnetoexciton and density fluctuation operators for electrons $\rho(Q)$ and holes $D(Q)$ on the base of Hamiltonian (1) in the quasiaverages theory approximation (QATA).

4. Equations of motion for the two-particle operators and for the corresponding Green's functions

The starting Hamiltonian in QATA has the form

$$\hat{\mathcal{H}} = \frac{1}{2}\sum_{\vec{Q}} W_{\vec{Q}}\left[\rho(\vec{Q})\rho(-\vec{Q}) - \hat{N}_e - \hat{N}_h\right] - \mu_e\hat{N}_e - \mu_h\hat{N}_h -$$

$$-\tilde{\eta}\sqrt{N}\left(e^{i\varphi}d^{\dagger}(k) + e^{-i\varphi}d(k)\right) + \frac{1}{2}B_{i-i}\widehat{N} - \tag{48}$$

$$-\frac{1}{4N}\sum_{Q} V(Q)\hat{\rho}(\vec{Q})\hat{\rho}(-\vec{Q}) - \frac{1}{4N}\sum_{Q} U(Q)\hat{D}(\vec{Q})\hat{D}(-\vec{Q}).$$

The density fluctuation operators (24) with different wave vectors P and Q do not commute, which is related with the helicity or spirality accompanying the presence of a strong magnetic field [18]. They are expressed by the phase factors in the structure of operators (6) and by the vector-product of two 2D wave vectors P and Q and its projection on the direction of the magnetic field. These properties considerably influence the structure of the equations of motion for the operators and determine new aspects of the 2D electron-hole (e-h) physics.

The equation of motion for the creation and annihilation operators $d^{+}(k)$, $d(k)$ (28) and for the density fluctuation operators (24) are

$$i\hbar\frac{d}{dt}d(\vec{P}) = [d(\vec{P}), \hat{\mathcal{H}}] = (-\bar{\mu} + E(\vec{P}) - \Delta(\vec{P}))d(\vec{P}) - 2i\sum_{\vec{Q}}\tilde{W}(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)\hat{\rho}(\vec{Q})d(\vec{P}-\vec{Q}) -$$

$$-\frac{1}{N}\sum_{\vec{Q}} U(\vec{Q})Cos\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)D(\vec{Q})d(\vec{P}-\vec{Q}) - \tilde{\eta}\sqrt{N}e^{i\varphi}\delta_{kr}(\vec{P},0) + \tilde{\eta}e^{i\varphi}\frac{D(\vec{P})}{\sqrt{N}};$$

$$i\hbar\frac{d}{dt}d^{\dagger}(-\vec{P}) = [d^{\dagger}(-\vec{P}), \hat{\mathcal{H}}] = (\bar{\mu} - E(-\vec{P}) + \Delta(-\vec{P}))d^{\dagger}(-\vec{P}) + \tag{49}$$

$$+2i\sum_{\vec{Q}}\tilde{W}(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)d^{\dagger}(-\vec{P}-\vec{Q})\hat{\rho}(-\vec{Q}) +$$

$$+\frac{1}{N}\sum_{\vec{Q}} U(\vec{Q})Cos\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)d^{\dagger}(-\vec{P}-\vec{Q})D(-\vec{Q}) + \tilde{\eta}\sqrt{N}e^{-i\varphi}\delta_{kr}(\vec{P},0) - \tilde{\eta}e^{-i\varphi}\frac{D(\vec{P})}{\sqrt{N}};$$

$$i\hbar\frac{d}{dt}\hat{\rho}(\vec{P})=[\hat{\rho}(\vec{P}),\hat{\mathcal{H}}]=$$

$$=-i\sum_{\vec{Q}}\tilde{W}(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)[\hat{\rho}(\vec{P}-\vec{Q})\hat{\rho}(\vec{Q})+\hat{\rho}(\vec{Q})\hat{\rho}(\vec{P}-\vec{Q})]+$$

$$+\frac{i}{2N}\sum_{\vec{Q}}U(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)\Big[D(\vec{P}-\vec{Q})D(\vec{Q})+D(\vec{Q})D(\vec{P}-\vec{Q})\Big];$$

$$i\hbar\frac{d}{dt}\hat{D}(\vec{P})=[\hat{D}(\vec{P}),\hat{\mathcal{H}}]=$$

$$-i\sum_{\vec{Q}}\tilde{W}(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)[\hat{\rho}(\vec{Q})\hat{D}(\vec{P}-\vec{Q})+\hat{D}(\vec{P}-\vec{Q})\hat{\rho}(\vec{Q})]+$$

$$+\frac{i}{2N}\sum_{\vec{Q}}U(\vec{Q})Sin\left(\frac{[\vec{P}\times\vec{Q}]_z l^2}{2}\right)[\hat{D}(\vec{Q})\hat{\rho}(\vec{P}-\vec{Q})+\hat{\rho}(\vec{P}-\vec{Q})\hat{D}(\vec{Q})]+$$

$$+2\tilde{\eta}\sqrt{N}\Big[e^{-i\varphi}d(\vec{P})-e^{i\varphi}d^{\dagger}(-\vec{P})\Big].$$

Here

$$\tilde{\eta}=(\tilde{E}_{ex}(k)-\mu)\mathrm{v}=(E(k)-\Delta(k)-\bar{\mu})\mathrm{v};\quad \tilde{E}_{ex}(k)=E_{ex}(k)-\Delta(k)=-I_l-\Delta(k)+E(k);$$

$$E_{ex}(k)=-I_l+E(k);\quad E(K)=2\sum_{Q}W_Q Sin^2\left(\frac{[K\times Q]_z l^2}{2}\right);$$

$$\bar{\mu}=\mu+I_l;\qquad v=\mathrm{v}^2;\qquad N_{ex}=\mathrm{v}^2N;\qquad \widetilde{W}(Q)=W_Q-\frac{1}{2N}V(Q);$$

$$\Delta(k)=\sum_{s}\phi_{e-h}(p,-p-k_x,s)e^{-k_y s l^2}=\frac{1}{N}\sum_{Q}\psi_{e-h}(Q)\exp(i[k\times Q]_z l^2).$$

Following equations of motion (49), we will introduce four interconnected retarded Green's functions at $T=0$ [35, 36]

$$\begin{aligned}
G_{11}(\vec{P},t)&=\Big\langle\Big\langle d(\vec{P},t);\hat{X}^{\dagger}(\vec{P},0)\Big\rangle\Big\rangle;\\
G_{12}(\vec{P},t)&=\Big\langle\Big\langle d^{\dagger}(-\vec{P},t);\hat{X}^{\dagger}(\vec{P},0)\Big\rangle\Big\rangle;\\
G_{13}(\vec{P},t)&=\left\langle\left\langle\frac{\hat{\rho}(\vec{P},t)}{\sqrt{N}};\hat{X}^{\dagger}(\vec{P},0)\right\rangle\right\rangle;\\
G_{14}(\vec{P},t)&=\left\langle\left\langle\frac{\hat{D}(\vec{P},t)}{\sqrt{N}};\hat{X}^{\dagger}(\vec{P},0)\right\rangle\right\rangle.
\end{aligned}\tag{50}$$

They are determined by the relations

$$\begin{aligned}
G(t)&=\Big\langle\Big\langle\hat{A}(t);\hat{B}(0)\Big\rangle\Big\rangle=-i\theta(t)\Big\langle\big[A(t),B(0)\big]\Big\rangle;\\
\hat{A}(t)&=e^{\frac{iHt}{\hbar}}\hat{A}e^{-\frac{iHt}{\hbar}};\\
\Big[\hat{A},\hat{B}\Big]&=\hat{A}\hat{B}-\hat{B}\hat{A},
\end{aligned}\tag{51}$$

where $\hat{H}$ is Hamiltonian (48).

The average $\langle\ \rangle$ will be calculated at $T=0$ in the HFB approximation using the ground state wave function $\left|\psi_g(k)\right\rangle$ (32). The time derivative of the Green's function is calculated as follows

$$\begin{aligned} i\hbar\frac{d}{dt}G(t) &= i\hbar\frac{d}{dt}\left\langle\left\langle A(t);B(0)\right\rangle\right\rangle = \\ &= \hbar\delta(t)\left\langle\left[\hat{A}(0),\hat{B}(0)\right]\right\rangle + \left\langle\left\langle i\hbar\frac{d}{dt}A(t);B(0)\right\rangle\right\rangle = \\ &= \hbar\delta(t)C + \left\langle\left\langle\left[\hat{A}(t),\hat{H}\right];\hat{B}(0)\right\rangle\right\rangle. \end{aligned} \tag{52}$$

C will stand for the average values, which do not depend on *t*. They are not needed in an explicit form for the determination of the energy spectrum of the elementary excitations.

The Fourier transforms of Green's functions (50) will be denoted as

$$\begin{aligned} G_{11}(\vec{P},\omega) &= \left\langle\left\langle d(\vec{P})\middle|\hat{X}^{\dagger}(\vec{P})\right\rangle\right\rangle_{\omega}; \\ G_{12}(\vec{P},\omega) &= \left\langle\left\langle d^{\dagger}(-\vec{P})\middle|\hat{X}^{\dagger}(\vec{P})\right\rangle\right\rangle_{\omega}; \\ G_{13}(\vec{P},\omega) &= \left\langle\left\langle \frac{\hat{\rho}(\vec{P})}{\sqrt{N}}\middle|\hat{X}^{\dagger}(\vec{P})\right\rangle\right\rangle_{\omega}; \\ G_{14}(\vec{P},\omega) &= \left\langle\left\langle \frac{\hat{D}(\vec{P})}{\sqrt{N}}\middle|\hat{X}^{\dagger}(\vec{P})\right\rangle\right\rangle_{\omega}. \end{aligned} \tag{53}$$

The two representations are related as follows

$$G(\vec{P},\omega) = \int_{-\infty}^{\infty} e^{i\omega t}G(\vec{P},t)dt = \int_{0}^{\infty} e^{i\omega t-\delta t}G(\vec{P},t)dt,$$

where the infinitesimal value $\delta\to+0$ guarantees the convergence of the integral in the interval $(0,\infty)$ for the retarded Green's function $G(\vec{P},t)$.

The equation of motion in the frequency representation can be deduced on the base of equation (52)

$$\begin{aligned} &\int_{-\infty}^{\infty} dte^{i\omega t}i\hbar\frac{dG(t)}{dt} = i\hbar\int_{0}^{\infty} dte^{i\omega t-\delta t}\frac{dG(t)}{dt} = -i\hbar\int_{0}^{\infty} dtG(t)\frac{de^{i\omega t-\delta t}}{dt} = \\ &= (\hbar\omega+i\delta)G(\omega) = C + \int_{-\infty}^{\infty} dt\left\langle\left\langle\left[\hat{A}(t),\hat{H}\right];\hat{B}(0)\right\rangle\right\rangle e^{i\omega t} \end{aligned} \tag{54}$$

Green's functions (53) will be named one-operator Green's functions because, in the left hand side of the vertical line, they contain only one summary operator of the types $d(P)$, $d^{\dagger}(P)$, $\hat{\rho}(P)$ and $\hat{D}(P)$. At the same time, these Green's functions are two-particle Green's functions, because the summary operators are expressed through the products of two Fermi operators. In this respect, Green's functions (53) are equivalent to the two-particle Green's functions introduced by Keldysh and Kozlov in their fundamental paper [27], forming the base of the theory of high density excitons in the electron-hole description. However, in contrast to [27], we are using the summary operators which represent integrals on the wave vectors of relative motions.

The equations of motion for the Green's functions are

$$(\hbar\omega + i\delta + \bar{\mu} - E(P) + \Delta(P))G_1(P,\omega) =$$

$$= C - 2i\sum_Q \tilde{W}(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle \rho(Q)d(P-Q)\middle| X\right\rangle\right\rangle_\omega -$$

$$-\frac{1}{N}\sum_Q U(Q)Cos\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle D(Q)d(P-Q)\middle| X\right\rangle\right\rangle_\omega + \tilde{\eta}G_4(P,\omega)e^{i\varphi};$$

$$(\hbar\omega + i\delta - \bar{\mu} + E(-P) - \Delta(-P))G_2(P,\omega) =$$

$$= C + 2i\sum_Q \tilde{W}(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle d^\dagger(-P-Q)\rho(-Q)\middle| X\right\rangle\right\rangle_\omega +$$

$$+\frac{1}{N}\left\langle\left\langle d^\dagger(-P-Q)D(-Q)\middle| X\right\rangle\right\rangle_\omega - \tilde{\eta}G_4(P,\omega)e^{-i\varphi};$$

$$(\hbar\omega + i\delta)G_3(P,\omega) =$$

$$= C - i\sum_Q \tilde{W}(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle \frac{\rho(P-Q)\rho(Q)}{\sqrt{N}} + \frac{\rho(Q)\rho(P-Q)}{\sqrt{N}}\middle| X\right\rangle\right\rangle_\omega +$$

$$+\frac{i}{2N}\sum_Q U(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle \frac{D(P-Q)D(Q)}{\sqrt{N}} + \frac{D(Q)D(P-Q)}{\sqrt{N}}\middle| X\right\rangle\right\rangle_\omega ;$$

$$(\hbar\omega + i\delta)G_4(P,\omega) =$$

$$= C - i\sum_Q \tilde{W}(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle \frac{D(Q)\rho(P-Q)}{\sqrt{N}} + \frac{D(P-Q)\rho(Q)}{\sqrt{N}}\middle| X\right\rangle\right\rangle_\omega + \qquad (55)$$

$$+\frac{i}{2N}\sum_Q U(Q)Sin\left(\frac{[P\times Q]_z l^2}{2}\right)\left\langle\left\langle \frac{D(Q)\rho(P-Q)}{\sqrt{N}} + \frac{\rho(P-Q)D(Q)}{\sqrt{N}}\middle| X\right\rangle\right\rangle_\omega +$$

$$+2\tilde{\eta}\left[e^{-i\varphi}G_1(P,\omega) - e^{i\varphi}G_2(P,\omega)\right].$$

5. Dyson equation and self-energy parts

Using the Zubarev's procedure [36] for the Green's function, we obtain a closed system of Dyson equation for the Green's functions in the form

$$\sum_{j=1}^{4} G_{1j}(\vec{P},\omega)\Sigma_{jk}(\vec{P},\omega) = C_{1k}; \quad k = 1,2,3,4. \qquad (56)$$

There are 16 different components of the self energy part of the 4×4 matrix $\Sigma_{jk}(\vec{P},\omega)$ as follows

$$\Sigma_{11}(\vec{P},\omega) = \hbar\omega + i\delta + \bar{\mu} - E(\vec{P}) + \Delta(\vec{P}) -$$

$$-\frac{\langle D(A)D(-A)\rangle}{N^2}\sum_{\vec{Q}\neq\vec{P}}\frac{U^2(\vec{Q})Cos^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega + i\delta + \bar{\mu} - E(\vec{P}-\vec{Q}) + \Delta(\vec{P}-\vec{Q})};$$

$$\Sigma_{21}(\vec{P},\omega) = 0;$$

$$\Sigma_{31}(\vec{P},\omega) = i\frac{\left\langle D(A)d(-A)\sqrt{N}\right\rangle}{N^2}\sum_{\vec{Q}\neq\vec{P}}\frac{U(\vec{Q})U(\vec{Q}-\vec{P})Cos\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)Sin\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega + i\delta + \bar{\mu} - E(\vec{P}-\vec{Q}) + \Delta(\vec{P}-\vec{Q})};$$

$$\Sigma_{41}(\vec{P},\omega) = -\tilde{\eta}e^{i\varphi} + U(\vec{P})\frac{\langle d(0)\rangle}{\sqrt{N}} +$$

$$+2\frac{\langle D(A)d(-A)\sqrt{N}\rangle}{N}\sum_{\vec{Q}\neq\vec{P}}\frac{\tilde{W}(\vec{Q})(U(\vec{P})-U(\vec{Q}-\vec{P}))Sin^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta+\bar{\mu}-E(\vec{P}-\vec{Q})+\Delta(\vec{P}-\vec{Q})} -$$

$$-\frac{\langle D(A)d(-A)\sqrt{N}\rangle}{N^2}\sum_{\vec{Q}\neq\vec{P}}\frac{U(\vec{Q})U(\vec{P})Cos^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta+\bar{\mu}-E(\vec{P}-\vec{Q})+\Delta(\vec{P}-\vec{Q})};$$

$$\Sigma_{12}(\vec{P},\omega) = 0;$$

$$\Sigma_{22}(\vec{P},\omega) = \hbar\omega+i\delta-\bar{\mu}+E(\vec{P})-\Delta(-\vec{P}) -$$

$$-\frac{\langle D(A)D(-A)\rangle}{N^2}\sum_{\vec{Q}\neq-\vec{P}}\frac{U(\vec{Q})U(-\vec{Q})Cos^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta-\bar{\mu}+E(-\vec{P}-\vec{Q})-\Delta(-\vec{P}-\vec{Q})};$$

$$\Sigma_{32}(\vec{P},\omega) = i\frac{\langle d^{\dagger}(A)D(-A)\sqrt{N}\rangle}{N^2}\sum_{\vec{Q}\neq-\vec{P}}\frac{U(\vec{Q})U(-\vec{Q}-\vec{P})Cos\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)Sin\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta-\bar{\mu}+E(-\vec{P}-\vec{Q})-\Delta(-\vec{P}-\vec{Q})};$$

$$\Sigma_{42}(\vec{P},\omega) = \tilde{\eta}e^{-i\varphi} - U(-\vec{P})\frac{\langle d^{\dagger}(0)\rangle}{\sqrt{N}} -$$

$$-2\frac{\langle d^{\dagger}(A)D(-A)\sqrt{N}\rangle}{N}\sum_{\vec{Q}\neq-\vec{P}}\frac{\tilde{W}(\vec{Q})(U(-\vec{Q}-\vec{P})-U(\vec{P}))Sin^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta-\bar{\mu}+E(-\vec{P}-\vec{Q})-\Delta(-\vec{P}-\vec{Q})} -$$

$$-\frac{\langle d^{\dagger}(A)D(-A)\sqrt{N}\rangle}{N^2}\sum_{\vec{Q}\neq-\vec{P}}\frac{U(\vec{Q})U(-\vec{P})Cos^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)}{\hbar\omega+i\delta-\bar{\mu}+E(-\vec{P}-\vec{Q})-\Delta(-\vec{P}-\vec{Q})};$$

$$\Sigma_{13}(\vec{P},\omega) = 0;\ \ \Sigma_{23}(\vec{P},\omega) = 0;$$

$$\Sigma_{33}(\vec{P},\omega) = \hbar\omega+i\delta-\frac{\langle D(A)D(-A)\rangle}{N^2(\hbar\omega+i\delta)}\sum_{\vec{Q}\neq\vec{P}}U(\vec{Q})(U(-\vec{Q})-U(\vec{Q}-\vec{P}))Sin^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right);$$

$$\Sigma_{43}(\vec{P},\omega) = 0;$$

$$\Sigma_{14}(\vec{P},\omega) = -2\tilde{\eta}e^{-i\varphi};$$

$$\Sigma_{24}(\vec{P},\omega) = 2\tilde{\eta}e^{i\varphi};$$

$$\Sigma_{34}(\vec{P},\omega) = 0;$$

$$\Sigma_{44}(\vec{P},\omega) = \hbar\omega+i\delta-\frac{2\langle D(A)D(-A)\rangle}{N(\hbar\omega+i\delta)}\sum_{\vec{Q}\neq-\vec{P}}\tilde{W}(\vec{Q})(U(\vec{Q}-\vec{P})-U(\vec{P}))Sin^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right)+$$

$$+\frac{\langle D(A)D(-A)\rangle}{N^2(\hbar\omega+i\delta)}\sum_{\vec{Q}\neq\vec{P}}U(\vec{Q})(U(\vec{P})-U(-\vec{Q}))Sin^2\left(\frac{[\vec{P}\times\vec{Q}]l^2}{2}\right). \qquad (57)$$

The self-energy parts $\Sigma_{jk}(\vec{P},\omega)$ represented by formulas (57) contain the different average values of the two-operator products. They were calculated using the ground state wave function $\left|\psi_g(0)\right\rangle$ (32) taken with $k=0$ and have the expressions

$$\begin{gathered}\left\langle D(\vec{Q})D(-\vec{Q})\right\rangle = 4u^2v^2N;\\ \overline{\mu} = -\Delta(0) + 2v^2(B_{i-i} - 2A_{i-i} + \Delta(0));\\ \left\langle D(\vec{Q})d(-\vec{Q})\sqrt{N}\right\rangle = \left\langle d^{\dagger}(\vec{Q})D(-\vec{Q})\sqrt{N}\right\rangle = -2uv^3N;\\ \left\langle d(0)\right\rangle = \left\langle d^{\dagger}(0)\right\rangle = uv\sqrt{N}; \qquad \tilde{\eta} = -(\Delta(0)+\overline{\mu})v.\end{gathered} \tag{58}$$

All these averages are extensive values proportional to N or $\sqrt{N}$, they essentially depend on the small parameters of the types $u^2\mathrm{v}^2$ or $u\mathrm{v}^3$, or uv.

The cumbersome dispersion equation is expressed in a general form by the determinant equation

$$\det\left|\Sigma_{ij}(\vec{P},\omega)\right| = 0. \tag{59}$$

We can substitute the self-energy parts $\Sigma_{jk}(\vec{P},\omega)$ (57) in formula (59), and the determinant equation (59) disintegrates into two independent equations. One of them concerns only optical plasmons and has the simple form

$$\Sigma_{33}(\vec{P};\omega) = 0, \tag{60}$$

whereas the second equation contains the self-energy parts Σ_{11}, Σ_{22}, Σ_{44}, Σ_{14}, Σ_{41}, Σ_{24}, Σ_{42} and the quasi-average constant $\tilde{\eta}$

$$\Sigma_{11}(\vec{P};\omega)\Sigma_{22}(\vec{P};\omega)\Sigma_{44}(\vec{P};\omega) - \Sigma_{41}(\vec{P};\omega)\Sigma_{22}(\vec{P};\omega)\Sigma_{14}(\vec{P};\omega) - \Sigma_{42}(\vec{P};\omega)\Sigma_{11}(\vec{P};\omega)\Sigma_{24}(\vec{P};\omega) = 0. \tag{61}$$

The solutions of dispersion equation (61) will be discussed in two limiting cases. One of them is the point $v^2 = 0$, where the system behaves as an ideal Bose gas and the other case is $v^2 \neq 0$.

All contributions to the self-energy parts contain the averages $\left\langle D(\vec{Q})D(-\vec{Q})\right\rangle$, $\left\langle D(\vec{Q})d(-\vec{Q})\sqrt{N}\right\rangle$, $\left\langle d(0)\right\rangle$, which do not vanish in the point $\vec{k}=0$. The 2D magnetoexciton system now is not at all a pure ideal gas. It was an ideal gas when the influence of ELLs was neglected. This unusual result was revealed for the first time by Lerner and Lozovik [15-17] and was confirmed by Paquet, Rice and Ueda [19]. In the case $v^2 = 0$, due to the vanishing of averages (58), the self-energy parts become

$$\begin{aligned}&\sigma_{11}(\vec{P},\omega) = \hbar\omega - E(P); && \overline{\mu} + \Delta(0) = 0;\\ &\sigma_{22}(\vec{P},\omega) = \hbar\omega + E(-P); && \tilde{\eta} = 0;\\ &\sigma_{33}(\vec{P},\omega) = \hbar\omega; && \Delta(\vec{P}) \approx \Delta(0);\\ &\sigma_{44}(\vec{P},\omega) = \hbar\omega; && v^2 = 0;\ \vec{k} = 0,\end{aligned} \tag{62}$$

the excitonic part of the dispersion relation and the acoustic plasmon frequency look like

$$\begin{gathered}\hbar\omega_{ex}(P) = \pm E(P);\\ \hbar\omega_A(P) = \hbar\omega_O(P) = 0.\end{gathered} \tag{63}$$

The acoustical and optical plasmon branches have the frequencies equal to zero. This case is presented in Fig. 1.

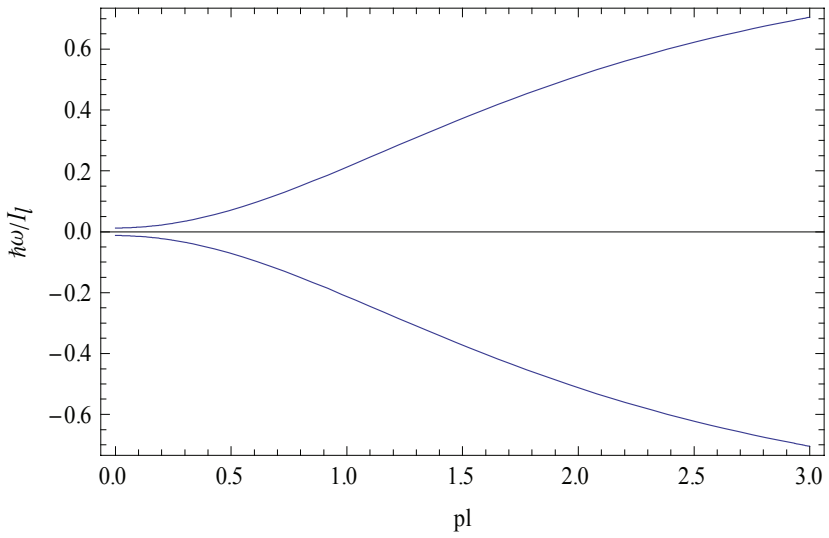


Fig. 1. The energy spectrum of elementary excitations of magnetoexcitons in the case when concentration corrections are not taken into account, the filling factor equals zero.

If we keep terms proportional to uv in formulas (58), then the self-energy parts included in (56) can be rewritten in following form

$$\begin{gathered}
\sigma_{11}(\vec{P},\omega)=\hbar\omega+\bar{\mu}-E(\vec{P})+\Delta(\vec{P});\\
\sigma_{41}(\vec{P},\omega)=-\tilde{\eta}+U(\vec{P})\frac{\langle d(0)\rangle}{\sqrt{N}};\\
\sigma_{22}(\vec{P},\omega)=\hbar\omega-\bar{\mu}+E(\vec{P})-\Delta(-\vec{P});\\
\sigma_{42}(\vec{P},\omega)=\tilde{\eta}-U(-\vec{P})\frac{\langle d^{\dagger}(0)\rangle}{\sqrt{N}};\\
\sigma_{14}(\vec{P},\omega)=-2\tilde{\eta};\\
\sigma_{24}(\vec{P},\omega)=2\tilde{\eta};\\
\sigma_{44}(\vec{P},\omega)=\hbar\omega+i\delta;\\
\bar{\mu}=-\Delta(0)+2v^2(B_{i-i}-2A_{i-i}+\Delta(0));\\
\tilde{\eta}=-(\Delta(0)+\bar{\mu})v=-2v^3(B_{i-i}-2A_{i-i}+\Delta(0));\\
\Delta(\vec{P})\approx\Delta(0);\quad \vec{k}=0;\quad v^2\neq 0;\quad U(\vec{P})\cong U(0)e^{-\frac{P^2l^2}{2}},\quad U(0)=2A_{i-i}.
\end{gathered}\tag{64}$$

Dispersion equation (61) in this case looks like

$$\hbar\omega=\pm\sqrt{\left(\bar{\mu}-E(\vec{P})+\Delta(0)\right)^2+4\tilde{\eta}\left(\tilde{\eta}-\frac{U(\vec{P})\langle d(0)\rangle}{\sqrt{N}}\right)}.\tag{65}$$

In the Ref. [22] the coefficient $(B_{i-i}-2A_{i-i}+\Delta(0))/I_l$ was determined to be 0.025 at the ratio $r=I_l/\hbar\omega=1/2$. It was used in the present calculations leading to the main parameters $\bar{\mu}$ and $\tilde{\eta}$ $(\bar{\mu}+\Delta(0))/I_l=2v^2\bullet 0.025=0.05v^2$, $\tilde{\eta}=-(\bar{\mu}+\Delta(0))v/I_l=-2v^3\bullet 0.025=-0.05v^3$, and $U(0)\langle d(0)\rangle/I_l\sqrt{N}=2A_{i-i}uv/I_l=0.15uv$. Introducing the dimensionless energies $\hbar\tilde{\omega}=\hbar\omega/I_l$ and $\tilde{E}(P)=E(P)/I_l$ one can transcribe solution (65) as follows

$$\hbar\omega = \pm\sqrt{\left(0.05v^2 - \tilde{E}(P)\right)^2 + 0.2v^2\left(0.05v^2 + 0.15uve^{-\frac{P^2l^2}{2}}\right)}. \qquad (66)$$

Their solutions are presented in Fig. 2. The excitonic branch of elementary excitations is characterized by a roton-type behavior at the small and intermediary values of the wavevectors and by a monotonic increasing at higher values of the wavevectors. The acoustical and optical plasmon branches have frequencies equal to zero because the sums in expressions (57) containing the coefficients $U^2(\vec{Q})$, as well as $\tilde{W}(\vec{Q})U(\vec{Q})$, were not included in these calculations.

The neglected terms in the expressions for $\Sigma_{33}(\vec{P},\omega)$ and $\Sigma_{44}(\vec{P},\omega)$ can be calculated using the approximation (27) $U(\vec{Q}) \cong U(0)\exp[-Q^2l^2/2]$, whereas the terms proportional to $U^2(\vec{Q})$ in the expressions $\Sigma_{11}(\vec{P},\omega)$ and $\Sigma_{22}(\vec{P},\omega)$ must be summarized together with the denominators of the types $\hbar\omega + i\delta \pm \bar{\mu} \mp E(\pm\vec{P} - \vec{Q}) \pm \Delta(\pm\vec{P} - \vec{Q})$. They were represented as

$$\frac{\hbar\omega + i\delta \pm \bar{\mu} \mp E(\pm\vec{P} - \vec{Q}) \pm \Delta(\pm\vec{P} - \vec{Q}) - i\gamma}{(\hbar\omega + i\delta \pm \bar{\mu} \mp E(\pm\vec{P} - \vec{Q}) \pm \Delta(\pm\vec{P} - \vec{Q}))^2 + \gamma^2},$$

and the approach was proposed

$$\frac{Pf}{\hbar\omega \pm \bar{\mu} \mp E(\pm\vec{P} - \vec{Q}) \pm \Delta(\pm\vec{P} - \vec{Q})} \approx \frac{\hbar\omega \pm \bar{\mu} \mp E(\pm\vec{P}) \pm \Delta(\pm\vec{P})}{\gamma^2}.$$

In this rude approximation for denominators, using the known expressions for $W_{\vec{Q}}$ and $U(\vec{Q})$, the sums on $\vec{Q}$ were calculated; the results are represented in Figs. 3-5.

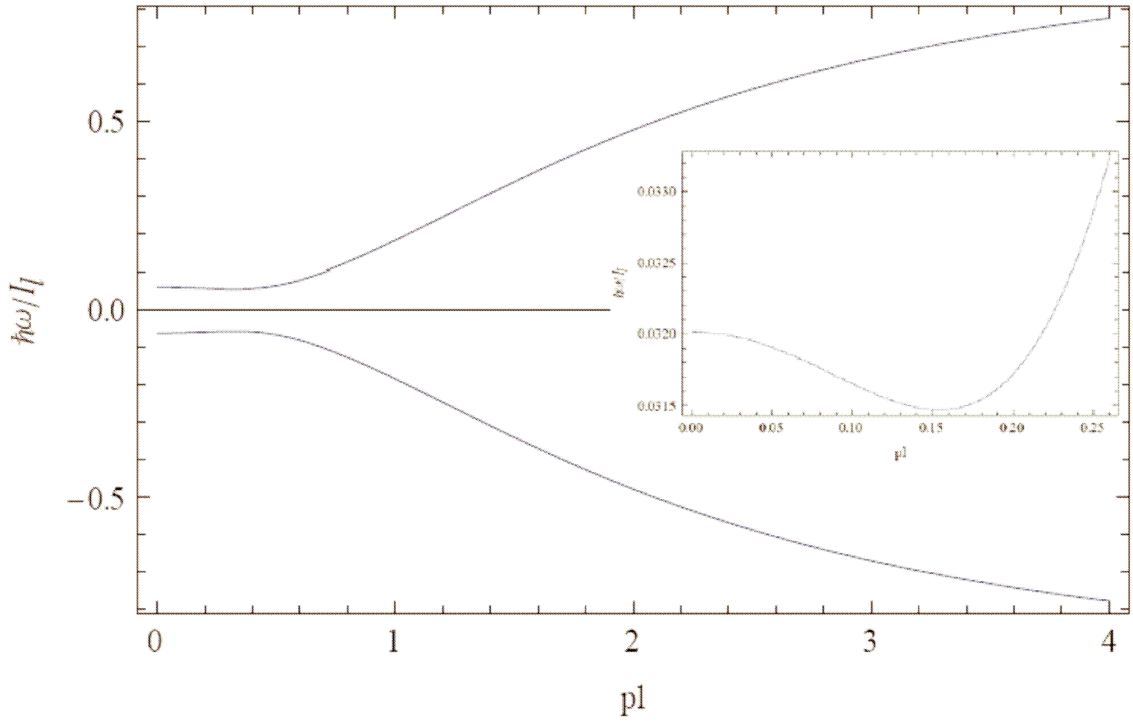


Fig. 2. Two exciton branches of the energy spectrum of collective elementary excitations of the Bose-Einstein condensed magnetoexcitons on the wave vector $\vec{k} = 0$ calculated in HFBA using self-energy parts (64) and the filling factor $v^2 = 0.1$.

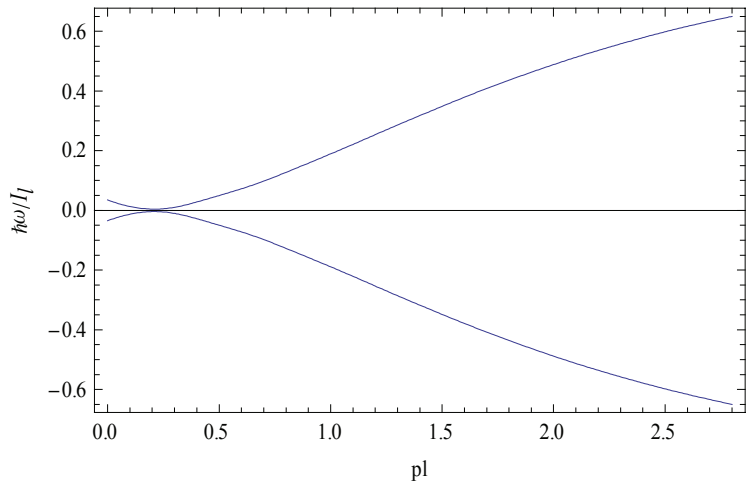


Fig. 3. Two exciton branches of the energy spectrum of collective elementary excitations of the Bose-Einstein condensed magnetoexcitons on the wave vector $\vec{k} = 0$ calculated in HFBA using self-energy parts (57) and the filling factor $\nu^2 = 0.1$.

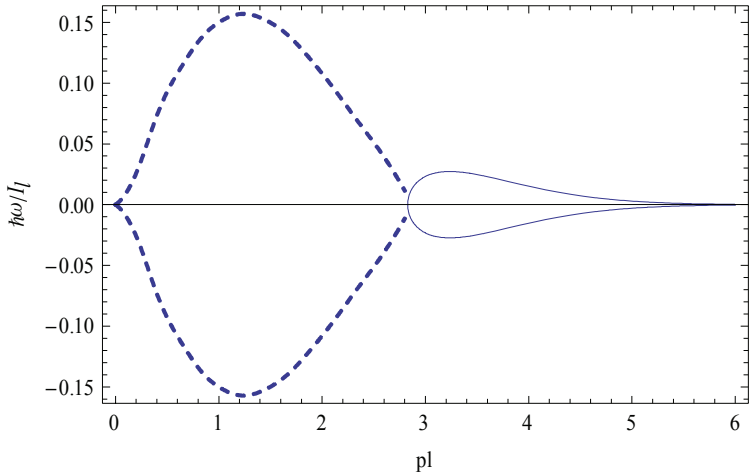


Fig. 4. The dispersion law of acoustical plasmon branch in the presence of the BEC of magnetoexcitons on the wave vector $\vec{k} = 0$ calculated in HFBA using self-energy parts (57) and filling factor $\nu^2 = 0.1$.

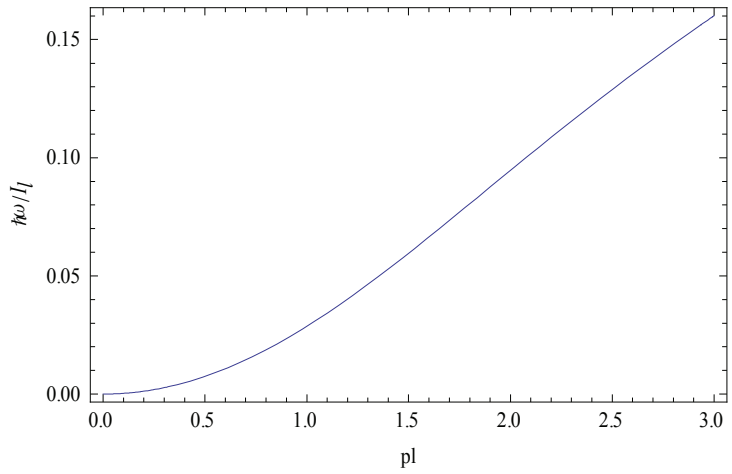


Fig. 5. The dispersion law of optical plasmon branch in the presence of the BEC of magnetoexcitons on the wave vector $\vec{k} = 0$ calculated in HFBA using self-energy parts (57) and the filling factor $\nu^2 = 0.1$.

Conclusions

The energy spectrum of the collective elementary excitations of a 2D e-h system situated in a strong perpendicular magnetic field in a state of BEC with wave vector $\vec{k} = 0$ was investigated within the frame of Bogoliubov theory of quasiaverages. The starting Hamiltonian describing the e-h system contains not only the Coulomb interaction between the particles lying on the LLLs but also the supplementary interaction due to their virtual quantum transitions from the LLLs to the ELLs and return back. This supplementary interaction generates after the averaging on the ground state BCS-type wave function the direct Hartree-type terms with attractive character, the exchange Fock-type terms giving rise to repulsion as well as the similar terms arising after the Bogoliubov $u - v$ transformation. The interplay of these three parameters gives rise to the resulting nonzero interaction between the magnetoexcitons with wave vector $\vec{k} = 0$ and to stability of their BEC as regards the collapse. It influences also the energy spectrum as well as the collective elementary excitations. It consists of four branches. Two of them are excitonic-type branches, one of them being the usual energy branch whereas the other is the quasienergy branch representing the mirror reflection of the energy branch. The other two branches are the optical and acoustical plasmon branches. The exciton energy branch has an energy gap due to the attractive interaction terms, which is needed to be got over during the excitation as well as a roton-type section in the range of intermediary values of the wave vectors. At higher values of wave vector, its dispersion law tends to saturation. The optical plasmon dispersion law is gapless with quadratic dependence in the range of small wave vectors and with saturation-type dependence in the remaining part of the spectrum. The acoustical plasmon branch reveals the absolute instability of the spectrum in the range of small and intermediary values of the wave vectors. In the remaining range of the wave vectors, the acoustical plasmon branch exhibits a very small real value of the energy spectrum tending to zero in the limiting case of high wave vectors.

References

[1] M.A. Liberman and B. Johanson, Uspekhi Fiz. Nauk, 165, 121, (1995).
[2] Dong Lai, Rev. Mod. Phys., 73, 629, (2001).
[3] H.L. Stormer, Rev. Mod. Phys., 71, 875, (1999).
[4] S. Das Sarma, and A. Pinczuk, Ed. by, Perspectives in Quantum Hall Effects, John Wiley & Sons, Inc., New York, 430 p., 1997.
[5] E.I. Rashba, Pure and Applied Chem., 67, 409, (1995).
[6] D.C. Tsui, H.L. Stormer, and A.C. Gossard, Phys. Rev. Lett., 48, 1559, (1982).
[7] R.B. Laughlin, Phys. Rev. Lett., 50, 13, (1983).
[8] A.H. MacDonald, E.H. Rezayi, and D. Keller, Phys. Rev. Lett., 68, 1939, (1992).
[9] F.D.M. Haldane, Phys. Rev. Lett., 51, 605, (1983).
[10] B.I. Halperin, Phys. Rev. Lett., 52, 1583, (1984).
[11] J.K. Jain, Phys. Rev. Lett., 63, 199, (1989).
[12] D. Arovas, J.R. Schriffer, and F. Wilczek, Phys. Rev. Lett., 53, 722, (1984).
[13] S.M. Girvin, A.H. MacDonald, and P.M. Platzman, Phys. Rev. B, 33, 2481, (1986).
[14] F.G.H. Haldane and E.H. Rezayi, Phys. Rev. Lett., 54, 359, (1985).
[15] I.V. Lerner and Yu.E. Lozovik, Zh. Eksp. Teor. Fiz., 78, 1167, (1980).
[16] I.V. Lerner and Yu.E. Lozovik, J. Low Temper. Phys., 38, 333, (1980).
[17] I.V. Lerner and Yu.E. Lozovik, Zh. Eksp. Teor. Fiz. 80, 1488, (1981); Sov.Phys.-JETP 53, 763, (1981).

[18] A.B. Dzyubenko and Yu.E. Lozovik, Fiz. Tverd. Tela (Leningrad), 25, 1519, (1983); 26, 1540, (1984); Sov. Phys. Solid State 25, 874, (1983); 26, 938, (1984); J. Phys. A, 24, 415, (1991).
[19] D. Paquet, T.M. Rice, and K. Ueda, Phys. Rev. B, 32, 5208, (1985); T.M. Rice, D. Paquet and K. Ueda, Helv. Phys. Acta, 58, 410, (1985).
[20] S.A. Moskalenko, M.A. Liberman, D.W. Snoke, and V. Botan, Phys. Rev. B, 66, 245316, (2002).
[21] S.A. Moskalenko, M.A. Liberman, D.W. Snoke, V. Botan, and B. Johansson, Physica E, 19, 278, (2003); V. Botan, M.A. Liberman, S.A. Moskalenko, D.W. Snoke, and B. Johansson, Physica B, 346-347 C, 460, (2004).
[22] S.A. Moskalenko, M.A. Liberman, P.I. Khadzhi, E.V. Dumanov, Ig.V. Podlesny, and V. Botan, Sol. State Comm., 140/5, 236, (2006); S.A. Moskalenko, M.A. Liberman, P.I. Khadzhi, E.V. Dumanov, Ig.V. Podlesny, and V. Botan, Physica E, 39/1, 137, (2007).
[23] V.M. Apalkov and E.I. Rashba, Phys. Rev. B, 46, 1628, (1992); ibid. 48, 18312, (1993); idem. Pisma Zh. Eksp. Teor. Fiz., 54, 160, (1991); ibid. 55, 38, (1992).
[24] X.M. Chen and J.J. Quinn, Phys. Rev. Lett., 70, 2130, (1993).
[25] A. Griffin, D.W. Snoke, and S. Stringari, Ed. by, Bose-Einstein Condensation, Cambridge University Press, Cambridge, 602 p., 1995.
[26] L.V. Butov, A.L. Ivanov, A. Imamoglu, P.B. Littlewood, A.A. Shashkin, V.T. Dolgopolov, K.L. Campman, and A.C. Gossard, Phys. Rev. Lett., 86, 5608, (2001).
[27] L.V. Keldysh, and A.N. Kozlov, Zh. Eksp. Teor. Fiz. 54, 978, (1968); Sov. Phys.- JETP, 27, 52, (1968).
[28] S. Das Sarma and A. Madhukar, Phys. Rev. B, 23, 805, (1981).
[29] R.P. Feynman, Statistical Mechanics, Chap. 11, Benjamin, Reading, Mass, 354 p., 1972.
[30] C. Kallin and B.J. Halperin, Phys. Rev. B, 30, 5655, (1984).
[31] H.A. Fertig, Phys. Rev. B, 40, 1087, (1989).
[32] Y. Kuramoto and C. Horie, Solid State Commun., 25, 713, (1978).
[33] S.A. Moskalenko and D.W. Snoke, Bose-Einstein condensation of excitons and biexcitons and coherent nonlinear optics with excitons, Cambridge University Press, Cambridge UK, New York USA, 415 p., (2000).
[34] N.N. Bogoliubov, Collection of papers in three volumes, [in Russian], Naukova Dumka, Kiev, vol. 2, 674 p. and vol. 3, 487 p., 1971.
[35] A.A. Abrikosov, L.P. Gor'kov, and I.E. Dzyaloshinskii, Methods of quantum field theory in statistical physics, Dover, New York, 444 p., 1975.
[36] D.N. Zubarev, Sov. Phys. Uspekhi Fiz. Nauk, 71, 71, (1960).

THE KINETICS OF PHASE TRANSITIONS: I. THE ROLE OF HETEROGENEITY

F. Paladi* and V. Gamurari

Department of Theoretical Physics, Moldova State University, 60, A. Mateevici str., MD-2009, Chisinau, Republic of Moldova
**Tel.: +373 22 577584; fax: +373 22 244248; E-mail: fpaladi@usm.md*
(Received 7 December 2009)

A kinetic model involving one order parameter was developed to study the impact of heterogeneity on the nucleation. The role of asymmetry on phase transitions in the presence of an intermediate metastable state is described in comparison with the generic model. In particular, the discovery of anomalous generation and extinction of crystal nuclei in nonequilibrium supercooled liquids reflects experimentally the effect of the real heterogeneous structure and its irreversible relaxation on the stability of nuclei. It was found that the cubic term in the order parameter of the Landau-type potential associated with the asymmetry of the system serves to increase the stability of crystalline state for the suitably varied control parameter. The cases which allow analytical solution of the problem and the realization of different transition scenarios are also discussed.

1. Introduction

Nucleation is the process the formation of new phases begins with; thus, it is a widely spread phenomenon in both nature and technology. Condensation and evaporation, crystal growth, deposition of thin films, and overall crystallization are only some of the processes in which nucleation plays a prominent role. *Fundamentally*, it is a topic in the physics of the phase transitions of first order. *Applicationally*, nucleation is a topic in the modern production of traditional and new materials and coatings for the needs of various technologies; therefore, nucleation theory, experiment, and practice are an interdisciplinary topic. *Academically*, it is an ingredient in the more general university courses on thermodynamics, solid state physics, physical chemistry, colloid and surface science, biophysics, etc. [1].

Discovery of the generation and extinction of crystal nuclei at very low temperatures [2-4] explores quite a different way to crystallize any liquid which is hard to crystallize, since it has been commonly recognized that the only way to crystallization is to wait for a long time at some temperatures. Considering that the crystal nucleation is just one extreme event in the fluctuation of a clustered structure, another metastable liquid phase with a structure different from the ordinary one would be also potentially nucleated in a similar procedure. Thus, stochastic generation of crystal nuclei can be considered a result of fluctuation of the cluster structure of a supercooled liquid. Another liquid phase with a structure different from the ordinary one could be generated in a similar procedure, and a liquid-liquid phase transition found experimentally in real systems is closely related to this phenomenon [5, 6]. The role of an intermediate metastable state in the fluctuation-induced transitions between two stable states was analyzed using kinetic models involving one or two order parameters [7, 8], the main aim being to highlight the conditions for which the transition rates can be amplified by the presence of an intermediate metastable state. The intermolecular attractive interactions are taken to be weak and short ranged, so the crystallization proceeding with crystal nucleation and growth processes is a very slow process in such substances. Meanwhile, the intrinsic transition dynam-

ics can be really enhanced for particular combinations of control parameters. This situation might correspond to crystallization in a temperature range much below $T_{g\alpha}$ produced by molecular rearrangements at the interface between clusters due to the β relaxation process [9]: the α process is related to the rearrangement motion of molecules within the clusters, while the β process is due to the rearrangement of molecules in the gap between the clusters; so the α process, being just the β process modified through the formation of clusters as a result of intermolecular interactions, becomes equal to the β process at high temperature [10].

In the present work, the generic kinetic model involving one order parameter [8] is developed to study the impact of heterogeneity on phase transitions in the presence of intermediate metastable state. The Landau-type potential possessing a single order parameter associated to the fluid phase and three control parameters is considered, and we aim mainly to assess the above mentioned effect in the presence of an intermediate metastable state rather than to study the nucleation process quantitatively. In general, the universal features of transition scenario can be described by a six-degree potential with four control parameters which are associated with coupling the system to an external field, diffusion, asymmetry, and viscosity. Depending on the values of its control parameters, the potential has one, two, or three possible minima, and we study the problem by construction of the equilibrium phase diagrams. Pointing out the range of parameter values for which the enhancement or decline in nucleation rate is observed, it is possible to analyze the impact of heterogeneity on nucleation. In fact, the occurrence of structural fluctuations in a temperature range much below $T_{g\alpha}$ indicates that a supercooled liquid would really have a heterogeneous structure composed of regions of molecules giving rise to both the α and β relaxation processes.

2. Intrinsic mechanism of phase transitions in the presence of an intermediate metastable state

In the following, the process of crystal nuclei formation is associated to the dynamics of a set of order parameters in a force field in the presence of thermal fluctuations. In the approximation of a local equilibrium, it is assumed that the force field has the structure of the derivative of a Landau-type free energy function, up to proportionality factors related to diffusion, viscosity, etc. [8]. Firstly, one can limit the model to the case of a single order parameter x, and two control parameters, and let $U(x)$ be the corresponding effective potential. Evolution of x is described by Langevin equation

$$\frac{dx}{dt} = -\frac{\partial U}{\partial x} + F(t), \tag{1}$$

where $U(x) = -\lambda x^2/2 + \mu x^4/4 + x^6/6$, and $F(t)$ is the force associated with thermodynamic fluctuations (so-called "white noise" or stochastic thermal fluctuations). Assuming that white noise process accounting for the thermodynamic fluctuations has the mean value $\langle F(t) \rangle = 0$, the intrinsic transition dynamics as a function of the above-mentioned two control parameters, namely λ and μ, is described by steady-state solutions of the equation

$$x_s^5 + \mu x_s^3 - \lambda x_s = 0 \tag{2}$$

up to a proportionality factor related to the diffusion or the viscosity coefficient, which is absorbed in this equation in a redefinition of the time scale. Eq. (2) has the trivial solution $x_0 = 0$ for any possible value of λ. Meanwhile, the non-trivial solutions are $x_{1,2,3,4} = \pm\sqrt{(-\mu \pm \sqrt{\mu^2 + 4\lambda})/2}$. These solutions are physically unacceptable for $\lambda < 0$ and $\mu > 0$, because the problem gets a negative value for x_s^2. At $\lambda > 0$ there are two solutions

$$x_s^2 = \frac{-\mu + \sqrt{\mu^2 + 4\lambda}}{2}, (\lambda > 0). \quad (3)$$

For $\lambda < 0$ and $\mu < 0$ four non-trivial solutions are physically acceptable while condition $\mu^2 + 4\lambda > 0$ holds

$$x_+^2 = \frac{-\mu + \sqrt{\mu^2 + 4\lambda}}{2}, \quad (4)$$

$$x_-^2 = \frac{-\mu - \sqrt{\mu^2 + 4\lambda}}{2} \ (\lambda < 0, \mu < 0, \ \mu^2 + 4\lambda > 0). \quad (5)$$

Figure 1 depicts the bifurcation diagram for these steady-state solutions versus λ, i.e., the dependence of the order parameter x on the control parameter λ for $\mu = -2$. As can be seen, the negative and positive branches resulting from x_+^2 are associated with the liquid, $L1$ and the crystalline, C phases, respectively. The branch $x_0 = 0$ is stable for $\lambda < 0$ and unstable for $\lambda > 0$; so, in the previous case it corresponds to the second fluid state $L2$. Notice that there are two limit point bifurcations at $\lambda = -\mu^2/4$, while two branches x_-^2 start from $x_s = 0$ through a bifurcation. Branches $L1$, C, and $L2$ (for $\lambda < 0$) correspond to minima of $U(x)$, whereas the remaining ones corespond to maxima and are unstable. So, the transition from $L1$ to C for negative λ values involves the intermediate state $L2$. The overall dynamics involves three transitions $L1 \rightarrow L2$, $L2 \rightarrow L1$, and $L2 \rightarrow C$: the first two transitions imply the crossing of an unstable state given by the negative root of x_-^2, whereas the third one refers to the crossing of the branch obtained from its positive root. It was shown [8] that the mean transition time between state $L1$ and state C decreases when the system enters the region of existence of $L2$. The particular value of λ for which the transition dynamics is the fastest one, i.e., the mean transition time goes through a minimum, corresponds to the "coexistence" of states $L1$ and $L2$: the potential barriers for the $L1 \rightarrow L2$ and the $L2 \rightarrow L1$ transitions become identical. Therefore, the presence of such intermediate metastable state, namely, the fluid phase $L2$, enhances the nucleation process.

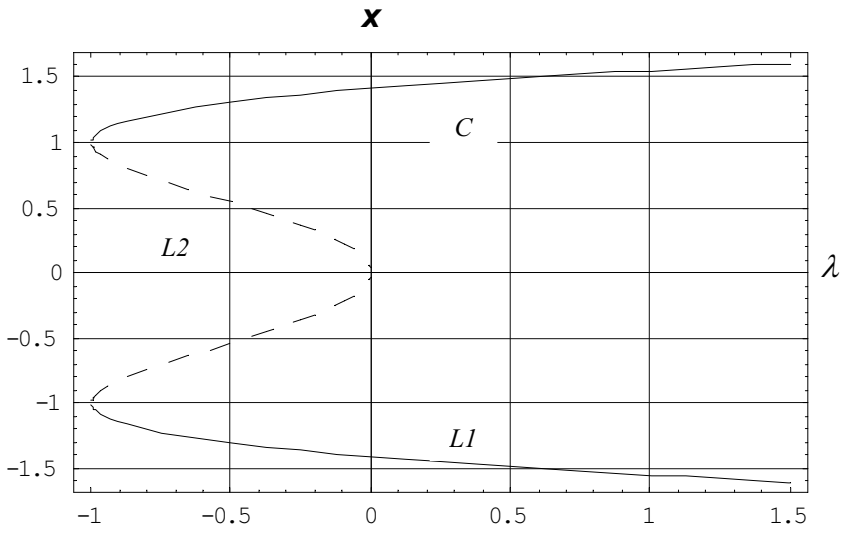


Fig. 1. Bifurcation diagram for the intrinsic transition dynamics of a homogeneous system. Parameter value $\mu = -2$. The solid and broken lines stand for the stable and metastabile states; C, $L1$, and $L2$ represent crystalline and two liquid phases, respectively.

3. The role of heterogeneity in the first-order phase transitions

It was experimentally found that the nucleation starts at temperatures much below those calculated for the homogeneous nucleation [1]; so, we shall consider the heterogeneous nu-

cleation which takes place when the old phase contacts and/or contains other phases and/or molecular species. While the supercooled liquid is microscopically heterogeneous in structure due to the presence of clusters, there should be an effect of asymmetry on the transition between phases in the real systems. The kinetic potential involving a single order parameter x and the coefficient of asymmetry ξ has the form

$$U(x) = -\lambda \frac{x^2}{2} + \xi \frac{x^3}{3} + \mu \frac{x^4}{4} + \frac{x^6}{6}, \tag{6}$$

where λ and μ are the control parameters related to the intrinsic transition dynamics [8]. Assuming that white noise process accounting for the thermodynamic fluctuations has the mean value $\langle F(t) \rangle = 0$, the possible states of the system x_s described by Eq. (6) are solutions of the equation $dx/dt = -\partial U/\partial x = 0$, so

$$x_s^5 + \mu x_s^3 + \xi x_s^2 - \lambda x_s = 0. \tag{7}$$

A trivial solution is $x_0 = 0$, which satisfies equation (7) for any values of parameters μ, ξ, and λ; the other four non-trivial solutions will be expressed in terms of control parameters as calculated as roots of the equation

$$x_s^4 + \mu x_s^2 + \xi x_s - \lambda = 0. \tag{8}$$

Note that (8) is already an incomplete four-degree equation, and this simplifies its analytical solution according to the Descartes-Euler method for solving polynomial equations. The four steady states solutions of Eq. (8) will thus take the form

$$x_{1,2,3,4} = \pm\sqrt{y_1} \pm \sqrt{y_2} \pm \sqrt{y_3}, \tag{9}$$

where the combination of signs is taken so as the following expression takes place

$$\sqrt{y_1}\sqrt{y_2}\sqrt{y_3} = -\frac{\xi}{8}, \tag{10}$$

and y_1, y_2 and y_3 are the solutions of the cubic equation

$$y^3 + \frac{\mu}{2} y^2 + \frac{\mu^2 + 4\lambda}{16} y - \frac{\xi^2}{64} = 0. \tag{11}$$

Next, we note $p = \mu/2$, $q = (\mu^2 + 4\lambda)/16$, and $r = -\xi^2/64$. Then Eq. (11) takes the form

$$y^3 + py^2 + qy + r = 0, \tag{12}$$

and, upon substitution $y = z - p/3$, this equation takes its final incomplete form

$$z^3 + az + b = 0, \tag{13}$$

where $a = -p^2/3 + q$ and $b = 2(p/3)^3 - pq/3 + r$. Solutions z_1, z_2, and z_3 of the incomplete cubic equation are

$$z_1 = A + B, \quad z_{2,3} = -\frac{A+B}{2} \pm i\sqrt{3}\frac{A-B}{2}, \tag{14}$$

where $A = \sqrt[3]{-\frac{b}{2} + \sqrt{Q}}$, $B = \sqrt[3]{-\frac{b}{2} - \sqrt{Q}}$, and $Q = \left(\frac{a}{3}\right)^3 + \left(\frac{b}{2}\right)^2$. Here A and B are, in general, complex quantities which satisfy the equation $AB = -a/3$ [11]. Next, we examine all possible cases for the model parameters; if $Q > 0$, then Eq. (13) has a physically acceptable solution and two complex conjugate roots; if $Q = 0$, then there are three physically acceptable solutions, at least two of which are equal. The solutions of these two cases are subject to the sign of the parameter a. Thus for $Q \geq 0$ and $a > 0$ one obtains the following results

$$z_1 = -2\sqrt{a/3}\,\mathrm{ctg}2\alpha, \tag{15}$$

$$z_{2,3} = \sqrt{a/3}(\mathrm{ctg}2\alpha \pm i\sqrt{3}\,\mathrm{cosec}2\alpha), \tag{16}$$

where

$$\mathrm{tg}\alpha = \sqrt[3]{\mathrm{tg}\frac{\beta}{2}}\left(|\alpha| \le \frac{\pi}{4}\right), \quad \mathrm{tg}\beta = \frac{2}{b}\sqrt{\left(\frac{a}{3}\right)^3}\left(|\beta| \le \frac{\pi}{2}\right). \tag{17}$$

For $Q \ge 0$ and $a < 0$ we can write

$$z_1 = -2\sqrt{-a/3}\,\mathrm{cosec}2\alpha, \tag{18}$$

$$z_{2,3} = \sqrt{-a/3}(\mathrm{cosec}2\alpha \pm i\sqrt{3}\,\mathrm{ctg}2\alpha), \tag{19}$$

where

$$\mathrm{tg}\alpha = \sqrt[3]{\mathrm{tg}\frac{\beta}{2}}\left(|\alpha| \le \frac{\pi}{4}\right), \quad \sin\beta = \frac{2}{b}\sqrt{\left(-\frac{a}{3}\right)^3}\left(|\beta| \le \frac{\pi}{2}\right). \tag{20}$$

If $Q < 0$, then $a < 0$ and there are three distinct real solutions

$$z_1 = 2\sqrt{-a/3}\,\cos\frac{\alpha}{3}, \tag{21}$$

$$z_{2,3} = -2\sqrt{-a/3}\,\cos\left(\frac{\alpha}{3} \pm \frac{2\pi}{3}\right), \tag{22}$$

where

$$\cos\alpha = -\frac{b}{2\sqrt{-(a/3)^3}}. \tag{23}$$

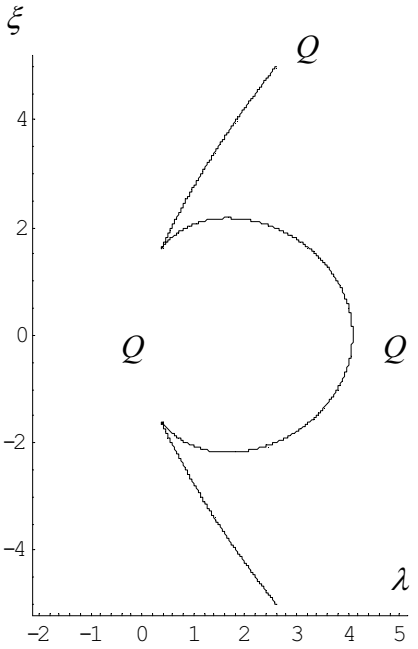


Fig. 2. Dependence of the model parameter Q on control parameters λ and ξ. There are two distinct regions where $Q < 0$ and $Q > 0$. Values in the curve correspond to $Q = 0$. Parameter value $\mu = -2$.

In all cases, one should consider the physically acceptable solutions. Note that the solution of initial equation (8) is voluminous, wherein additionally Eq. (10) is satisfied. Thus, depending on the values of the parameter Q, the solutions $x_{1,2,3,4}$ can be both physically acceptable and unacceptable; so, we shall examine further the dependence of Q on the initial control parameters μ, ξ, and λ by replacing the interim notations a, b, p, q, and r. Then the final equation for Q is given by

$$Q = \frac{1}{1728}\left(\lambda - \frac{\mu^2}{12}\right)^3 - \frac{1}{1728^2}\left(\frac{27\xi^2}{2} + 36\lambda\mu + \mu^3\right)^2. \tag{24}$$

Since $Q(\mu, \xi, \lambda)$ depends on three control parameters, we determine the sign of Q for different values of parameters λ and ξ, where $\mu = -2$. This will allow us to compare these results with the previously reported case for the intrinsic mechanism of phase transitions [8]; so, it will be possible to analyze the influence of asymmetry on the transition dynamics in the presence of an intermediate metastable state. The dependence of model parameter Q on control parameters λ and ξ is shown in Fig. 2. In this figure, one can easily follow the evolution of Q depending on the set of control parameters (λ, ξ): the output in the curve corresponds to $Q = 0$ (there are two distinct physically acceptable solutions), the region on the left contains values for

$Q < 0$ (there are three distinct physically acceptable solutions), and that on the right – for $Q > 0$ (there is only one such solution). The function $Q(\lambda, \xi)$ is symmetrical at $\xi = 0$ for any λ, because the dependence on ξ in Eq. (24) is squared. However, a particular interest in the construction of bifurcation diagram for the steady-state solutions of Eq. (8) versus λ is focused on the dependence of order parameter x in a special range of values, for example, $|\lambda| \in [0,4]$, $|\xi| \in [0,2]$, and, obviously, $\mu = -2$ as well as two limit point bifurcations at (0.33 ± 1.54). In this range of values for λ, as shown in Fig. 2 for $\xi = \pm 2$, there is a transition from $Q < 0$ to $Q = 0$, and then to $Q > 0$, and, again, $Q = 0$ and $Q < 0$, and, finally, $Q = 0$ and $Q > 0$, respectively, from left to right. The results analyzed in this section are summarized below in Table 1.

Table 1. Steady-state solutions of Eq. (8) obtained for the kinetic potential involving a single order parameter and the coefficient of asymmetry.

Solution	Condition	Parameters
$x_{1,2,3,4} = \pm\sqrt{y_1} \pm \sqrt{y_2} \pm \sqrt{y_3}$	$\sqrt{y_1}\sqrt{y_2}\sqrt{y_3} = -\frac{\xi}{8}$	$p = \mu/2$, $q = (\mu^2 + 4\lambda)/16$, $r = -\xi^2/64$
$y = z - \frac{p}{3}$		
$z_1 = A + B$, $z_{2,3} = -\frac{A+B}{2} \pm i\sqrt{3}\frac{A-B}{2}$		$a = -\frac{p^2}{3} + q$, $b = 2\left(\frac{p}{3}\right)^3 - \frac{pq}{3} + r$, $A = \sqrt[3]{-\frac{b}{2} + \sqrt{Q}}$, $B = \sqrt[3]{-\frac{b}{2} - \sqrt{Q}}$, $Q = \frac{1}{1728}\left(\lambda - \frac{\mu^2}{12}\right)^3 - \frac{1}{1728^2}\left(\frac{27\xi^2}{2} + 36\lambda\mu + \mu^3\right)^2$
Possible cases for $\mu = -2$		
$x_{1,2,3,4} = \pm\sqrt{y_1} \pm \sqrt{y_2} \pm \sqrt{y_3}$	$\sqrt{y_1}\sqrt{y_2}\sqrt{y_3} = -\frac{\xi}{8}$	$p = -1$, $q = (1+\lambda)/4$, $r = -\xi^2/64$
$y = z + \frac{1}{3}$		$b = \frac{\lambda}{12} - \frac{\xi^2}{64} + \frac{1}{108}$, $A = \sqrt[3]{-\frac{b}{2} + \sqrt{Q}}$, $B = \sqrt[3]{-\frac{b}{2} - \sqrt{Q}}$, $Q = \frac{1}{1728}\left(\lambda - \frac{1}{3}\right)^3 - \frac{1}{1728^2}\left(8 - \frac{27\xi^2}{2} + 72\lambda\right)^2$
$z_1 = A + B$, $z_{2,3} = -\frac{A+B}{2} \pm i\sqrt{3}\frac{A-B}{2}$		

1.	$Q < 0$	values (λ, ξ) on the left side of the curve in Fig. 2; physically acceptable distinct solution $z_1 \neq z_2 \neq z_3$
2.	$Q = 0$	values (λ, ξ) on the curve in Fig. 2; physically acceptable solutions z_1 and $z_2 = z_3$
3.	$Q > 0$	values (λ, ξ) on the right side of the curve in Fig. 2; only one physically acceptable solution z_1

Figure 3 depicts the bifurcation diagram of the steady-state solutions of Eq. (8) versus λ, i.e., the dependence of the order parameter x on the control parameter λ in the range of values $\lambda \in [-4,4]$, and for $\mu = -2$ and $\xi = 0$, 1, and 2. As can be seen, the lowest negative and the highest positive branches are associated with the liquid, $L1$ and the crystalline, C phases, re-

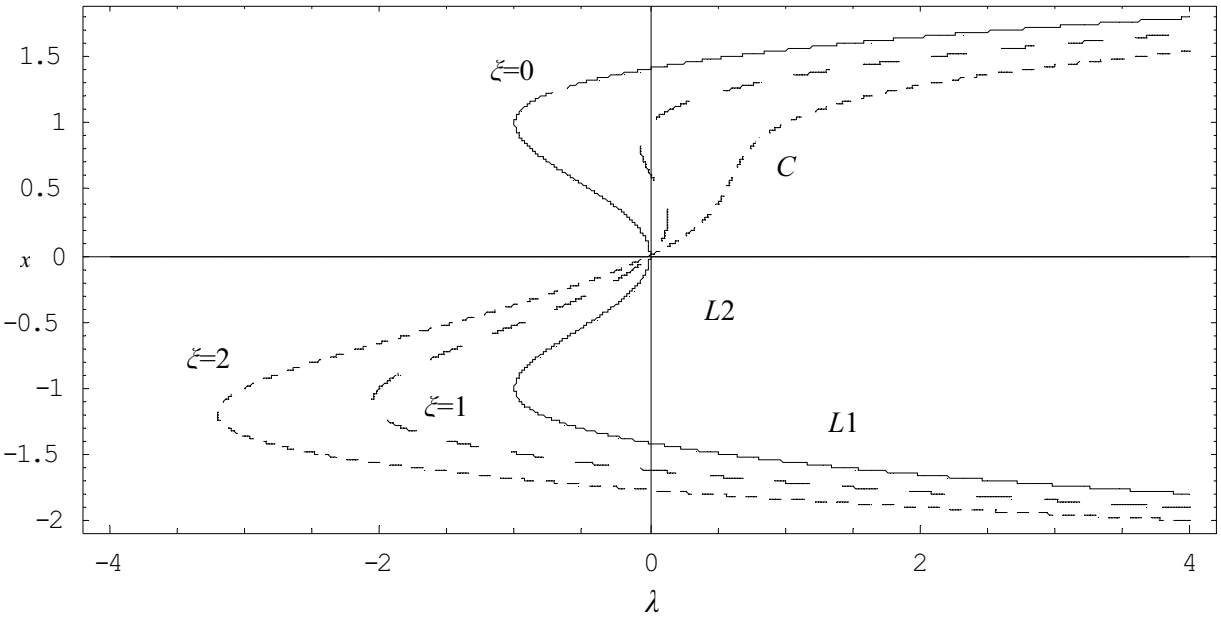


Fig. 3. Bifurcation diagram for a heterogeneous system. Solid line stands for the intrinsic transition dynamics ($\xi = 0$), and the broken and dotted curves correspond to the kinetic potential involving an additional asymmetry effect with $\xi = 1$ and 2. C, $L1$, and $L2$ represent crystalline and two liquid phases, respectively. Parameter value $\mu = -2$.

spectively. The branch $x_0 = 0$ for $\lambda < 0$ corresponds to the second fluid phase $L2$. The solid line corresponds to the referred case of the intrinsic transition dynamics as a function of the remaining two control parameters denoted by λ and μ, when $\xi = 0$. Branches $L1$, C, and $L2$ (for $\lambda < 0$) correspond to minima of the Landau-type potential $U(x)$, whereas the remaining ones are unstable and correspond to maxima. We see that, with an increase in the asymmetry of the system, i.e., as the absolute value of parameter ξ gets bigger, there is an increase in the stability of the fluid phase $L1$ for $\xi > 0$, whereas the stability of the crystalline C phase decreases, and vice versa for $\xi < 0$, when the stability of the liquid phase $L1$ decreases, there is an increase in the stability of the corresponding solid phase. Two limit point bifurcations change accordingly. These facts are represented in the figure by the broken and dotted curves drawn for $\xi = 1$ and 2, respectively. First-order phase transitions occur by the nucleation mechanism; based on the 'cluster structure model' [9, 10], this kinetic aspect of structure relaxation is understood as follows: the supercooled liquids and glasses are pictured to have a statically heterogeneous structure composed somehow of ordered clusters and the interfacial parts (of the clusters) with somewhat disordered arrangement. Furthermore, this study supports the idea that the role of heterogeneity in nucleation is to help the formation of the clusters of the new phase.

4. Conclusions

In this paper, a Landau-type potential based kinetic model involving one order parameter has been developed to study the effect of asymmetry on phase transitions in the presence of an intermediate metastable state. The analytical solutions of the problem were investigated, and the model provides a set of three control parameters necessary to realize different transition scenarios, which therefore assign a high degree of universality to this model as long as the respective control parameters are suitably varied.

The intrinsic transition dynamics shows that there is a particular value of λ for which the transition dynamics is the fastest one, and it corresponds to the "coexistence" of fluid phases

$L1$ and $L2$. The presence of such intermediate metastable state, namely, the fluid phase $L2$, enhances therefore the nucleation process. While the asymmetry of the system increases, there is an increase in the stability of the fluid phase for $\xi > 0$, whereas the stability of the crystalline phase decreases; and vice versa for $\xi < 0$, when the stability of the liquid phase decreases, there is an increase in the stability of the solid one. First-order phase transitions occur by the nucleation mechanism; so, these results are primarily related to the nucleation phenomena. In particular, our results support the 'cluster structure model' of supercooled liquids and glasses, according to which the role of heterogeneity in nucleation is to enhance the formation of crystal nuclei as clusters of the new phase. Finally, we can conclude that the model refers also to the problem of crystallization in substances where fluctuating metastable states have been already observed experimentally [2-4, 12-14]; also, it offers possible solutions for controlling the phase transitions in the presence of an intermediate metastable state.

Acknowledgements

F.P. is grateful to Prof. Masaharu Oguni for the guidance in this field of research during his stay in the Tokyo Institute of Technology, Japan, and thereafter.

References

[1] D. Kashchiev, Nucleation. Basic Theory with Applications, Butterworth-Heinemann, Oxford, 529 p., 2000.
[2] F. Paladi and M. Oguni, Phys. Rev. B, 65, 144202, (2002).
[3] F. Paladi and M. Oguni, J. Phys. Condens. Matter, 15, 3909, (2003).
[4] S. Tomitaka, M. Mizukami, F. Paladi, and M. Oguni, J. Therm. Anal. Cal., 81, 637, (2005).
[5] P.H. Poole, F. Sciortino, U. Essmann, and H.E. Stanley, Nature, 360, 324, (1992).
[6] S. Aasland and P.F. McMillan, Nature, 369, 633, (1994).
[7] G. Nicolis and C. Nicolis, Physica A, 323, 139, (2003).
[8] G. Nicolis and C. Nicolis, Physica A, 351, 22, (2005).
[9] M. Oguni, J. Non-Cryst. Solids, 210, 171, (1997).
[10] H. Fujimori and M. Oguni, Solid State Commun., 94, 157, (1995).
[11] G.A. Korn and T.M. Korn, Mathematical Handbook for Scientists and Engineers, McGraw-Hill, New York, 1130 p., 1968.
[12] P.R. ten Wolde and D. Frenkel, Science, 277, 1975, (1997).
[13] V. Talanquer and D.W. Oxtoby, J. Chem. Phys., 109, 223, (1998).
[14] O. Galkin and P.G. Vekilov, J. Cryst. Growth, 232, 63, (2001).

THE KINETICS OF PHASE TRANSITIONS: II. THE ROLE OF INTERFACE BETWEEN CLUSTERS

F. Paladi*

Department of Theoretical Physics, Moldova State University, 60, A. Mateevici str., MD-2009, Chisinau, Republic of Moldova
**Tel.: +373 22 577584; fax: +373 22 244248; E-mail: fpaladi@usm.md*
(Received 7 December 2009)

The present developments in the physics of complex systems, in particular, the structural relaxation of supercooled liquids and glasses, are discussed by using a stochastic cluster-based model. We are able to depict the impact of the interface between the nucleus considered as a cluster of a certain number of molecules and the liquid phase for the enhancement of the overall nucleation process. It is also shown that even a relatively simple stochastic model, which appears phenomenological if it is not agent-based, can describe precisely the outcomes from multiple agent-based simulation runs where probabilistic insight is lacking and which should be long enough to equilibrate the states of large systems.

1. Introduction

Clustering is often described by Ewens sampling formula, which admits an interpretation in terms of rational versus herding behavior [1]. An alternative model was inspired from Simon's approaches to the problem. It differs from the Ewens model both in destruction and in creation. In particular, the probability of herding does not depend on the size of the herd; this assumption destroys the exchangeability of the random partitions and forbids an analytical solution [2]. The herding models proposed by Rodgers et al. [3−5] incorporate either coagulation and fragmentation, attachment and fragmentation, growth and addition or growth and coagulation of groups of agents at each time step. We have also introduced an alternative kinetic model based on the cluster concept [6]. In terms of clusters, growth/removal is the process in which a new cluster or a free agent is introduced/removed to/from the system. Fragmentation can be defined as a process in which a cluster breaks up into isolated agents. Coagulation takes place when two clusters join making a single one. Addition can be described as a special type of coagulation in which a free agent already in the system is added at random to another cluster. Attachment is the process in which new incoming agents attach themselves to an existing cluster, and this allows both the system and the clusters to grow in size. Only restricted combinations of these ingredients cause the models to differ.

One aim of this paper is to get new insights into microscopic explanations of stochastic models, which may be compared with the agent-based computational models, and to bridge the gap between agent-based models (ABM) and stochastic processes. The application refers to the nucleation process, a widely spread phenomenon in both nature and technology, which may be considered as a representative of the aggregation phenomena in complex systems. The recent discovery of the generation and extinction of crystal nuclei at very low temperatures [7−9] suggests that stochastic generation of crystal nuclei would be considered as a result of fluctuation of complex cluster structure of the supercooled liquid. The role of both heterogeneity and the interface between clusters in the enhancement of nucleation rate has still to be explained. In particular, it was observed that nuclei could almost always be formed near the surface of the cluster instead of in the interior, and one factor favoring nucleation

near the surface would be the greater freedom of motion and, hence, a larger nucleation probability [10−12]. This is surprising, because it is known that the surface layers of the nuclei tend to be disordered and melt at significantly lower temperatures than their cores.

2. A generic stochastic model for crystal nucleation

Let us consider N atoms, which can be in three different states (cluster, liquid, and their interface), and can perform four possible moves: liquid to interface, interface to liquid, interface to cluster, and cluster to interface. One can identify four different combinations denoted with probabilities $p_1...p_4$. That is, drawing randomly one particle, it will be of type i with probability p_i. Let N=1,2,…∞ be the total number of atoms in the system, and $\{n_1, n_2, n_3, n_4\}$ are their partition into four subsets. Each subset can be called cluster; the process proper, clustering. Schematic distribution of particles in clusters is shown in Fig. 1. This diagram illustrates a potential process for the generation of a crystal nucleus in the course of irreversible structural relaxation of the nonequilibrium supercooled liquid. The number of possible partitions in this case is $P(N) = \frac{1}{3!}\prod_{i=1}^{3}(N+i)$, where $n_i = \overline{0, N}$, $i = \overline{1,4}$ and $\sum_{i=1}^{4} n_i = N$. For example, in a system of N=1000 atoms, $P(N)$ equals to 167668501! Such an interaction in the ABM model always involves an active agent and a passive one: the agents have preferences over their states and they can play both roles interchangeably. Accordingly, the number of repeated computer runs due to different possible partitions would be very large. As already mentioned in the previous section, we are able to overcome this problem by developing similar stochastic analytical models which can describe exactly the results of the agent-based computational models, and, finally, by bridging the gap between ABM modelling and stochastic processes.

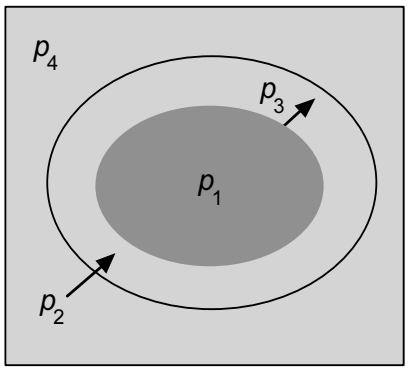


Fig. 1. Schematic diagram illustrating the distribution of particles with corresponding share in cluster, p_1, liquid, p_4, and at the interface, p_2 and p_3. The darkness represents the degree of order in the molecular arrangements, and the encircled part stands schematically for a more ordered cluster, namely, a crystal nucleus.

Let us consider further that each particle interacts with the entire group both as an aggressor (in terms of the Kolmogorov theory) and as a passive agent (in terms of the ABM models) as well. Then the mean π, namely, the stability index, takes here the form

$$\pi = p_1(p_1 - p_3) + (p_2 + p_3)(-p_1 + p_2 + p_3 - p_4) + p_4(p_4 - p_2). \quad (1)$$

or, taking into account that $\sum_{i=1}^{4} p_i = 1$, one can exclude one probability, for example, p_4, from the above equation

$$\pi = p_1(p_1 - p_3) + (1 - p_1 - 2p_2 - p_3)(1 - p_1 - p_2 - p_3) + (p_2 + p_3)(-1 + 2p_2 + 2p_3). \quad (2)$$

One can represent the distribution of states as a three-dimensional point $d(l,r,c) \equiv d(p_1, p_2, p_3, p_4) = \sqrt{l^2 + r^2 + c^2}$, where the axes are labeled l, c, and r: $l = p_1 - p_3$, $c = p_2 + p_3 - p_1 - p_4$, $r = p_4 - p_2$, and as before $l + c + r = 0$ and $d \in \left[0, \sqrt{2}\right]$. Thus different distributions of states can lead to the same point in the sphere; that is, different microscopic partitions can generate the same result on aggregate inside a sphere around the origin.

Preliminary results for the two limit cases are obvious: if all particles show the same behavior, then $d = \sqrt{2}$ and there is a maximum stability of states in such a completely asymmetrical system, but $\pi = 0$ for a homogeneous system, $p_1 = p_2 = p_3 = p_4 = 1/4$, and for combinations, such as $p_1 = p_3$ and $p_2 = p_4$ in the case of unstable steady-states.

To represent a two-dimensional graph describing the stability/instability of the system, we consider some fixed probabilities. Figure 2 depicts such dependence of the cluster instability, $1–\pi$, on the probability p_1 for a system of 100 particles, where the atoms at the liquid-cluster interface are missing (a); share of atoms at the interface p_2+p_3 is just 1/10 of their total number in the system (b); and corresponding share of particles at the interface is equal to 3/10, i.e. $p_2 = p_3 = 0.15$ (c), $p_2 = 0.1$, $p_3 = 0.2$ (d), and $p_2 = 0.2$, $p_3 = 0.1$ (e). Note that, in the absence of an interface between liquid and cluster, as we expected, the system is in a state of maximum instability for $p_1 = p_4 = 0.5$. While the number of particles at the interface, i.e., p_2+p_3 increases, the stability of the system decreases simultaneously, regardless of whether the particle flow at the interface is achieved at the expense of the liquid phase or particles in the cluster. However, particles at the liquid-cluster interface definitely accelerate the formation of clusters due to the displacement of the maximum instability in the region of smaller values for p_1. In other words, the nucleus formation is indeed a random event with a chance largely determined by the nucleation work which increases with the subnuclei size [13]; thus, a decline in share of atoms, p_1 in such a cluster, namely a decline in the critical nucleus size, would be followed by the appearance of a crossover point to the supernuclei at the lower value of the energy consumed for the cluster formation. The curves (d) and (e) in this figure also show that a greater flow from the liquid phase ($p_2 > p_3$) or from the cluster ($p_2 < p_3$) causes a minor increase in the instability of the related branches, but for $p_2 + p_3 =$ const the value of maximum π remains constant too.

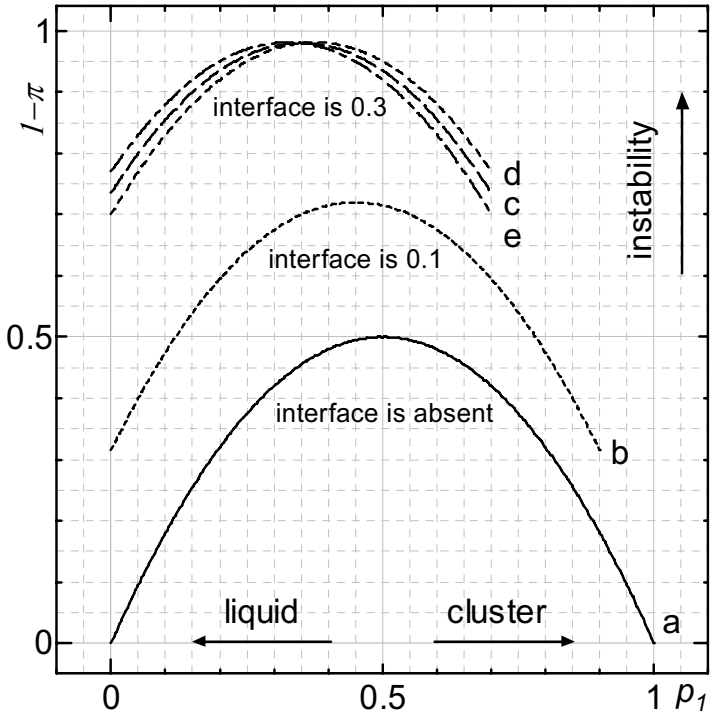


Fig. 2. Dependence of the cluster instability, $1–\pi$, on the probability p_1 for the cases: (a) particles at the liquid-cluster interface are missing; (b) share of atoms at the interface represents 1/10 of their total; corresponding share of atoms at the interface represents 3/10, i.e. $p_2 = p_3 = 0.15$ (c), $p_2 = 0.1$, $p_3 = 0.2$ (d), and $p_2 = 0.2$, $p_3 = 0.1$ (e). A total number of 100 agents in a similar ABM model is considered.

3. Conclusions

We propose a generic stochastic model for crystal nucleation. It is generally known that first-order phase transitions occur by the nucleation mechanism and that both the nucleus, a cluster of molecules or atoms, and the nucleation work, an energy barrier to the phase transition, are

basic thermodynamic quantities in the theory of nucleation. However, the critical nucleus formation is statistically a random event with a probability largely determined by the nucleation work. It is shown that, while the number of particles at the liquid-cluster interface increases, the stability of the entire system decreases simultaneously, and the nucleus formation would be definetely enhanced due to the displacement of the bifurcation point in the region of smaller clusters.

References

[1] W.J. Ewens, Theor. Pop. Biol., 3, 87, (1972).
[2] D. Costantini, S. Donadio, U. Garibaldi, and P. Viarengo, Comput. Econ., 27, 115, (2006).
[3] G.J. Rodgers and Y.J. Yap, Eur. Phys. J. B, 28, 129, (2002).
[4] G.J. Rodgers and D. Zheng, Physica A, 308, 375, (2002).
[5] S. Rawal and G.J. Rodgers, Physica A, 344, 50, (2004).
[6] F. Paladi and V. Eremeev, Physica A, 348, 630, (2005).
[7] F. Paladi and M. Oguni, Phys. Rev. B, 65, 144202, (2002).
[8] F. Paladi and M. Oguni, J. Phys. Condens. Matter, 15, 3909, (2003).
[9] S. Tomitaka, M. Mizukami, F. Paladi, and M. Oguni, J. Therm. Anal. Cal., 81, 637, (2005).
[10] S. Valkealahti and M. Manninen, J. Phys. Condens. Matter, 9, 4041, (1997).
[11] Y.G. Chushak and L.S. Bartell, J. Phys. Chem. B, 103, 11196, (1999).
[12] Y.G. Chushak and L.S. Bartell, J. Phys. Chem. A, 104, 9328, (2000).
[13] D. Kashchiev, Nucleation. Basic Theory with Applications, Butterworth-Heinemann, Oxford, 529 p., 2000.

CHAOS COMMUNICATION USING SEMICONDUCTOR LASERS SUBJECT TO DIFFERENT KINDS OF OPTICAL FEEDBACK

V.Z. Tronciu

Technical University of Moldova, 168, Stefan cel Mare ave., MD-2005, Chisinau, Republic of Moldova
Weierstrass Institute for Applied Analysis and Stochastics, 39, Mohren str., 10117, Berlin, Germany
Institute of Applied Physics, Academy of Sciences of Moldova, 5, Academiei str., MD-2028, Chisinau, Republic of Moldova
E-mail: tronciu@mail.utm.md
(Received 11 July 2009)

Abstract

In this review, we report the results of numerical investigations of the dynamical behavior of an integrated device composed by a semiconductor laser and different cavities that provide optical feedback. Due to the influence of the feedback, under the appropriate conditions the systems display chaotic behavior appropriate for chaos based communications. The optimal conditions for chaos generation are identified. It is found that the longitudinal double cavity feedback requires lower feedback strengths for developing high complexity chaos as compared with a single cavity. The synchronization of two unidirectional coupled (master-slave) systems and the influence of parameters mismatch on the synchronization quality are also studied. Examples of message encoding and decoding within chaos modulation technique for longitudinal double cavity optical feedback are presented and discussed. We find that the resynchronization time for the T-like double cavity optical feedback scheme can be two orders of magnitude shorter as compared with that of the single-cavity feedback case. Very good conditions for message encoding by using the on/off phase shift keying encryption method are identified, and examples of message encoding/decoding are presented.

1. Introduction

The synchronization of chaotic oscillators has been a subject of deep studies in the last years due to its fundamental and applied properties [1]. In terms of application, chaos based communications have become an option to improve privacy and security in data transmission, especially after the recent field demonstration on the metropolitan fiber networks of Athens [2]. In optical chaos based communications, the chaotic waveform is generated by using semiconductor lasers with either all-optical [3-7] or electro-optical [8-10] feedback loops. In particular, semiconductor lasers subject to the influence of optical feedback from a distant mirror have been investigated extensively for the past two decades, and different dynamical behaviors have been characterized, including periodic and quasi-periodic pulsations, low frequency fluctuations, and coherent collapse (for more detail, see Ref. [11]). In the Conventional all-Optical Feedback case (COF) typically, to achieve chaotic behavior a delay round trip time of at least a few hundreds of ps is needed. So, in the air, the external cavity should be about a few cm long, which is a drawback for the design of compact chaotic sources. Integrated lasers with ultra-short feedback cavities have also revealed similar characteristics if the

feedback is properly amplified [12]. In this context, multi-section lasers with an amplified feedback section could be suitable candidates for integrated chaotic emitters. Due to the continuing technological progress, multi-section lasers have stable and compact configurations, which include integrated sections with common waveguides and a tunable phase shift [13]. However, the simplest configuration, a two-section laser with one active section and one passive section acting as an external cavity, is not suitable since the length of the passive section is typically too short to achieve chaotic dynamics.

Lasers subject to feedback from two cavities have been considered in several configurations [14-18]. In particular, feedback from a second cavity has been used to control the chaotic dynamics of semiconductor lasers with optical feedback. Control in the low frequency fluctuation regime has been achieved by adjusting properly both the length and the feedback strength of the second external cavity. Configurations using Fabry-Perot resonators to provide feedback have also been studied [19, 20]. In this case, the feedback can destabilize the laser emission but can also improve the stability of continuous wave (CW) emission by enhancing the damping of relaxation oscillations or allowing the control of the chip in a non-invasive way.

Several ways for encoding and decoding a message within the chaotic carrier has been proposed in the literature, including chaos modulation [3], chaos shift keying [21], chaos masking [22], etc. Here we consider two kinds; the chaos modulation (CM) technique applied to longitudinal double cavity feedback scheme and On/Off Phase Shift Keying (OOPSK) method applied to T-like double cavity optical feedback setup (see Fig. 1). Finally, we mention that this review summarizes and completes the main results of [23] and [24].

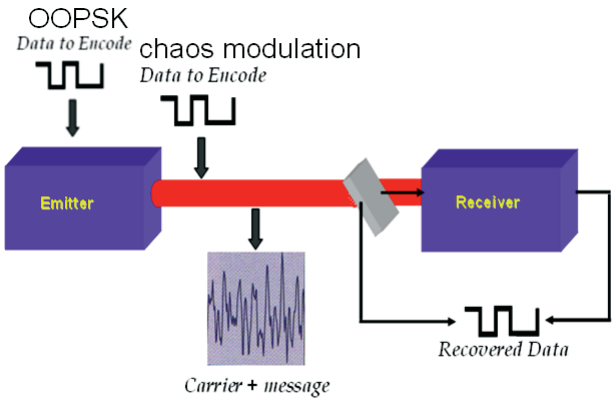


Fig. 1. Different schemes of message encryption.

The paper is structured as follows. In Section 2 we introduce an appropriate model to describe the dynamics of semiconductor laser under the influence of the longitudinal double cavity feedback (LDCF). We study the synchronization of two of such systems. The influence of the mismatch in the feedback phases on the synchronization quality is also discussed. Section 3 presents a study of the dynamics of a laser under the influence of a T-like double cavity optical feedback (TDCF). We highlight the advantages of the proposed setup when compared with the COF case and the OOPSK encryption method is demonstrated for the TDCF. The summary and conclusions are given in Section 4.

2. Longitudinal double cavity feedback. Chaos modulation technique

In this Section, we consider an integrated device composed of a semiconductor laser subject to feedback from a double cavity grown in longitudinal direction with the aim of generat-

ing a complex chaotic waveform suitable for applications in chaos based communications. The scheme of the system is depicted in Fig. 2. It consists of a single mode semiconductor laser coupled to an external passive cavity of the same III-V material through an air gap. Here we choose two external cavities (air gap and III-V material) of the same length, in this case 1 cm. To avoid diffraction losses in the air cavity, a micro lens should be placed at the laser facet to collimate the beam (not shown in our set-up). Assuming a refractive index of 3 for the material, the total delay time in the two cavities amounts approximately 0.266 ns. The advantage of this longitudinal double cavity configuration is the existence of two feedback phases, one in the air gap cavity and one in the material cavity, the latter can be easily adjusted to destabilize the dynamics of the laser. Moreover, while we assume that the first reflectivity is defined by the air-material facet, the outer facet of the material cavity can be coated to increase its reflectivity. We study the conditions for which the behavior of the system is chaotic due to the influence on the laser dynamics of the feedback from the double cavity. We also study the synchronization of two such systems under unidirectional coupling. In the absence of coupling, the behavior of transmitter and receiver systems is uncorrelated. When a certain amount of light is injected from the transmitter into the receiver, the latter is able to synchronize to the emitter under the appropriate conditions. Typically, the receiver does not synchronize identically to the emitter due to the injected field. Therefore, synchronization is not complete but generalized [25]. Synchronization is robust to small perturbations of the carrier. A message of small amplitude can then be included in the carrier which will be filtered-out by the receiver. As mentioned above, several message-encoding schemes have been proposed in the literature. Here we include the message as a modulation in the amplitude of the chaotic carrier (chaos modulation) [3] (see Fig. 1). The message can be decoded at the receiver by comparing its input (carrier mixed with message) with its output (which ideally reproduces only the carrier).

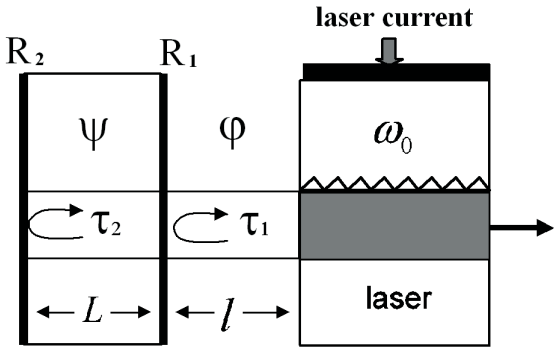


Fig. 2. A sketch of the proposed setup for chaos synchronization and message encoding, using semiconductor lasers under the influence of a feedback from a double cavity grown in longitudinal direction. R_1 and R_2 are the reflectivity of the air-material facet and the outer facet of the material cavity, respectively. The length of the air cavity l and the material cavity L are taken to be the same: $L = l = 1$ cm, $n_L = 3$; ω_0 is the free running frequency of CW laser.

2.1. Model and equations

To model the set-up shown in Fig. 2 we consider a single mode CW laser coupled to a longitudinal double cavity. The first mirror is located at distance l from the laser facet. The distance between the first and second mirror is L. The optical feedback phase in the second cavity ψ can be controlled by injecting current into the passive section. We assume that the current injected in

the passive section is small enough to affect only the refractive index, so that the optical length of the resonator is changed in the sub-wavelength range. In this way, while the feedback phase ψ can be tuned, the change in the delay time between the two mirrors τ_L is negligible. Alternatively, this phase could be also controlled with a piezo actuator. In principle, multiple reflections may take place. However, for the feedback parameters we will use, it is sufficient to consider a single reflection in both cavities. This approximation strongly simplifies the calculations.

The laser dynamics can be analyzed in the framework of the extended Lang-Kobayashi equations for the complex field amplitude E and an excess carrier density N [26, 27]

$$\begin{aligned}\frac{dE_{t,r}}{dt} &= (1+i\alpha)\left[\frac{g(N_{t,r}-N_0)}{1+\varepsilon\left|E_{t,r}\right|^2}-\frac{1}{\tau_{ph}}\right]\frac{E_{t,r}}{2}+F_{LDCF}+k_rE_t \\ \frac{dN_{t,r}}{dt} &= \frac{I_{t,r}}{e}-\frac{1}{\tau_e}N_{t,r}-\frac{g(N_{t,r}-N_0)}{1+\varepsilon\left|E_{t,r}\right|^2}\left|E_{t,r}\right|^2\end{aligned} \tag{1}$$

where the feedback term is written as

$$F_{LDCF}=\gamma_{t1,r1}e^{-i\varphi}E_{t,r}(t-\tau_l)+\gamma_{t2,r2}e^{-i(\varphi+\psi)}E_{t,r}(t-(\tau_l+\tau_L)). \tag{2}$$

The subscripts t and r refer to transmitter and receiver lasers, respectively. The last term in equation (1) is present only in the receiver laser and describes the unidirectional coupling; k_r is the coupling parameter of the injected field into the receiver laser given by $\kappa_r=\sqrt{1-R}\eta_{ext}\Big/\left(\tau_c\sqrt{R}\right)$, where R is the facet power reflectivity of the slave laser (R = 30%), τ_c is the cavity roundtrip time of the light within the laser ($\tau_c=10$ ps), η_{ext} accounts for losses different than those introduced by the laser facet ($\eta_{ext}=0.5$); τ_l and τ_L are the air gap and passive material roundtrip times, respectively; $\gamma_{t1,r1}$ and $\gamma_{t2,r2}$ are the feedback strengths governed by the reflectivity R_1 and R_2, respectively. For simplicity, we assume R_1 and R_2 such that $\gamma_{t1}=\gamma_{t2}=\gamma_{r1}=\gamma_{r2}=\gamma$; $\varphi=\omega_0\tau_l$ (whose value can strongly vary in different devices), and $\psi=\omega_0\tau_L$ are the optical phase accumulated in the air gap and material cavities, respectively. The other parameter values are the linewidth enhancement factor α = 5, the differential gain parameter $g=1.5\times10^{-8}$ ps^{-1}, the gain saturation coefficient $\varepsilon=5\times10^{-7}$, the photon and carrier lifetimes τ_{ph} = 3 ps and τ_e = 2 ns, respectively, and the round trip times τ_l = 0.066 ns and τ_L = 0.2 ns. The injected current is fixed at $I=50$ mA (the threshold current I_{th} = 11.5 mA) and the currier number at the transparency at $N_0=1.2\times10^8$. The parameter values are used for the calculated results that are shown in all figures of the paper.

2.2. Stationary states

In the subsequent analysis, we consider the stationary lasing states of system (1)-(2). They are given by rotating wave solutions, usually called External Cavity Modes (ECMs)

$$E(t)=E_Se^{i\omega_St},\ N=N_S. \tag{3}$$

By substituting (3) into (1)-(2), we obtain the transcendental equation for the emission frequency ω_S

$$\begin{aligned}F(\omega_S) = &-\omega_S+\alpha\gamma\left[-\cos(\varphi+\omega_S\tau_l)-\cos(\varphi+\psi+\omega_S(\tau_l+\tau_L)\right]- \\ &-\gamma\left[\sin(\varphi+\omega_S\tau_l)+\sin(\varphi+\psi+\omega_S(\tau_l+\tau_L))\right],\end{aligned} \tag{4}$$

ω_S is obtained from $F(\omega_S) = 0$. Finally, E_s and N_s can be obtained by substituting the value of ω_S into (1) and (2) and by setting roots to zero.

When only one cavity is present (COF), eq. (4) provides a finite number of solutions which are located on top of an ellipse in the N_S vs. ω_S plane. This elliptical locus for the solutions is independent of the feedback phase $\omega_S\tau$, being τ the external cavity round trip time. The feedback phase then determines the exact number of solutions as well as its precise location on the ellipse. If the feedback phase is changed, the location of the solutions moves around the ellipse.

In contrast to the COF case, the feedback from Fabry-Perot resonator implies a non elliptic location of modes [19]. For different phases φ, the location of the ECMs moves along an eight-shape figure, with the solitary laser mode located in the waist. The nature of the bifurcations and the stability of the solutions for the resonant feedback from FPR have been analyzed in more detail in [19].

The situation is again different for the case of LDCF. The solid lines in Fig. 3 show the locus of ECMs in the plane ($N_S - \omega_S$) for two different levels of feedback strength and three different values of one of the feedback phases, φ. The other feedback phase, ψ, determines then the exact number of solutions and its precise location of the solutions on the geometrical locus shown in Fig. 3. We first consider small feedback strength $\gamma = 5\ \mathrm{ns}^{-1}$ (left column in Fig. 3). For $\varphi = -\pi/2$, the location of the fixed points is similar to that of COF case, i.e., the modes are located on the ellipse although now the ellipse is distorted (see Fig. 3a left). Different symbols show the precise location of ECMs at particular feedback phases ψ. Note that the number of ECM depends on the values of the feedback phase. For $\psi = -\pi$ two modes and one anti-mode coexist (circles). On the other hand, only one mode is present for $\psi = 0$ (triangle) and $\psi = \pi/4$ (square). For $\varphi = \pi/6$ the system exhibits almost the tilted eight-shape (see Fig. 3b left) which is similar to that found in a Fabry–Perot resonator [18]. Although the level of feedback is still the same as before, the number of solutions has increased for all the values of the feedback phase ψ. Note also that the size of the locus for the ECMs is clearly larger than in the previous case. Finally, for $\varphi = \pi/2$ the tilted eight shape opens in the center leading to a "peanut" shape for the locus of the ECMs (see Fig. 3c left). The overall size of the locus, as well as the approximate number of solutions, is the same as for $\varphi = \pi/6$. This clearly illustrates that, in the case of the LDCF, the location of the modes becomes more complicated as compared with that of COF. When the feedback strength is increased to $\gamma = 15\ \mathrm{ns}^{-1}$ (right column in Fig. 3) new satellite bubbles of ECMs appear on left and right sides of the deformed ellipse. The different satellite bubbles span for a range of frequencies which is much larger than for the weak feedback case $\gamma = 5\ \mathrm{ns}^{-1}$. The onset of these bubbles reflects the existence of frequency gaps for which no ECM solutions exist. These frequency gaps are originated from destructive interference in the feedback coming from the two cavities. Figure 4a shows the total reflected light for the parameters corresponding to the right panel of Fig. 3a. Within the range of frequencies in which ECM solutions are found, there are two regions of vanishing reflected light which correspond to the two regions that separate the three bubbles shown in the right panel of Fig. 3a.

The interference of the feedback from the two cavities leads to an effective vanishing feedback, so there are no external cavity modes. Figure 4b shows the value of the roots of Eq. (4) for the parameters of the left panel of Fig. 3a and $\varphi = -\pi/2$. It can be seen as a fast oscillatory behavior on the top of a slower one. The fast and slow oscillatory periods are determined by the two feedback times. The intersection of this curve with the diagonal signals the

solutions of Eq. (4). As γ increases, the amplitude of the oscillations becomes larger and therefore more ECMs exist for any given value of the feedback phases. The feedback strength increase leads also to the emergence of additional isles of ECMs. The precise location of the bubbles and their shape depend on the feedback phase as shown in the right column of Fig. 3.

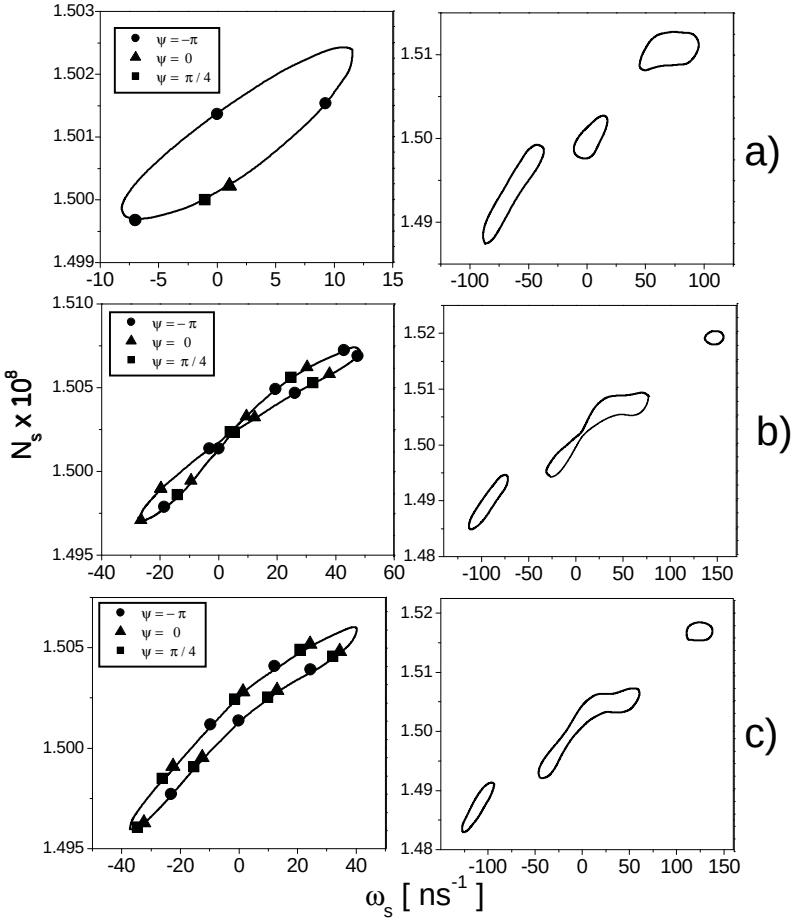


Fig. 3. Locus of the ECMs in the plane ($N_S - \omega_S$) for different phases: (a) $\varphi = -\pi/2$, (b) $\varphi = \pi/6$, and (c) $\varphi = \pi/2$ and two levels of feedback strength: $\gamma = 5\ \mathrm{ns}^{-1}$ (left) and $\gamma = 15\ \mathrm{ns}^{-1}$ (right). Symbols indicate the external cavity modes for a specific value of the feedback phase ψ.

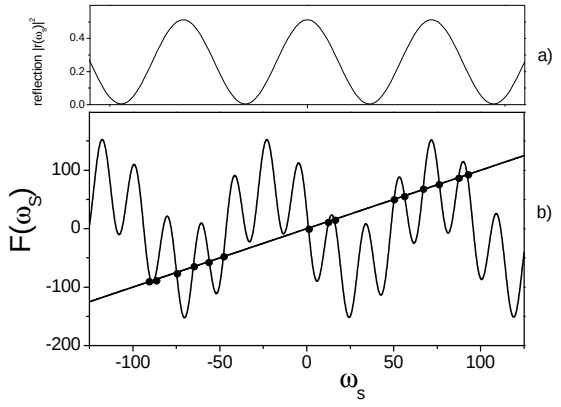


Fig. 4. Reflection spectrum of a double cavity with $\gamma = 15\ \mathrm{ns}^{-1}$ (a) and graphical solution of (4) (b) for $\gamma = 15\ \mathrm{ns}^{-1}$, $\varphi = -\pi/2$, and $\psi = \pi/4$. The circles indicate the modes.

2.3. Chaotic behavior of the LDCF transmitter laser

For feedback strength small enough, a laser under the influence of COF or LDCF shows either CW or pulsating operation. Chaotic behavior appears as the feedback strength is increased. Figure 5 illustrates typical time traces (left) and the power spectra (right) of a semiconductor laser under the influence of COF and LDCF for identical laser parameters in the chaotic regime. It can be observed that the LDCF makes the laser behavior more complex. This fact was further confirmed by calculation of the autocorrelation time from Eqs. (3) $T_c = \int_0^\infty d\tau\ \Gamma_{ii}^2(\tau)$ and (4) $\Gamma_{ij}(\tau) = \left\langle \left(P_i(t) - \langle P_i \rangle\right)\left(P_j(t-\tau) - \langle P_j \rangle\right)\right\rangle \Big/ \sqrt{\left\langle \left(P_i(t) - \langle P_i \rangle\right)^2 \right\rangle \left\langle \left(P_j(t) - \langle P_j \rangle\right)^2 \right\rangle}$ of [12]. These calculations yield to $T_c^{COF} \approx 0.1$ ns and $T_c^{DCF} \approx 0.037$ ns. Moreover, larger amplitude fluctuations when compared with COF can be observed.

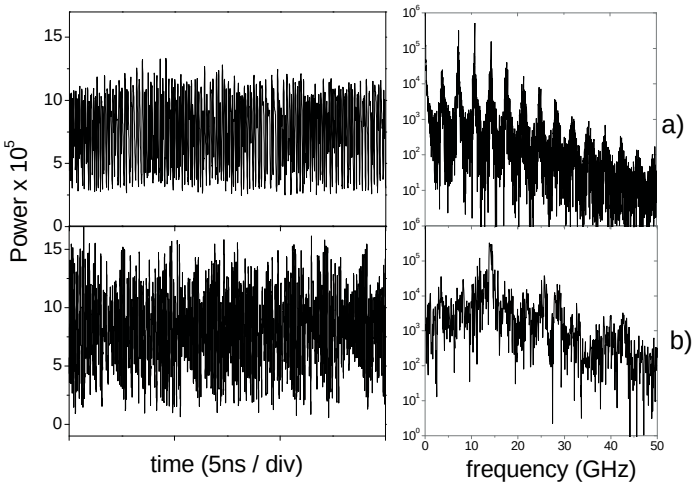


Fig. 5. A typical optical power time trace (left) and the power spectrum (right) of a semiconductor laser under the influence of (a) COF for $\gamma = 30\ \text{ns}^{-1}$, $\varphi = -\pi/2$, $\tau = 0.266$ ns and (b) DCF for $\gamma_1 = \gamma_2 = 30\ \text{ns}^{-1}$, $\varphi = -\pi/2$, $\psi = \pi$, $\tau = \tau_l + \tau_L = 0.266$ ns, $\tau_l = 0.066$ ns, $\tau_L = 0.2$ ns.

Figure 6 displays the bifurcation diagrams of the semiconductor laser subject to LDCF for two feedback phases. As the feedback strength is increased, several instabilities take place. For a given value of the feedback strength, the figure displays the values of all the local maxima of the time traces of the emitted power. Considering $\varphi = -\pi/2$ and $\psi = \pi$, for low values of the feedback strength, CW operation is observed, which is depicted as a single value for the maxima of the power in Fig. 6a. At the feedback strength $\gamma = 10\ \text{ns}^{-1}$ a Hopf bifurcation appears and the output power develops an oscillatory behavior. Since the oscillations are periodic, for a given feedback strength, all the local maxima of the output power have the same value and consequently a single point appears in Fig. 6. The Hopf bifurcation is supercritical and, as expected, the oscillation amplitude grows with the square root of the distance from the bifurcation point. As the feedback strength is further increased, a scenario compatible with quasi-periodic route to chaos is obtained. However, the range and amplitude of this behavior are small. When the feedback strength reaches the value $\gamma = 15\ \text{ns}^{-1}$, a jump to a new *P* periodic operation region is observed. As the feedback strength increases, a second scenario compatible with quasi-periodic route to chaos appears. For large values of the feedback strength, the system exhibits a chaotic behavior.

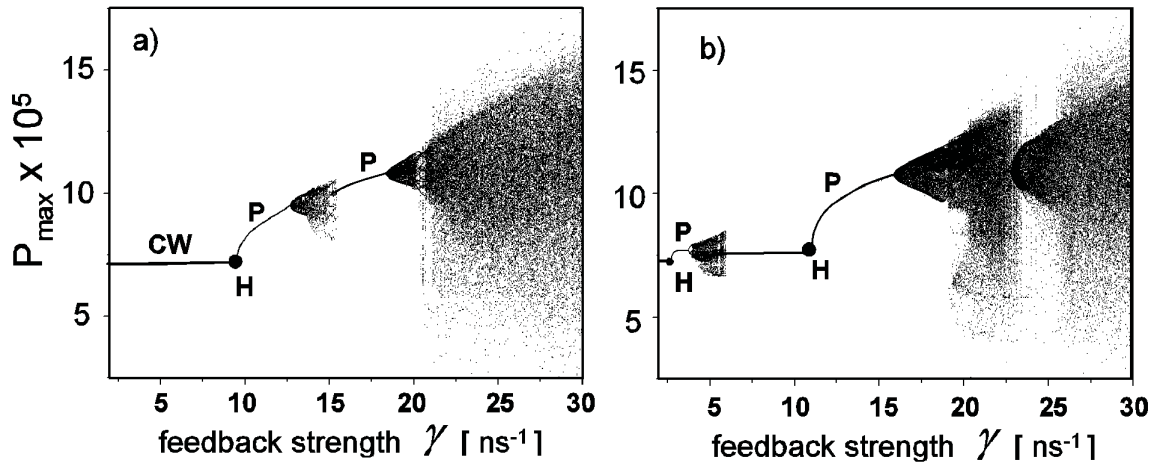


Fig. 6. Numerical bifurcation diagrams for different values of phases: (a) $\varphi = -\pi/2,\ \psi = \pi$ and (b) $\varphi = \pi/2,\ \psi = \pi/4$. CW shows the continuous-wave operation, the circle H indicates Hopf bifurcation, and P shows the peak of stable periodic solution.

For $\varphi = \pi/2$ and $\psi = \pi/4$ (see Fig. 6b) the system behavior is slightly different; the Hopf bifurcation is shifted to a lower feedback level involving the appearance of low amplitude chaotic behavior for low feedback strengths followed by the CW operation and a scenario compatible with quasi-periodic route to chaos. We mention that the numerical calculations show that, in this parameter region and for any value of the feedback strength larger than $25\ \text{ns}^{-1}$ and combination of phases φ and ψ, the laser behavior is chaotic and robust.

2.4. Synchronization and mismatch in the laser parameters

In the previous section, we have clarified different aspects of the dynamics of a semiconductor laser with integrated DCF for obtaining chaotic behaviors. In what follows, we are interested in the transmitter–receiver configuration and in the evaluation of their synchronization properties. Synchronization can be quantified by measuring the cross correlation coefficient:

$$\left[C = \langle P_m(t)P_s(t)\rangle / \left(\langle |P_m(t)|\rangle \langle |P_m(t)|\rangle\right)\right]. \tag{5}$$

Figure 7 shows the emitted power of slave system versus the power of the master (synchronization diagram) for feedback strength $\gamma = 30\ \text{ns}^{-1}$ and different levels of the coupling parameter k. We first consider the case of identical Master and Slave systems, so we take the same parameter values for both of them. When the coupling parameter is equal to zero, as shown in Fig. 7a, the trajectories of the master and slave lasers depart from each other and the synchronization map is a cloud of points showing the lack of correlation between outputs. Upon increasing the coupling until $100\ \text{ns}^{-1}$, the synchronization improves and cross correlation coefficient increases approaching one (see Fig. 7c).

Figure 8 shows the dependence of the synchronization quality as a function of the feedback phases. It displays the value of the correlation function in the parameter space $(\varphi - \psi)$ for feedback strength $\gamma = 30\ \text{ns}^{-1}$ and coupling coefficient $\kappa = 75\ \text{ns}^{-1}$. It can be clearly seen that the region of high correlation coefficients is wide while regions of low correlation hardly appear. The white star in Fig. 8 corresponds to the operating point that will be considered for message encoding and decoding in the next section.

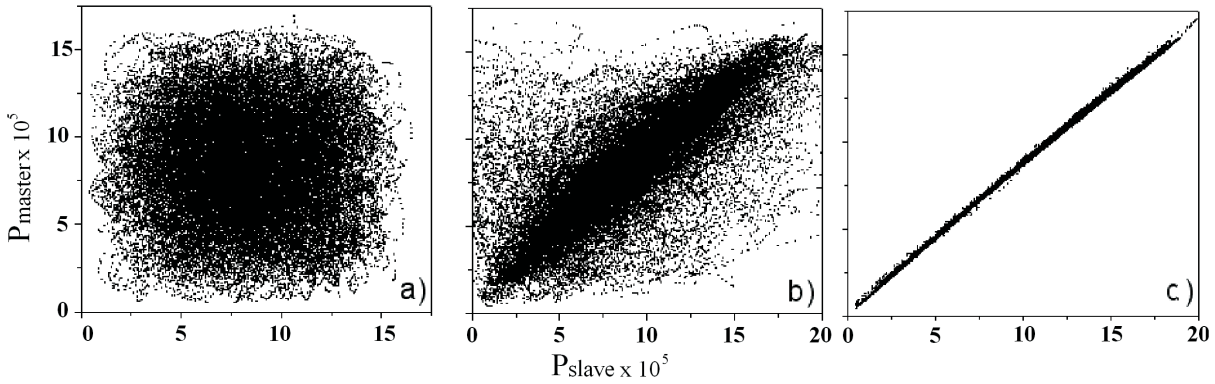


Fig. 7. Synchronization diagrams for different levels of the coupling parameter k: (a) $k = 0\ \mathrm{ns}^{-1}$ (the systems are uncorrelated), (b) $k = 50\ \mathrm{ns}^{-1}$ (C=0.75), and (c) $k = 100\ \mathrm{ns}^{-1}$ (C=0.9995). The feedback strength is taken as $\gamma = 30\ \mathrm{ns}^{-1}$ and the feedback phases are $\varphi = -\pi/2$ and $\psi = \pi$.

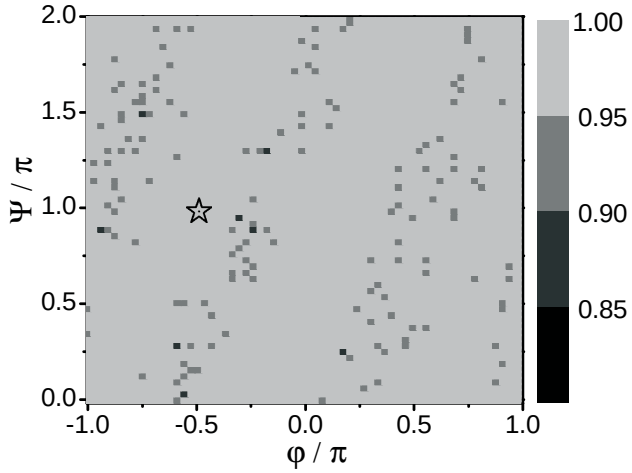


Fig. 8. Cross correlation coefficient as a function of the feedback phases for $\gamma = 30\ \mathrm{ns}^{-1}$ and $\kappa = 75\ \mathrm{ns}^{-1}$. The white star is the operating point for the message encoding and decoding.

It is well known that the quality of the synchronization depends on the similarity between master and slave lasers. The influence of the internal laser parameters mismatch on the synchronization quality has been studied in [22, 28], so here we focus on the influence of the mismatch on the two feedback phases. Figure 9 shows the dependence of cross correlation coefficient on the phase difference (phase master–phase slave) for feedback strength $\gamma = 30\ \mathrm{ns}^{-1}$ and coupling strength $k = 75\ \mathrm{ns}^{-1}$ (a) and $k = 100\ \mathrm{ns}^{-1}$ (b). The solid line shows the degradation of the synchronization due to a mismatch in the material cavity feedback phase. We take $\varphi_m = \varphi_S = \pi$ and $\psi_s = 0$ while ψ_m is varied from 0 to π. The dotted line shows the effect of a mismatch in the air cavity feedback phase. We consider $\psi_s = \psi_m = \pi$ and $\varphi_S = 0$ while φ_m is varied from 0 to π. When the feedback phases coincide, the system exhibits perfect synchronization with C ~ 1 cross-correlation coefficient. An increase of the mismatch in any of the feedback phases induces a degradation of the synchronization which is indicated by a reduction of the cross correlation coefficient. For small mismatch, the degradation of the correlation is similar for the mismatch in any of the two phases.

As the mismatch is increased, the degradation is clearly more severe in the case of mismatch in the feedback phase of the air cavity φ.

This may be understood from the fact that φ is the phase of a shorter cavity and, in general, short cavities are more sensitive to phase variations than long cavities. For larger values of the coupling strength, the effect of the mismatch in the feedback phases is smaller and therefore the cross-correlation coefficient decreases slower as the mismatch is increased. High values of the cross-correlation coefficient are usually required for efficient message encoding and decoding, therefore, the mismatch in the feedback phases should not exceed $0.1/\pi$. Alternatively, a small airgap phase mismatch can be compensated adjusting the material phase ψ of the slave laser. Figure 9c shows an example of this compensation. The mismatch in the airgap phase is 2.5% (point *B* in Fig. 9a), and a very good correlation can be achieved (point *C*).

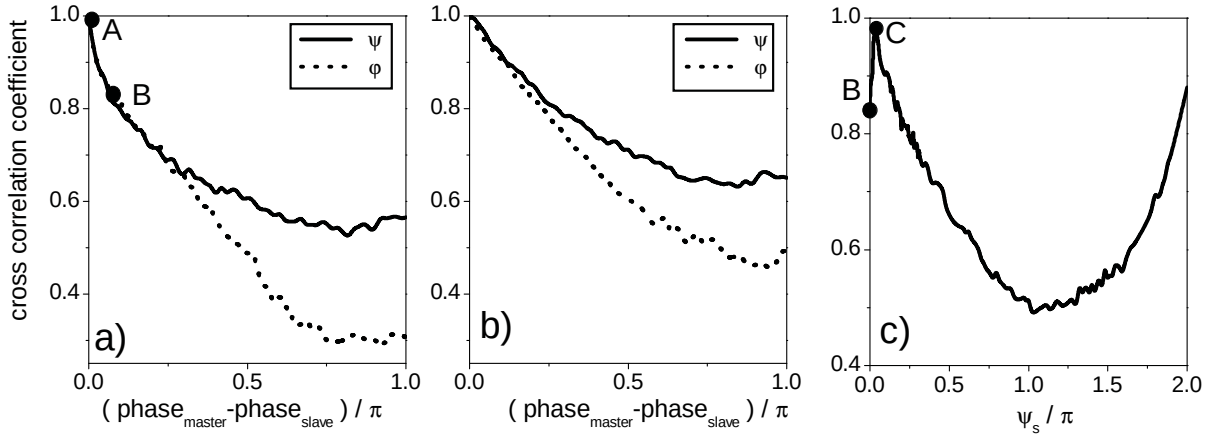


Fig. 9. Cross correlation coefficient as a function of the feedback phase difference (phase master–phase slave) for coupling $k = 75\ \text{ns}^{-1}$ (a), to $k = 100\ \text{ns}^{-1}$ (b). Panel (c) shows the cross correlation coefficient for $\varphi_m = \varphi_S = \pi$, and $\psi_m = 0$ as a function of the material phase for the slave laser ψ_s. For both master and slave systems, the feedback strength was taken as $\gamma = 30\ \text{ns}^{-1}$ (from [23]).

2.5. Message transmission

In this section, we consider the use of these integrated devices for message encoding and decoding in chaos based communications. A message is encoded as a small amplitude modulation of the emitted field of the master, so that the signal transmitted to the receiver is

$$E_T = E_t(1+\varsigma m(t)), \tag{6}$$

where $m(t)$ is the message and ς is the message amplitude. In the receiver system, a message is decoded comparing the input of the receiver with its output, which is ideally synchronized to the carrier

$$M_{decoded} = \sqrt{P_T / P_S} - 1. \tag{7}$$

Figure 10 illustrates the transmission of a non return to zero pseudorandom message. The system parameters correspond to the operating point shown by the white star in Fig. 8. Panel (a) shows the input message. Panels (b) and (c) show the chaotic carrier without the message and the transmitted signal (carrier with message). Panel (d) shows the decoded message as indicated in Eq. (7), and filtered by an appropriate low-pass filter [29]. As can be seen from the figure, the message is well recovered. Panel (e) shows the recovered message for a 2.5% mis-

match in the airgap phases between master and slave lasers. It can be clearly seen that a part of message cannot be recovered. On the other hand the airgap phase mismatch can be compensated by the controllable phase of slave laser. As can be seen in panel (f), the message is now well recovered.

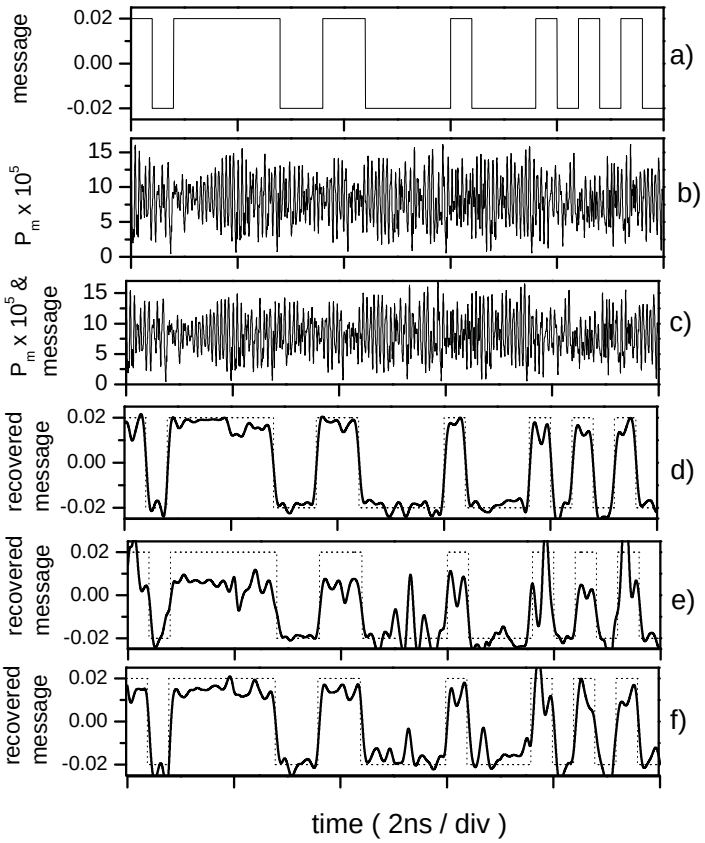


Fig. 10. Numerical results of encoding and decoding of a 2.5 Gbit/s digital message for a closed loop scheme [23]: (a) encoded message, (b) output of master laser, (c) output of master laser with message, (d) recovered message after filtering (solid line) and input message (dotted line) for identical parameters for master and slave lasers (point *A* in Fig. 9a), (e) recovered message (solid line) after filtering for lasers with 2.5% mismatch of airgap phases (point *B* in Fig. 9a), (f) recovered message (solid line) after filtering for lasers with 2.5% mismatch of airgap phases but the slave air gap phase is compensated by a different ψ_s phase (point *C* in Fig. 9c). Parameters $\gamma = 30\ \mathrm{ns}^{-1}$, $\kappa = 75\ \mathrm{ns}^{-1}$, $\varphi = -\pi/2$, $\psi = \pi$, $\tau_l = 0.066$ ns, $\tau_L = 0.2$ ns.

It is noteworthy that these simple examples of the chaos modulation encoding technique, within a chaotic waveform, obtained from a double cavity feedback are efficient and simple and could be easily applied to a chaos-based communication system by using an external modulation.

3. T-like double cavity optical feedback. On/off phase shift keying method

One of the most attractive schemes in terms of security is the OOPSK method [21, 30] where the codification is achieved by slight modulating the phase of the optical feedback of the emitter (see Fig. 1). The physical basis for OOPSK is that the synchronization behavior of the receiver acts as a sensitive detector for variations of the transmitter feedback phase: suitable discrete changes yield the dynamics of the receiver to jump between synchronized and desynchronized states. In contrast to these drastic changes in the receiver dynamics, changes in the emitter dynamics should be noticeable neither in the intensity dynamics nor in the RF nor optical spectra. The principle of the ON/OFF phase shift keying encryption is as follows. The message is encoded by switching between two states of the master system that yield highly correlated (synchronized) states (Bit "0") or less correlated (desynchronized) states (Bit "1") in the receiver system. Hence, the message can be simply recovered by monitoring the synchronization error. The controlled variations in the master system can be accomplished by inserting, e.g., an electrooptical modulator within the external cavity of the transmitter. The message is decoded by detecting whether the receiver synchronizes or not with the input carrier [26]. Up to now, this technique has the disadvantage that the maximum modulation rate is only a few tens of Mbit/s [21, 30, 31].

3.1. Laser and feedback model

The proposed setup for implementation of OOPSK method is depicted schematically in Fig. 11. It consists of a semiconductor laser coupled to the external reflectors R_1 and R_2, that could be implemented, e.g., by using two fiber cavities. The advantage of the proposed scheme is that we can control two feedback strengths, two feedback phases, and two delay times independently. The feedback branch governed by reflectivity R_1 is called conventional feedback branch (CFB); the one governed by R_2, the modulated feedback branch (MFB). Assuming fiber based cavities with a refractive index of 1.5, we consider the delay time in the CFB to be $\tau_l = 0.5$ ns and that of MFB $\tau_L = 0.3$ ns. In the model, we only account for single reflexions in both branches. In the absence of coupling, the correlation between the transmitter and receiver outputs is negligible. When a certain amount of light from the transmitter is injected into the receiver, the latter is able to synchronize to the emitter under appropriate conditions. Once synchronized, a message can be encoded into the carrier. At the receiver side, the message can be recovered via the chaos pass filtering process [29].

The laser dynamics is analyzed in the framework of the extended Lang-Kobayashi equations well explained in Section 2 with the following feedback term

$$F_{TDCF} = \gamma_1 e^{i\varphi} E_{t,r}(t-\tau_1) + \gamma_2 e^{i\psi} E_{t,r}(t-\tau_2). \quad (8)$$

The cavity roundtrip time of the light within the laser is ($\tau_c = 8.5$ ps), η_{ext} accounts for losses different than those introduced by the laser facet ($\eta_{ext} = 0.5$) resulting in $\kappa = 90\ \text{ns}^{-1}$; τ_1 and τ_2 are roundtrip time in the CFB and MFB, respectively; γ_1 and γ_2 are the feedback strengths governed by the reflectivities R_1 and R_2, respectively; $\varphi = \omega_0 \tau_l$ and $\psi = \omega_0 \tau_L$ are the accumulated optical phases in the CFB and MFB, respectively, which, without loss of generality, can be assumed to take values between 0 and 2π. The other parameter values are: the linewidth enhancement factor $\alpha = 5$, the differential gain parameter $g = 1.5\times10^{-5}\ \text{ns}^{-1}$, the gain saturation coefficient $s = 4\times10^{-7}$, the photon and carrier lifetimes $\tau_{ph} = 2$ ps and $\tau_e = 2.0$ ns, respectively, and the carrier number at the transparency $N_0=1.5x10^8$. These parameters, which are considered identical for both lasers, are used for the calculated results shown in all figures in the paper. The injection current is fixed at $I = 45$ mA (I_{th}=14.7 mA). For the model given by equations (8) and (9), if $\tau_1 = \tau_2$ the feedback term in (8) can be reduced to a COF term with an equivalent feedback coefficient given by $\tilde{\gamma} e^{i\tilde{\varphi}} = \gamma_1 e^{i\varphi} + \gamma_2 e^{i\psi}$.

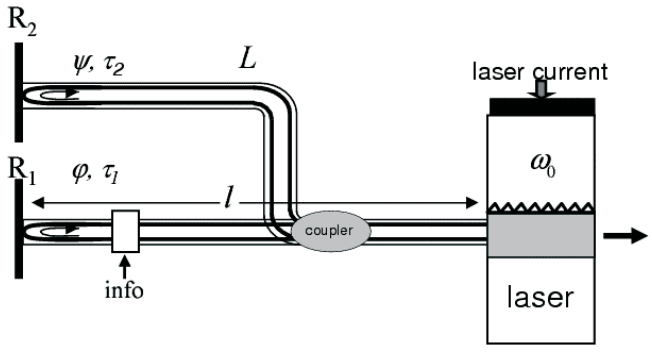


Fig. 11. Setup under study: a laser with fiber-based external cavities. The cavity lengths are l = 0.05 m (τ_1 = 0.5 ns) and L = 0.03 m (τ_2 = 0.3 ns). The refractive index of the optical fiber is n = 1.5.

3.2. TDCF transmitter laser dynamics

In this section, we discuss the behavior of a semiconductor laser under the influence of a TDCF. For feedback strengths small enough, semiconductor lasers under the influence of either COF or TDCF show CW or pulsating operations. Chaotic behavior appears if the feedback strength is increased enough. Figure 12b illustrates typical time traces (left) and the power spectra (right) of a laser under the influence of a DCF operating in a robust chaotic regime. We mention that the behavior shown in Fig. 12a is similar to that of a laser under the influence of COF with $\gamma_1 = 40\ \text{ns}^{-1}$, $\tau = 0.5\ \text{ns}$ and identical laser parameters. It is well known that the autocorrelation time is related to the complexity of the generated chaos. The correlation time accounts for the complexity of the generated chaos. The shorter is the correlation time, the more chaotic and less predictable is the dynamics. The calculations of the autocorrelation time for the traces shown in Figs. 12a and 12b [12] yield similar result for both COF and TDCF with values of $T_{ac}^{COF} \sim T_{ac}^{TDCF} \sim 100$ ps for our parameter values.

A confirmation of this property is given below. Figure 13a shows the autocorrelation time as a function of feedback strength for COF for $\tau = 0.5\ \text{ns}$ (solid line) and $\tau = 2\ \text{ns}$ (dotted line). It can be clearly seen that, as the feedback strength and delay time are increased, the autocorrelation time decreases; this is an indication that the laser dynamics becomes more chaotic. Figure 13b shows the calculated autocorrelation time for a laser under the influence of a TDCF. The feedback strength of CFB is fixed to $\gamma_1 = 30\ \text{ns}^{-1}$ while that of MFB is varied. For zero MFB strength, the resynchronization and autocorrelation times coincide with that of COF for $\gamma = 30\ \text{ns}^{-1}$. An increase in feedback strength of MFB leads to a decrease in autocorrelation time up to 0.1 ns similar to that of COF. However, when the MFB is introduced, the resynchronization time can be expected to become much shorter as a result of only distortion of the chaotic attractor generated by the CFB.

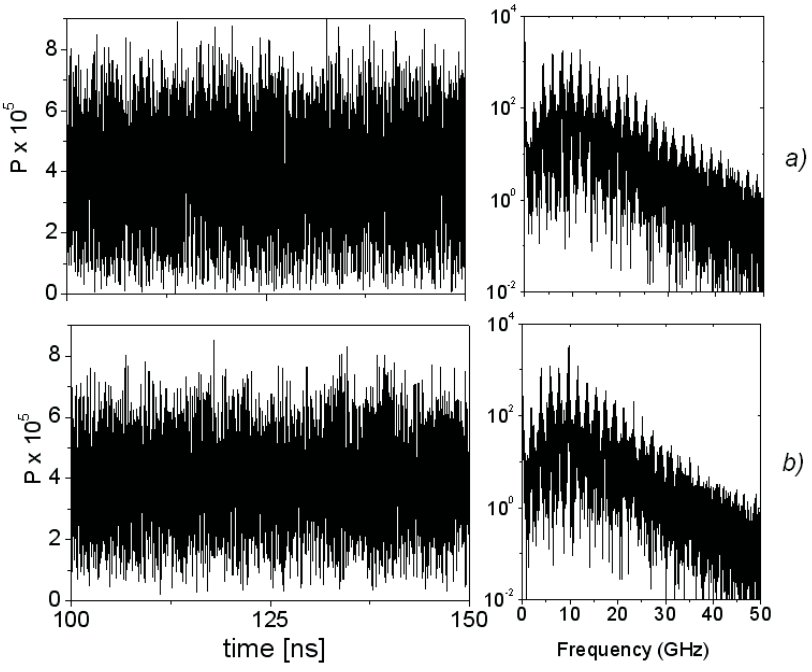


Fig. 12. Time traces of the output power P (left) and the power spectrum (right) for (a) COF at $\gamma = 40\ \text{ns}^{-1}$, $\tau = 0.5\ \text{ns}$, and $\varphi = 0$ and for (b) TDCF at $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$, $\tau_1 = 0.5\ \text{ns}$, $\tau_2 = 0.3\ \text{ns}$, $\varphi = 0$, and $\psi = \pi/2$.

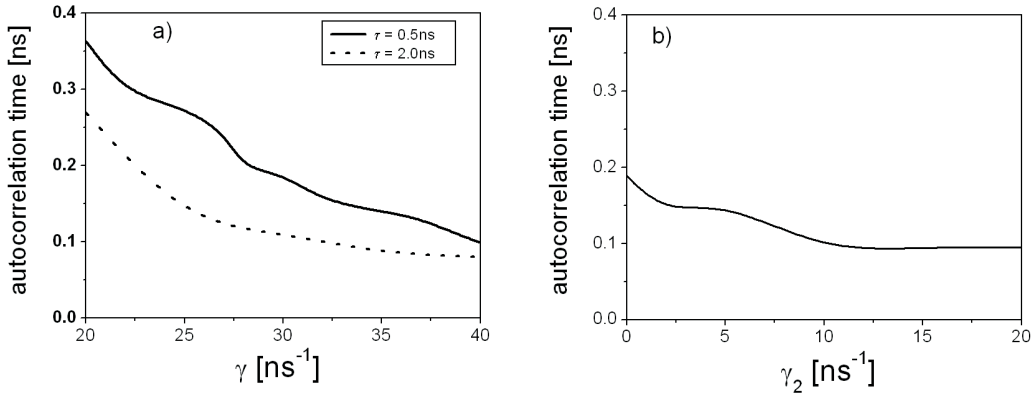


Fig. 13. The autocorrelation time as a function of feedback strength for (a) COF at $\varphi = 0$ and different values of delay time and (b) TDCF of MFB at $\gamma_1 = 30\ \text{ns}^{-1}$, $\varphi = 0$, $\psi = \pi/2$, $\tau_1 = 0.5$ ns, $\tau_2 = 0.3$ ns.

Figure 14a displays a typical bifurcation diagram of a semiconductor laser under the influence of COF with the feedback strength acting as a bifurcation parameter and $\tau = 0.5$ ns, $\varphi = 0$. As the feedback strength is increased, several bifurcations take place. For each value of the feedback strength, the figure displays the values of the maxima of the time traces of the emitted power. It is well known that, as the feedback strength is increased, a scenario compatible with quasi-periodic route to chaos appears [32].

Figures 14b and 14c display the bifurcation diagrams of a semiconductor laser subject to DCF for the feedback strength and feedback phase acting as bifurcation parameters. Let us consider, e.g., the case of feedback strength for the CFB fixed to $\gamma_1 = 30\ \text{ns}^{-1}$ while the feedback strength of MFB is increased. Considering $\varphi = 0$ and $\psi = \pi/2$, as shown in Fig. 14b, even for low values of the feedback strength γ_2 the dynamics of the laser is already chaotic due to the influence of the feedback of CFB. It can be noticed from the figure that the amplitude of the chaotic oscillations slightly increases with the feedback strength γ_2 (see Fig. 14b). When both feedback strengths are fixed to $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$ and the phase $\varphi = 0$, as shown in Fig. 14c, fully developed chaotic dynamics is found for any value of MFB phase ψ.

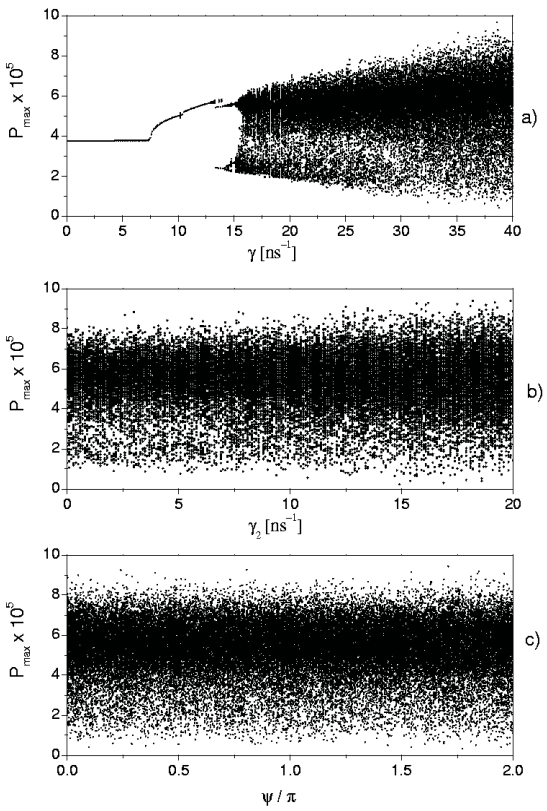


Fig. 14. Bifurcation diagram of the output power for(a) COF with the feedback strength γ as bifurcation parameter; (b) TDCF with γ_2 as bifurcation parameter; $\gamma_1 = 30\ \text{ns}^{-1}$, $\varphi = 0$ and $\psi = \pi/2$; (c) DCF with the MFB phase ψ as bifurcation parameter; $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$ and $\varphi = 0$. Each dot represents a peak of the output power [24].

3.3. Synchronization and message transmission

So far we have clarified different aspects of the transmitter laser dynamics under a TDCF. In what follows, we focus on the transmitter–receiver configuration and evaluate the synchronization properties. Since our final aim is to use the auxiliary branch to perform OOPSK encryption, it is important to characterize in advance the resynchronization time, i.e., the time required by the setup to synchronize when the link between master and slave lasers is interrupted. The inverse of the resynchronization time is an estimation of the maximum modulation rate that can be achieved with the OOPSK technique. We estimate the resynchronization time as the time needed by the system to achieve a correlation coefficient of 0.98 when starting from an initial uncoupled configuration, for which the correlation between emitter and receiver is close to zero [24]. Figure 15 shows the resynchronization time as a function of feedback strength for a laser under the influence of COF for different values of delay time τ and different coupling coefficient κ in a region where the system displays a chaotic behavior.

As shown in Fig. 15a, two regimes are observed for delay time $\tau = 0.5$ ns. For $\gamma > \kappa_r$ the resynchronization time grows linearly with a small slope and it is on the order of a few round-trip time. When γ becomes larger the resynchronization time grows exponentially. This last regime appears for $\gamma > 27\ \text{ns}^{-1}$ when $\kappa_r = 60\ \text{ns}^{-1}$ and for $\gamma > 33\ \text{ns}^{-1}$ when $\kappa_r = 90\ \text{ns}^{-1}$. The separation in the regimes comes from a competition between the coupling strength κ_r and the feedback strength γ. For $\kappa_r >> \gamma$ the first regime is dominant leading to a small resynchronization time; when both are similar, there is a transient competition, which induces large resynchronization time. Finally, no synchronization is observed for $\gamma >> \kappa_r$. Similar behavior is observed for a large delay time $\tau = 2$ ns (see Fig. 15b), although now the resynchronization time in the first regime grows linearly with a larger slope.

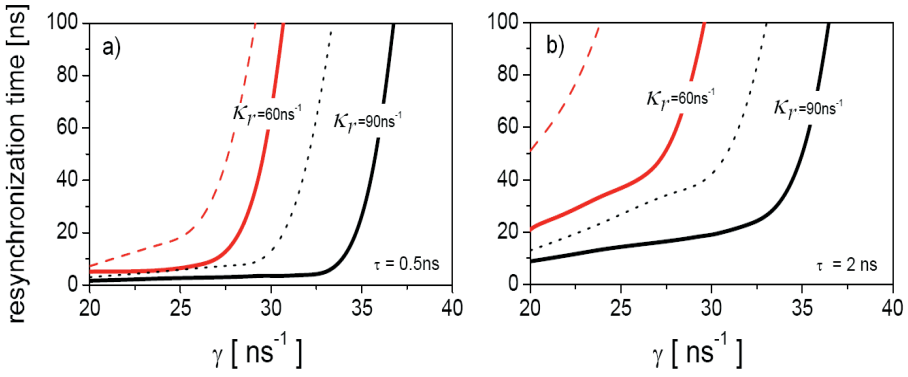


Fig. 15. Resynchronization time as a function of feedback strength for COF for $\varphi = 0$ and the delay time $\tau = 0.5$ ns (a) and $\tau = 2$ ns (b). Results are obtained from 100 random initial conditions. Thick lines show the average resynchronization time, while thin lines show the maximum value of those 100 realizations.

Figure 16a shows the calculated resynchronization time for a laser under the influence of a TDCF. The feedback strength of CFB is fixed to $\gamma_1 = 30\ \text{ns}^{-1}$, while that of MFB is varied. For zero MFB feedback strength the resynchronization time coincides with that of COF for $\gamma = 30\ \text{ns}^{-1}$. An increase in the MFB feedback strength up to $\gamma_2 = 20\ \text{ns}^{-1}$ leads only to a smooth increase in the resynchronization time up to a value of approximately 7 ns.

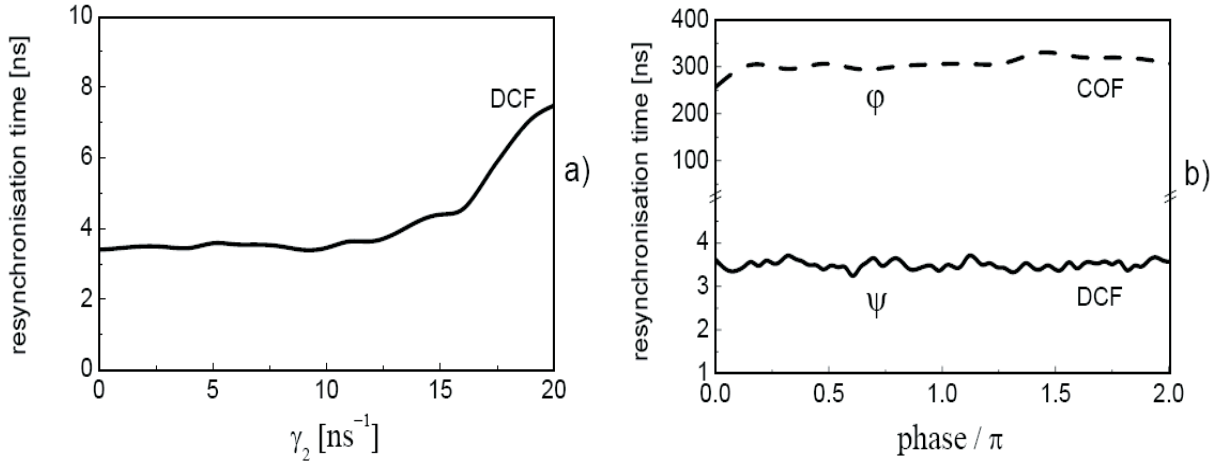


Fig. 16. Resynchronization recovery time as a function of different parameters (a) the feedback strength γ_2 of MFB for TDCF for $\gamma_1 = 30\ \text{ns}^{-1}$, $\varphi = 0$, $\psi = \pi/2$, (b) the feedback phase for the COF (dashed line) and for the TDCF (solid line). The parameters for the COF are $\gamma = 40\ \text{ns}^{-1}$ and $\tau = 0.5\ \text{ns}$. The parameters for the DCF are $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$, $\varphi = 0$, $\tau_1 = 0.5\ \text{ns}$, $\tau_2 = 0.3\ \text{ns}$, $\kappa_r = 90\ \text{ns}^{-1}$.

A first comparison of Figs. 15 and 16 shows that TDCF has some advantages over COF. When the feedback strength of COF is $40\ \text{ns}^{-1}$, the resynchronization time is approximately 300 ns, while for the DCF with $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$ the resynchronization time is approximately 3 ns (see Fig. 16a). In fact, we have found that, when using the DCF setup, we can reduce the resynchronization time by two orders of magnitude as compared with the COF setup. This decrease in the resynchronization time can be attributed to the fact that, for low feedback strength values, the MFB acts as a weak perturbation of the strong chaotic attractor generated by the CFB. This fact yields shorter resynchronization times as compared to the COF case. The solid line in Fig. 16b shows the resynchronization time as a function of MFB phase ψ for TDCF, for $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$, $\psi_m = \psi_S = 0$. The dashed line shows the resynchronization time as a function of optical feedback phase of the COF case for $\gamma = 40\ \text{ns}^{-1}$, $\varphi = 0$. It can be clearly seen from these results that the TDCF system resynchronizes much faster than the COF system for any value of the feedback phase. We have checked that these results also hold for any value of the COF phase ψ.

Now we consider the influence of a mismatch between the phases ψ of the slave laser with respect to that of the master laser on the cross correlation coefficient. Figure 17 shows the values of this coefficient in the plane $(\psi_s - \psi_m)$ for feedback strengths $\gamma_1 = 30\ \text{ns}^{-1}$, $\gamma_2 = 10\ \text{ns}^{-1}$ and the coupling coefficient $\kappa_r = 90\ \text{ns}^{-1}$.

Other parameters are identical for the master and slave lasers. It can be clearly seen that highest correlation coefficients are obtained when the two phases coincide, i.e., $\psi_m = \psi_s$, while the correlation degrades when the phases start to be different. Points *A* and *B* in Fig. 17 correspond to the operating points that will be considered later for message encoding and decoding using OOPSK encryption. The point *A* is chosen to have high correlation while the point *B* (or *B'*) corresponds to a state with low correlation.

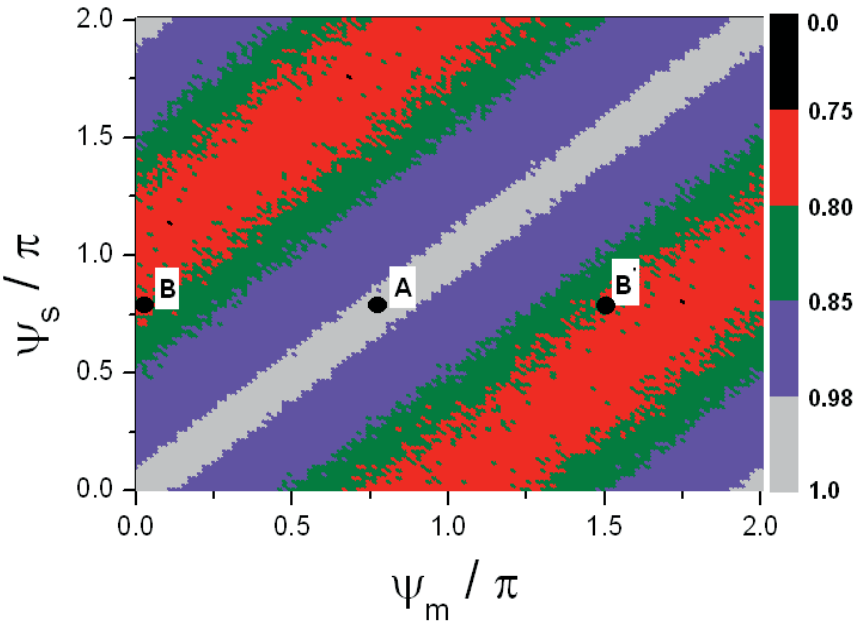


Fig. 17. Cross correlation coefficient in the ψ_m - ψ_s phase space. The other parameters are $\gamma_1 = 30\ \mathrm{ns}^{-1}$, $\gamma_2 = 10\ \mathrm{ns}^{-1}$, $\kappa_r = 90\ \mathrm{ns}^{-1}$, $\varphi_m = \varphi_S = 0$. A high degree of synchronization is characterized by light grey level. Phases are varied in 0.05 radian steps (taken from [26]).

3.4. Message transmission

An important issue in chaos based communication systems is the security of proposed setup. Schemes, such as chaos shift keying, chaos masking, and chaos modulation, require keeping the message amplitude small enough in order to avoid message recognition. In the OOPSK technique, the message is codified by changing the feedback phase of the master laser without introducing significant changes in the time trace or spectrum of the emitted light. In this setup, the slave laser for which the feedback phase is kept constant acts as a detector of the synchronization quality. When the feedback phases of the emitter and receiver coincide, the correlation between the outputs of the two systems is high, while it is low when the phases are different (as shown in Fig. 17). In our scheme, the phases φ of the COF branch is kept constant in both master and slave lasers while we study the phase shift keying method by varying the phase of the auxiliary branch. Figure 18 shows the pulse traces of the master laser operating in the chaotic regime at the point *A* (a) and at point *B* (b). It can be clearly seen that both time traces remain similar to each other. Figure 18c shows the power spectra of the time traces shown in Figs. 18a and 18b, while Fig. 18d shows the power spectra of the receiver system for the fixed phase $\psi_S = 0.75$ rad.

The power spectra of the emitter system for the operating points *A* and *B*, as shown in Fig. 18c, remain almost unchanged. On the contrary, the power spectrum of receiver laser changes, as shown in Fig. 18d, due to the coupling light that is uncorrelated with that generated by the receiver system. Figures 8e and 8f show the synchronization error defined as $\left|(P_m - P_s)/(P_m + P_s)\right|$ for different phases. For $\psi_m = \psi_s = 0.75$ rad (see Fig. 18e) the synchronization error is almost zero and the cross correlation coefficient approaches unity. On the other hand, for $\psi_m = 0$, $\psi_S = 0.75$ rad synchronization degrades, as shown in Fig. 18f, and the synchronization error is very high.

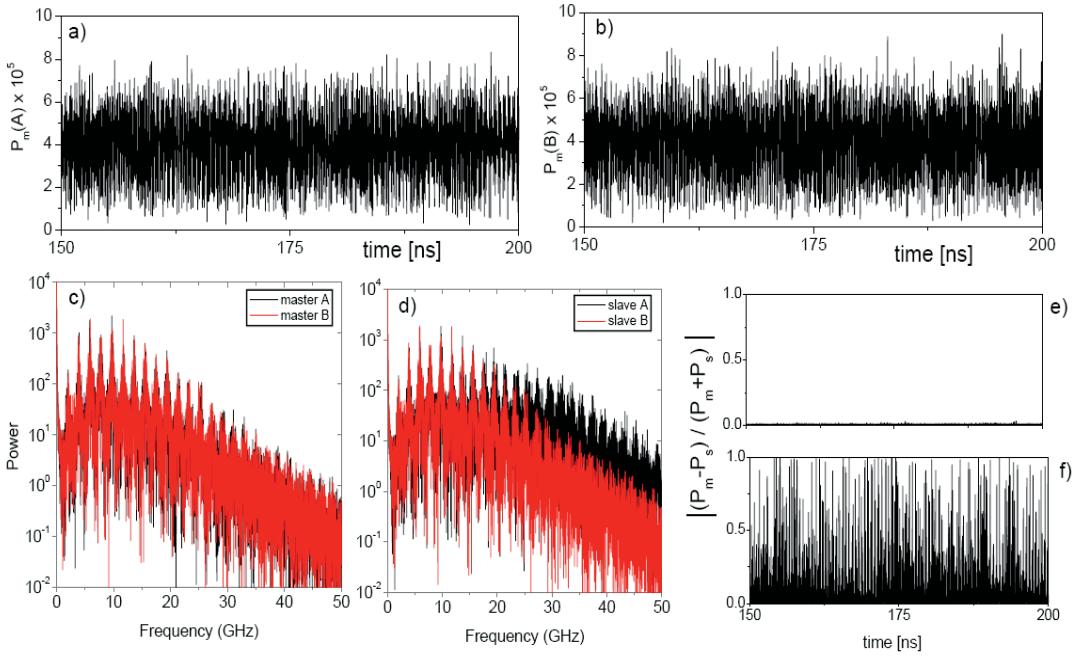


Fig. 18. (Color online) Calculated pulse traces of the emitter laser at point *A* (a) and point *B* (b) shown in Fig. 7. Power spectra of the master (c) and slave (d) lasers. Panels (e) and (f) show the synchronization error for $\psi_m = \psi_s = 0.75$ rad and $\psi_m = 0$, $\psi_S = 0.75$ rad, respectively. Parameters are the same as in Fig. 17.

Figure 19 depicts the process of 0.25-Gbit/s message OOPSK encryption. The top panel shows the digital message. Figure 19 (central panel) shows the synchronization error when the phase of the receiver laser is changed from 0.75 rad (bit "0") to 0 (bit "1"), i.e., from point *A* to point *B* in Fig. 17. Figure 19 (bottom panel) shows that the message can be successfully recovered after a standard filtering process. Thus, the proposed setup can distinctly increase the bit rate compared with that previously obtained in [21].

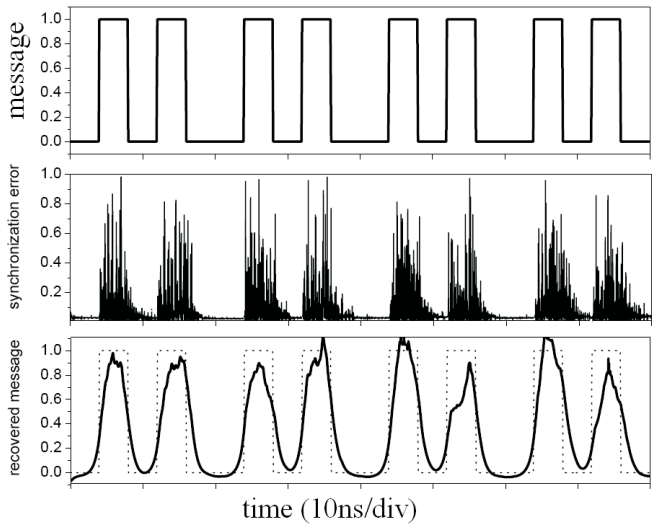


Fig. 19. On/off phase shift keying encoding and decoding of 0.25-Gb/s digital message. Top panel: encoded message. Central panel: decoded message represented by the synchronization error. Bottom panel: recovered message after filtering. The other parameters are the same as in Fig. 17 and [24].

4. Conclusions

In this review, we have studied the dynamics of an integrated device composed of a semiconductor laser and a longitudinal double cavity that provides delayed optical feedback. The double cavity feedback implies the existence of two feedback phases, which can play an important role in the dynamics. This extra degree of freedom leads to a more complex behavior, which in fact is already indicated when looking for the number and location of the fixed points. While in the more conventional case of using a single cavity, these steady states are located on the top of an ellipse in the $(N_S - \omega_S)$ plane, the ellipse in the double cavity case can be strongly distorted and can break into several bubbles. The number of coexisting steady states is also increased in the case of double cavity feedback, and chaotic behavior is found for lower values of the feedback strength. Furthermore, chaos appears already for quite short cavities, allowing for compact devices. We have shown that two of these devices can be synchronized when operating in the chaotic regime in a master-slave configuration. However, synchronization is degraded when there is a mismatch in the parameters of the master and slave system. Since the novelty of this scheme is the existence of two feedback phases, we have discussed the effect of a mismatch in these phases in detail. A mismatch in the air gap feedback phase turns out to have stronger effects in the master-slave cross-correlation than a mismatch in the material cavity feedback phase. However, for mismatches small enough, good quality synchronization is obtained. For the parameter values where good synchronization is achieved, it is possible to encode a message in the carrier using the chaos modulation technique. The message can be appropriately recovered at the receiver even for high bit rates. The codification method we employed is just an example of what can be done. While this codification technique is efficient and simple to implement, other codification methods could be used as well.

We have also studied the dynamics of a device composed of a semiconductor laser subject to a TDCF. The main advantages of this scheme over that of single cavity include the existence of two feedback strengths, two feedback phases, and two delay times that can be controlled separately. The results presented in this review show the following features: under appropriate conditions, this setup is capable of generating a robust chaotic waveform; two of these devices can be synchronized when operating in the chaotic regime in a master-slave configuration if some parameters are properly matched; a short resynchronization time, which is two orders of magnitude shorter than that of COF case, can be obtained with this scheme; and OOPSK encryption can be successfully applied at a rate of hundreds of Mbit/s. This means that such devices are promising candidates for fast on/off phase shift keying encryption.

We believe that our work provides a good basis for future study and, in particular, provides some pointers for more detailed investigations of multi-section integrated devices and their applications for chaos-based communication systems.

Acknowledgements

This study was supported by projects 307b/s of the Technical University of Moldova and Picasso IST-2005-34551. The author thanks C. Mirasso, P. Colet, and I. Fischer for useful discussions and expresses his gratitude for the hospitality in the group IFISC at the UIB, Palma de Majorca (Spain) and in the Laser Dynamics group at the WIAS, Berlin (Germany).

References

[1] A. Pikovsky, M. Rosenblum, and J. Kurths, Synchronization: A Universal Concept in Nonlinear Sciences, Cambridge University Press, Cambridge, England, 405 p., 2003.

[2] Apostolos Argyris, Dimitris Syvridis, Laurent Larger, Valerio Annovazzi-Lodi, Pere Colet, Ingo Fischer, Jordi García-Ojalvo, Claudio R. Mirasso, Luis Pesquera, and K. Alan Shore, Nature, 438, 343, (2005).
[3] C.R. Mirasso, P. Colet, and P. Garcia-Fernandez, IEEE Photon. Tech. Lett., 8, 299, (1996).
[4] V. Annovazzi-Lodi, S. Donati, and A. Scire, IEEE J. Quantum Electron., 32, 953, (1996).
[5] S. Sivaprakasam and K.A. Shore, Opt. Lett., 24, 466, (1999).
[6] I. Fischer, Y. Liu, and P. Davis, Phys. Rev. A, 62, 011801, (2000).
[7] A. Bogris, D.F. Kanakidis, A. Argyris, and D. Syvridis, IEEE J. Quant. Electron., 41, 469, (2005).
[8] S. Tang and J.M. Liu, Opt. Lett., 26, 596, (2001).
[9] N. Gastaud, S. Poinsot, L. Larger, J.M. Merolla, M. Hanna, J.P. Goedgebuer, and E. Malassenet, Elect. Lett., 40, 898, (2004).
[10] F.Y. Lin and M.C. Tsai, Opt. Express, 15, 302, (2007).
[11] B. Krauskopf and D. Lenstra (Eds.), Fundamental Issues of Nonlinear Laser Dynamics, AIP Conference Proceedings, 548, (2000).
[12] Toni Perez, Mindaugas Radziunas, Hans-Juergen Wuensche, Claudio R. Mirasso, and Fritz Henneberger, IEEE Photon. Technol. Lett., 18, 20, 2135, (2006).
[13] S. Bauer, O. Brox, J. Kreissl, B. Sartorius, M. Radziunas, J. Sieber, H.J. Wunsche, and F. Henneberger, Phys. Rev. E, 69, 016206, (2004).
[14] C. Simmendinger and O. Hess, Physics Lett. A, 216, 97, (1996).
[15] F.R. Ruiz-Oliveras and A.N. Pisarchik, Opt. Express, 14, 12859, (2006).
[16] Yun Liu and Junjii Ohtsubo, IEEE J. Quant. Electron., 33, 1163, (1997).
[17] F. Rogister, P. Mégret, O. Deparis, M. Blondel, and T. Erneux, Opt. Lett., 24, 1218, (1999).
[18] F. Rogister, D.W. Sukow, A. Gavrielides, P. Mégret, O. Deparis, and M. Blondel, Opt. Lett., 25, 808, (2000).
[19] V.Z. Tronciu, H.-J. Wünsche, M. Wolfrum, and M. Radziunas, Phys. Rev. E, 73, 046205, (2006).
[20] S. Schikora, P. Hoevel, H-J. Wuensche, E. Schoell, and F. Henneberger, Phys. Rev. Lett., 97, 213902, (2006).
[21] Tilmann Heil, Josep Mulet, Ingo Fischer, Claudio R. Mirasso, Michael Peil, Pere Colet, and Wolfgang Elsäßer, IEEE J. Quant. Electron., 38, 1162, (2002).
[22] C.R. Mirasso, Application of semiconductor lasers to secure communications, Fundamental Issue of Nonlinear Laser Dynamics, Ed. by B. Krauskopf and D. Lenstra, 112, (2000).
[23] V.Z. Tronciu, C.R. Mirasso, and P. Colet, J. Phys. B: Atom. Mol. & Opt. Phys., 41, 155401, (2008).
[24] V.Z. Tronciu, I. Ermakov, P. Colet, and C.R. Mirasso, Opt. Comm., 281, 4747, (2008).
[25] V. Raúl, T. Pérez, and C.R. Mirasso, IEEE J. Quant. Electron., 38, 1197, (2002).
[26] R. Lang and K. Kobayashi, IEEE J. Quant. Electron., QE-16, 347, (1980).
[27] C.R. Mirasso, J. Mulet, and C. Masoller, IEEE Photon. Tech. Lett., 14, 456, (2002).
[28] F. Rogister, D. Pieroux, M. Sciamanna, P. Megret, and M. Blondel, Opt. Comm., 207, 295, (2002).
[29] A. Sanchez-Diaz, C.R. Mirasso, P. Colet, and P. Garcia-Fernandez, IEEE J. Quant. Electron., 35, 292, (1999).
[30] M. Peil, T. Heil, I. Fischer, and W. Elsässer, Phys. Rev. Lett., 88, 1741011, (2002).
[31] V. Annovazzi-Lodi, M. Benedetti, S. Merlo, T. Perez, P. Colet, and C.R. Mirasso, Phot. Tech. Lett., 19, 76, (2007).
[32] E. Schöll, Nonlinear Spatio-Temporal Dynamics and Chaos in Semiconductors, Cambridge University Press, Cambridge, 424 p., 2001.

THERMOELECTRIC PROPRIETIES OF LEAD TELLURIDE MICROWIRES

D. Meglei

Institute of Electronic Engineering and Industrial Technologies, Academy of Sciences of Moldova, 3/3, Academiei str., MD-2028, Chisinau, Republic of Moldova
E-mail: meglei@iieti.asm.md, Phone: +(37322) 737097
(Received 4 September 2009)

Introduction

In recent years, in physics of narrow-gap semiconductors, much attention is paid to semiconductors of the group A^4B^6 and solid solutions on their basis. Practical significance of these substances is determined by the possibility to produce a series of electronic devices, such as thermoelectric, optoelectronic, strain-gauge, etc.

The present-day methods of planar technology for preparation of films make it possible to control the type of conductivity and concentration of carriers in the process of epitaxial growth. However, the structure of films obtained is not ideal due to defects at the material-support interface.

Therefore, it is of interest to obtain and to study quasi-one-dimensional crystalline structures—microwires in glass isolation.

Obtaining of PbTe microwires

As the initial material for obtaining of microwires, we used compounds of lead telluride (PbTe) with conductivity of *p*- or *n*-type. Initial compounds were synthesized in the form of polycrystals or single crystals grown by the Bridgman-Stockbarger method [1]. PbTe compounds were synthesized by consecutive melting of chemical elements in quartz glass ampoules [2]. Initial components Pb-0000 and tellurium purified by the method of zone crystallization were placed into quartz glass ampoules with a diameter of 2-2.5 mm and a length of 120 mm. Before the placing, the chemical elements had been chemically processed, washed, and dried. To avoid sticking of the compound to the ampoule walls, the inner surface of the ampoules was treated with pure alcohol or acetone. After grinding, compounds the introduced into the ampoule to obtain microwires by the Ulitovsky method [3]. The microwires obtained were spooled on a coil with a diameter of 30 cm at a rate of 10-15 m/min.

Results of the study of the microwires showed that there are many cracks along their length and that their number does not depend on sort of the glass and on technology of obtaining. The main causes of the cracks are the high rate of crystallization and the difference in the linear temperature coefficient of the glass and PbTe. The high rate of crystallization favors supercooling in the melt, which results in the appearance of spontaneous crystallization centers, which leads to the creation of new centers of crystallization. The melt in the glass capillary, being between two crystallization centers, is attracted between them in the process of crystallization, which results in the formation of microcracks in the microwires.

In addition, in the process of formation of microcracks, the atmospheric pressure inside the capillary decreases leading to elimination of intrinsic vapors of the PbTe melt, which also results in the formation of microcracks in the microwires.

Therefore, to obtain microwires of PbTe, we applied the method that involves the filling of preliminarily evacuated glass capillaries with a melt of PbTe by means of inert gas pressure as shown in Fig. 1 [1]. For this purpose, evacuated quartz ampoule *1* with weighted amount of PbTe *2*, above which sealed quartz capillaries *3* are placed, is introduced into a high-temperature furnace in such a way that capillaries with the material are situated in the operating zone of the furnace (its temperature is 10-15° higher than the melting temperature of the material). After melting of the material and the immersion of the capillaries in the melt, an excess pressure is generated by means of inert gas (P = (6. 8 - 49) x 10^4 Pa); as a result, the capillaries are filled with the melt. Afterwards, the oriented crystallization of the melt in the capillaries takes place.

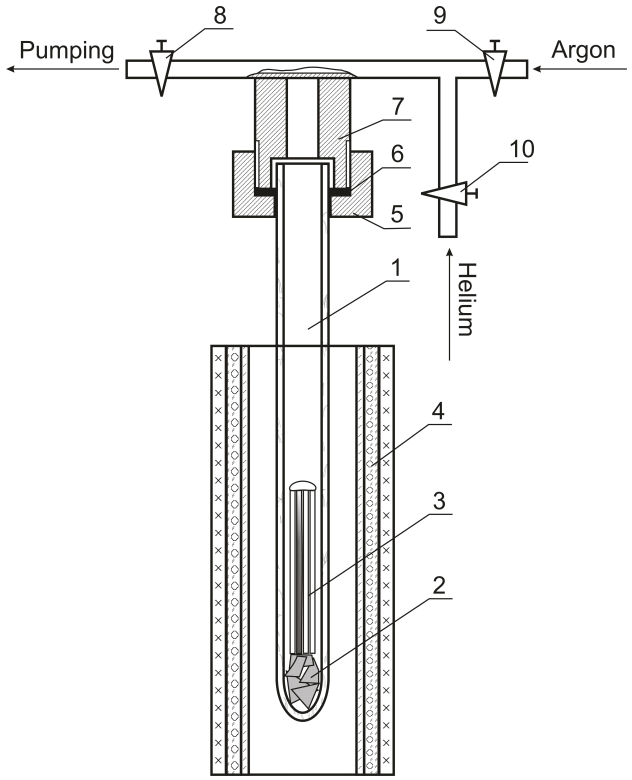


Fig. 1. Diagram of the installation for obtaining of PbTe microwires: *1* is the quartz ampoule; *2* is the PbTe material; *3* is the quartz capillaries; *4* is the furnace; *5, 6, 7* are the coupling device; *8, 9, 10* are taps for evacuation and introduction of inter gas.

It should be mentioned that high rates of growth of the microwires (higher than 15 cm/h) lead to the appearance of microcracks along the length of microwires as well as in the case of the Ulitovsky method. Low rates (lower than 1 cm/h) result in the formation of step-like pits due to evaporation of material from the microwire surface.

Qualitative PbTe microwires were obtained in the following mode: the growth rate was 1-10 cm/h; the temperature gradient was 10-15°/cm; the maximum melting temperature was maintained to be 10-15° higher than the temperature of crystallization; the excess of inert gas pressure was (6.8 - 49) x 10^4 Pa [4].

Electrothermal properties of PbTe microwires

Crystalline microwires of lead telluride (PbTe) with the *n*-type electric conductivity were obtained from the initial material being a chemical compound with an excess of lead with respect to the stoichiometric composition; the ones of the *p*-type, with an excess of tellurium [1].

Temperature dependence of resistivity of PbTe microwires of the *n*- and *p*-types (Figs. 2 and 3, respectively) in the temperature range of 80-350 K showed the following:

- The resistivity exhibits the metallic behavior; that is, it increases with growing temperature; the intrinsic conductivity for the crystals studied in this temperature range does not occur.
- The resistivity depends on microwire diameter; as the diameters decreases, the resistivity ρ decreases as well.

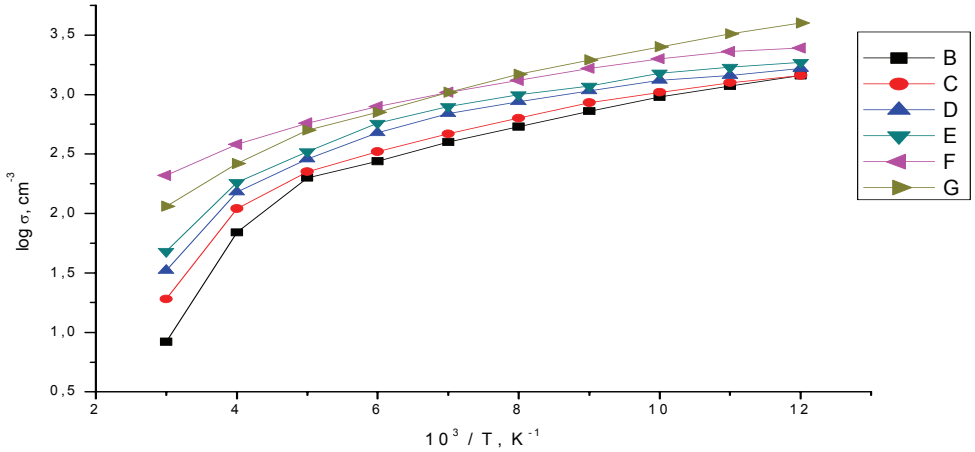


Fig. 2. Dependence of resistivity on temperature for PbTe microwires with the *n*-type conductivity: B - d = 125.0 μ; C - d = 92.0 μ; D - d = 51.0 μ; E - d = 33.0 μ; G - d = 15.0 μ; F - n = 1.2 10^{18} cm^{-3} at the temperature T = 77 K.

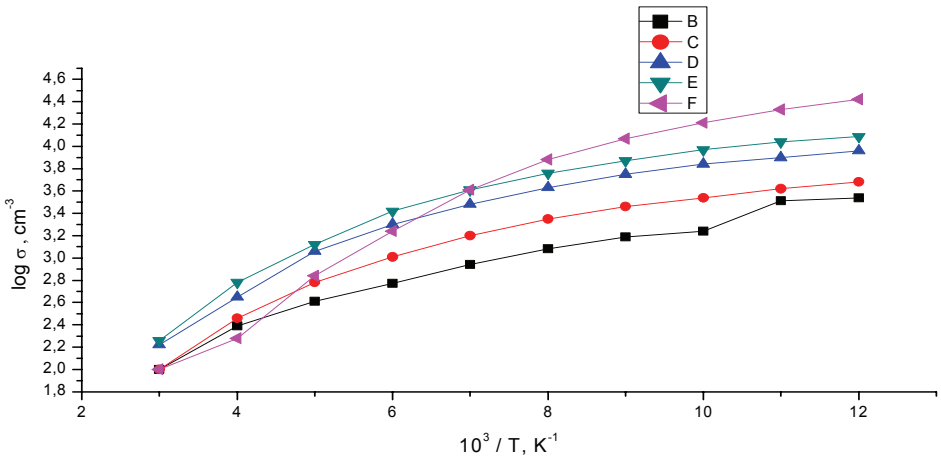


Fig. 3. Dependence of resistivity on temperature for PbTe microwires with the *p*-type conductivity: B - d = 103.0 μ; C - d = 45.0 μ; D - d = 29.3 μ; E - d = 17.8 μ; F - p = 1.47 10^{18} cm^{-3} at T = 77 K.

The decrease in the resistivity due to the decrease in the microwire diameter can be explained as follows. On the basis of the above, for current carriers of the p- or n-type, the dependences between resistivity (ρ), concentration (n), and carrier mobility (μ) is given by the following expression

$$\rho = (e\, n\, \mu)^{-1}. \qquad (1)$$

It follows from this expression that resistivity can decrease due to the increase in concentration (n) or in mobility (μ).

The studies of the temperature dependence of thermopower for the microwires with different diameters showed (Figs. 4 and 5, respectively) that, as the microwire diameter decreases, the thermopower coefficient (α) hardly changes at all.

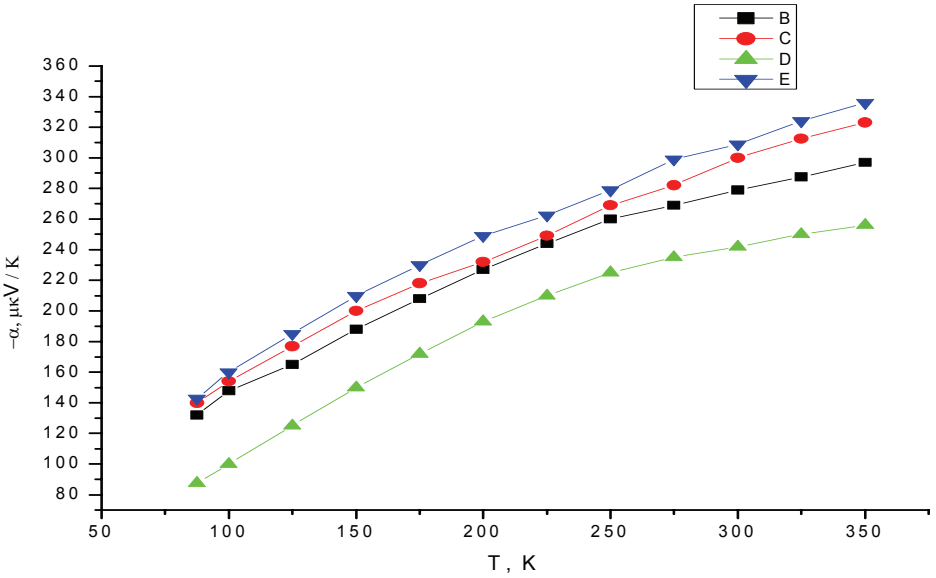


Fig. 4. Temperature dependence of the thermopower coefficient of PbTe microwires with the *n*-type conductivity: B - d = 125.0 μ; C - d = 51.0 μ; E - d = 15.0 μ; D - n = 1.2 10^{18} cm^{-3} at a temperature of 77 K.

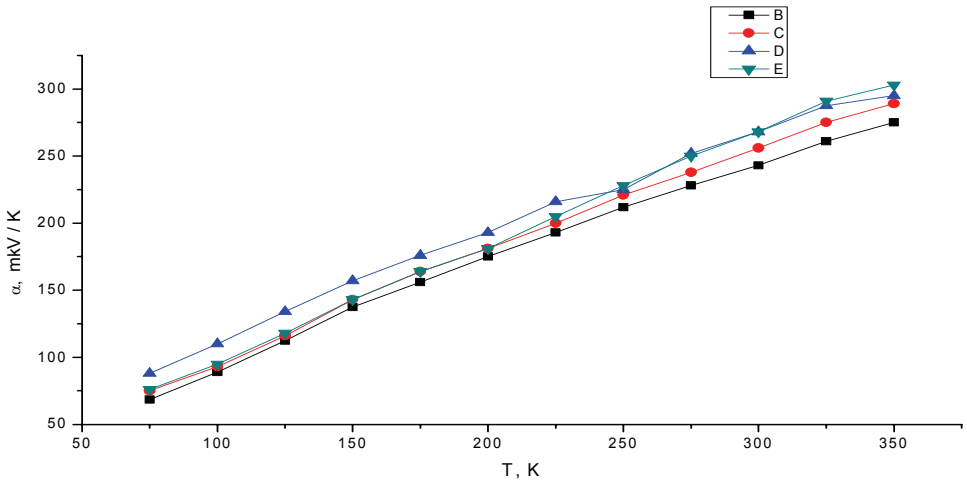


Fig. 5. Temperature dependence of the thermopower coefficient of PbTe microwires with the *p*-type conductivity: B - d = 103.0 μ; C - d = 45.0 μ; D - d = 17.8 μ; E - p = 1. 47 10^{18} cm^{-3} at temperature T = 77.

It follows from the relation of the coefficient α, which is

$$\alpha_{p(n)} = \pm k/e\,[A + \ln (2(2\pi m_{p(n)} kT)^{3/2} / p\,(n)\,h^3)], \qquad (2)$$

that the concentration of carriers of the p(n) type in the samples does not depend on diameter of microwires within the studied range. Thus, it follows from the obtained results that the decrease in the resistivity ρ is due to the increase in the mobility of carriers. Therefore, the increase in the mobility with decreasing diameters counts in favor of the enhancement of the structure perfection and the improvement of the homogeneity of microwires with decreasing diameter.

The comparison of the results of measurements of resistivity and thermopower coefficient of microwires and bulk crystals allows making the conclusion that the values for ρ and α range within the limits for the bulk crystals grown by the Bridgman method. This shows that the microwires obtained by this method retain the properties characteristic of bulk crystals.

Heat efficiency of PbTe microwires

According to the thermoelectricity energy theory, the efficiency of the thermal element branch material is determined by

$$Z = \alpha^2 \sigma / \kappa, \quad (3)$$

where Z is the heat efficiency; α is the thermopower; σ is the electric conductivity; κ is the heat conductivity of the material. One can see from the formula that the efficiency depends directly on the product $\alpha^2\sigma$ and is inversely proportional to the heat conductivity κ. Therefore, we will fix on the behavior of the parameter $\alpha^2\sigma$—the so-called thermoelectric power. Figures 6 and 7 show the temperature dependences of thermoelectric power for microwires of *n*- (Fig. 6) and *p*-type (Fig. 7).

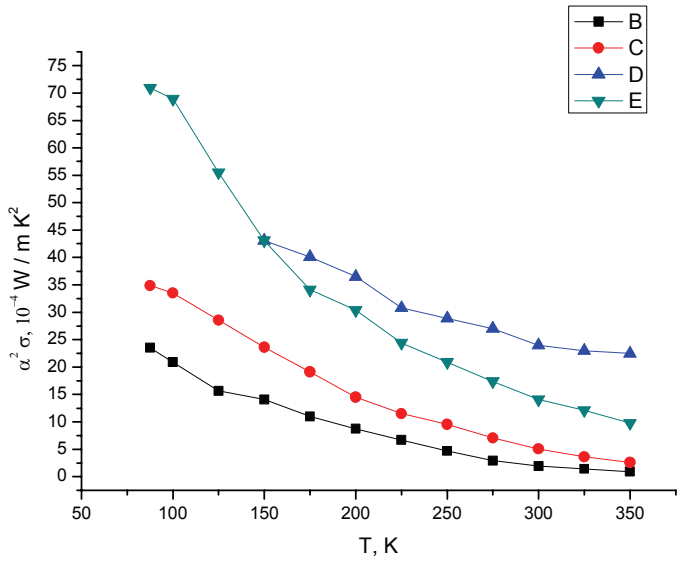


Fig. 6. Temperature dependence of the thermopower of *n*-type PbTe microwires: B - d = 125.0 μ; C - d = 51.0 μ; D - n = 1.2 10^{18} cm^{-3} at a temperature of 77 K; E - d = 15.0 μ.

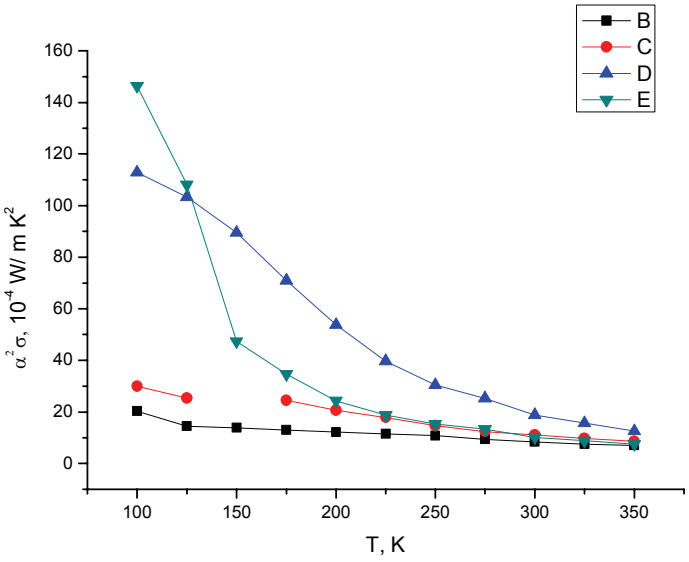


Fig. 7. Temperature dependence of the thermopower of *p*-type PbTe microwires B - d = 103.0 μ; C - d = 45.0 μ; E - p = 1.47 10^{18} cm^{-3} at a temperature of 77 K; D - d = 17.8 μ.

The calculation showed that if we compare the heat conductivity of microwires to the conductivity of bulk samples, then the efficiency of the material of *n*-type is $Z = 0.7\ 10^{-3}\ K^{-1}$ at a temperature of 300 K.

Conclusions

1. A technology for obtaining of quartz-glass covered microwires of thermoelectric material PbTe with conductivity of the *p*- and *n*-type, with a core diameter of 10-150 μ, has been developed.
2. A series of measurements of electrical and thermoelectrical properties of microwires with different diameters of the crystalline core in the temperature range 77-350 K has been carried out.
3. Analysis of the results of measurements of resistivity and thermopower in comparison with properties of bulk materials obtained by the Bridgman method has shown that the numerical values of α and ρ for microwires are within the same limits. This means that obtaining of PbTe microwires by the method of capillary filling preserves the same properties as in bulk materials obtained by melting (deviation from stoichiometry, similar limits of solubility of components, high concentration of current carriers).
4. The heat efficiency of the PbTe material is $Z = 0.7\ 10^{-3}\ K^{-1}$ at a temperature of 300 K.

References

[1] V.F. Banari, F.G. Donica, and A.I. Shkurpelo, Obtaining and Some Structural Characteristics of Thin Wires of Lead Telluride, Izv. Akad. Nauk MSSR, Ser. Fiz. Tekh. Mat. Nauk, 3b, 66-67, (1982).
[2] D.F. Miglei, Anisotropy of Transport Phenomena in Solid Solutions of Lead-Tin Telluride, Extended Abstract of Candidate Dissertation, 16 p., (1982).
[3] A.V. Ulitovsky and N.M. Averin, Method for Preparation of Metal Microwire, USSR Author's Certificate MKI g 03 C 29/-00, no. 426025/22-2, Published 19.03.64, Bulletin no. 7, C14.
[4] D.F. Meglei, M.P. Dyntu, and S.V. Donu, International Conference ICT+ "Information and Communication Technologies", Chisinau, 105-106, (2009).

APPLICATION OF AMORPHOUS MICROWIRES FOR ELECTROMAGNETIC SHIELDING

S.A. Baranov[1], M. Yamaguchi[2], K.L. Garcia[2, 3], and M. Vazquez[3]

[1]Institute of Applied Physics, Academy of Sciences of Moldova, 5, Academiei str., MD-2028, Chisinau, Republic of Moldova. Phone: +(37322) 731725, fax: +(37322) 731725; e-mail: baranov@phys.asm.md
[2]Graduate School of Engineering, Tohoku University, 980-8579, Sendai, Japan e-mail: yamaguti@ecei.tohoku.ac.jp
[3]Instituto de Ciencia de Materiales, CSIC, Campus Cantoblanco, 28049, Madrid, Spain e-mail: mvazquez@icmm.csic.es
(Received 9 December 2009)

1. Introduction

Recently, the demand for high density integration of electronic devices has led to serious noise and signal interference problems, promoting the research and development of thin films with specific features of electromagnetic noise absorption [1]. As this research is evolving, a need to control the magnetic anisotropy and resistivity for thin films came into attention since near field absorption is tailored by ferromagnetic resonance profile and eddy current losses [2]. Several attempts to control these features have been successfully made, especially for the cases of granular and patterned thin films, in which the demagnetizing factor, that is, the shape anisotropy plays an important role [3].

On the other hand, it has been widely documented that amorphous microwires exhibits very interesting high frequency properties due to the strong magnetoelastic anisotropy which is induced in the fabrication process [4]. By defining strictly the geometric characteristics and the alloy composition, the ferromagnetic resonance profiles can be accurately controlled (see below). Furthermore, the high production rate and high permeability of microwires without any rigid substrates are advantageous for the use as noise absorbers.

In this report, amorphous glass-coated microwire arrays in a sheet configuration are proposed to be used as absorbers. High frequency characteristics of thin Co-based amorphous microwires will be presented and estimated as noise absorbers.

2. Preparation of shielding

In the course of preparation of microwires using the Ulitovsky-Taylor method, the soften glass capillary is filled with a melted metal alloy and then quenched. The residual stresses arising in the metal core of amorphous glass-coated microwires in the course of preparation determine the magnetic behavior of the microwires. The metal is frozen into a very thin amorphous wire. Thus, a strong quenching stress (for example, $\sigma \sim 10^9$ Pa in ultra thin microwires) is introduced into the wire [5].

For materials with a negative magnetostriction, the orientation of magnetization vector is determined by the minimal component of the stress tensor. As was shown in [5, 6], the thermoelastic stresses are characterized by equal radial and tangential components which do not depend on r coordinate and are 2-3 times smaller than the axial ones as a result of the cylindrical symmetry of the microwire.

However, for very thin Co wires with negative but almost zero magnetostriction, the domain structure is far more complicated due to the complex stress distribution within a small diameter. The simplest domain structure reported so far for these very thin wires is mainly composed of radially magnetized domains [6, 7].

The composition of microwires investigated here is $Co_{68.15}$ $Fe_{4.35}$ $Si_{12.5}B_{15}$, which has a very small saturation magnetostriction constant (λ_s) of ~-1.11 x 10^{-7}.

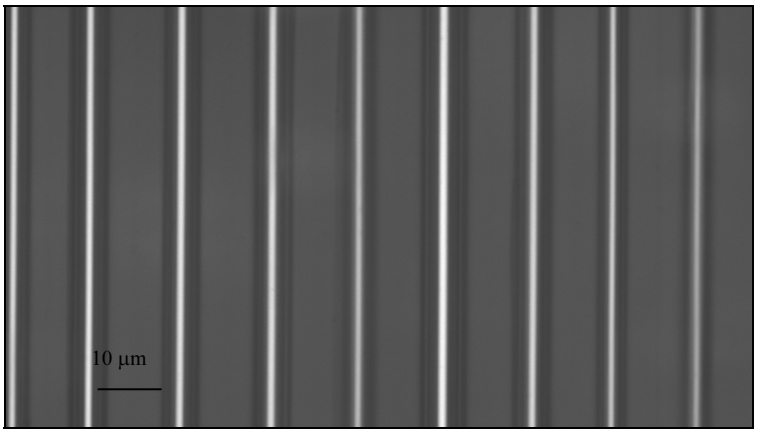


Fig. 1. Micrograph of an array of Co based microwires.

As for the microwire array configuration for noise suppression experiments, 223 microwires (with a total diameter of 10 μm and a length of 5 mm where the metal core diameter is 2 μm) are aligned in parallel within an area of 5 x 4 mm^2 as can be observed in Fig. 1. They are supported by a thin polymer sheet (~100 μm thickness). Axial and transverse hysteresis loops for the microwires were measured with a vibrating sample magnetometer as shown in Fig. 2.

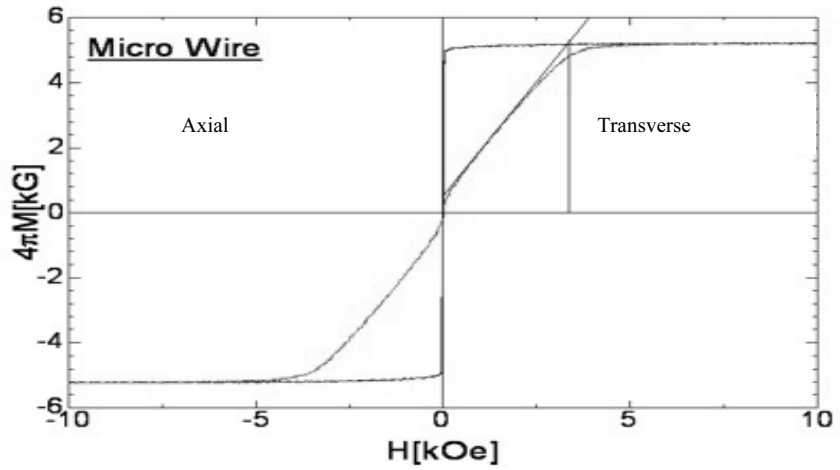


Fig. 2. Axial and transverse loops for a single microwire film.

The anisotropy field in the axial and perpendicular direction is ~10 Oe and 3 kOe, respectively. For the axial loop, the magnetization process is quasi-nonhysteretic (still, a low hysteresis appears near ± 5 Oe), and the magnetization process is mainly due to magnetization rotation.

Since the multilayer structures of the microwire array sheets are more practical as noise absorbers, samples with various layer structures, from single to four layers, were prepared.

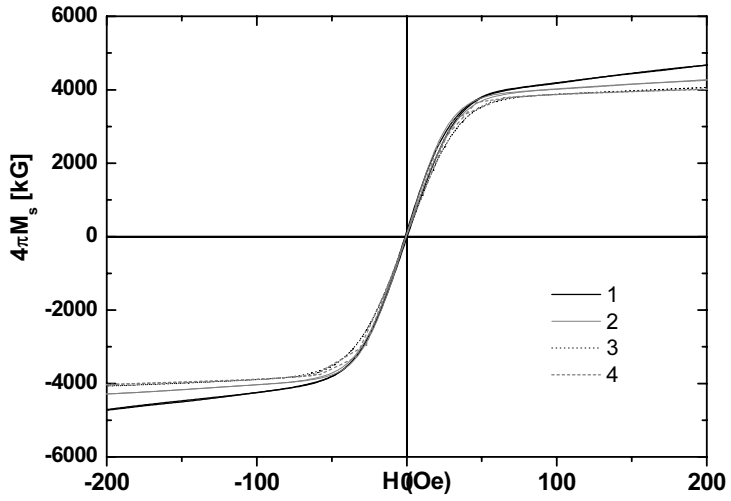


Fig. 3. Axial loops for one to four microwire films.

Figure 3 shows a slight increase in the anisotropy field, H_k, from 10 to 22 Oe. This suggests that, in a multilayer configuration, the initial permeability decreases slightly due to the dipole-dipole interaction between the wires.

The initial permeability for the single-layer sample was measured by the ferrite core method. It exhibited an initial relative permeability of 6000 at 1 MHz. A rough estimation of relative permeability from the slope of the curves in Fig. 3 showed values of 3600. This highly enhanced initial permeability suggests that the very thin microwires studied here have a complex domain structure which differs from the so-called "bamboo domains" [8]. Some authors have attributed this high value of permeability to the presence of regions with lower anisotropy within the 2 μm metallic wires. However, at this point, we could not confirm this assumption experimentally.

3. Permeability characterization

Initial permeability at high frequencies was measured using the in-lab 9 GHz permeameter. The initial permeability profile of a single microwire layer, thus obtained, is shown in Fig. 4.

A rapid decrease in the real part of the permeability μ' was observed. This can be attributed to the eddy current caused by the high initial permeability. The skin penetration depth is defined as $\delta = (\rho/\pi f \mu\mu_0)^{1/2}$, where $\mu\mu_0$ is the effective magnetic permeability and ρ is the microwire electrical resistivity of about 200 μΩ cm. Considering the measured initial permeability of 3600, δ takes values between 1.5-0.1 μm for frequencies of 0.5 to 9 GHz. This estimation is significant when compared to the microwire diameter of 2 μm.

However, a possible contribution due to wall displacement loss should not be excluded; domain walls could contribute significantly to initial permeability observed at lower frequencies, at higher frequencies domain wall movement is damped by microscopic eddy currents rapidly [9].

A multilayer configuration increases the value of field penetration depth δ as seen in Fig. 5. This behavior is a direct consequence of the subsequent decrease in the permeability and is consistent with the axial loops measured.

The resonance frequency of 0.8 GHz observed in Fig. 5 should be treated carefully since the value is in agreement with the theoretical prediction by Kittel [10], which assumes the magnetization component lying parallel to its axis.

When a cylinder without any magnetic anisotropy, except for the shape anisotropy, is submitted to an AC field along its axis,

$$\omega/\gamma \sim 2\pi\mathbf{M} + \mathbf{H_e}, \quad (1)$$

where $\mathbf{H_e} = \mathbf{H_y} + \mathbf{H_k}$, $\mathbf{H_y}$ is the outside field, $\mathbf{H_k}$ is the inside field, $\mathbf{M}$ is the magnetization (determined in [7]), and $\gamma = 2.8$ MHz/Oe is the gyromagnetic ratio.

The demagnetizing field for a precession by the magnetization lying parallel to the cylinder axis is $2\pi M$ which, for the present sample, predicts $\omega/2\pi$ to be ~ 1 - 2 GHz [7].

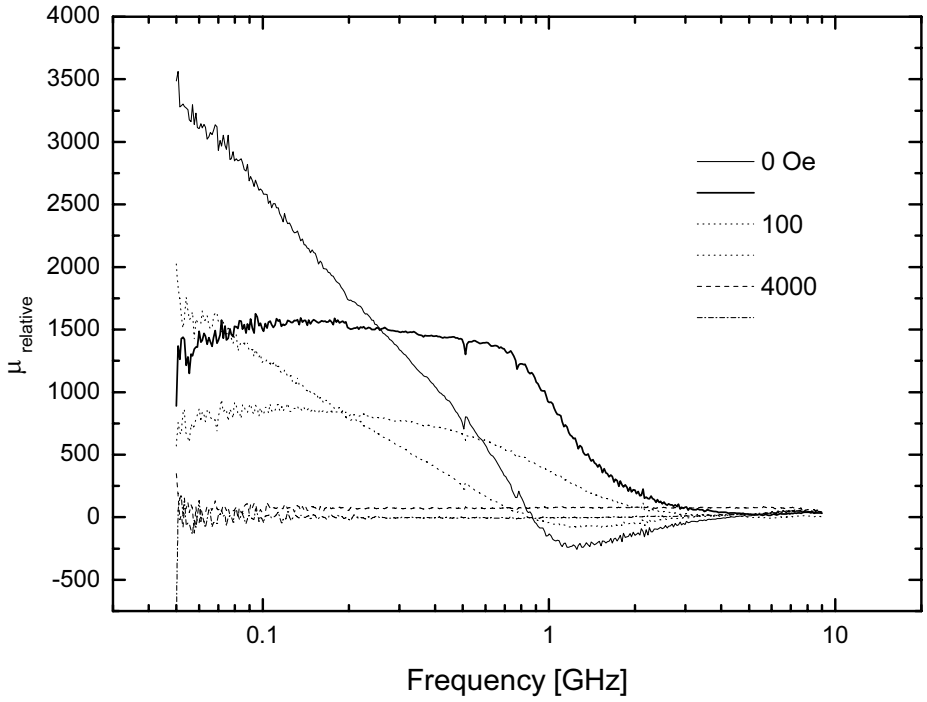


Fig. 4. Permeability of a single Co microwire thin film. H_y is the outside field as a parameter.

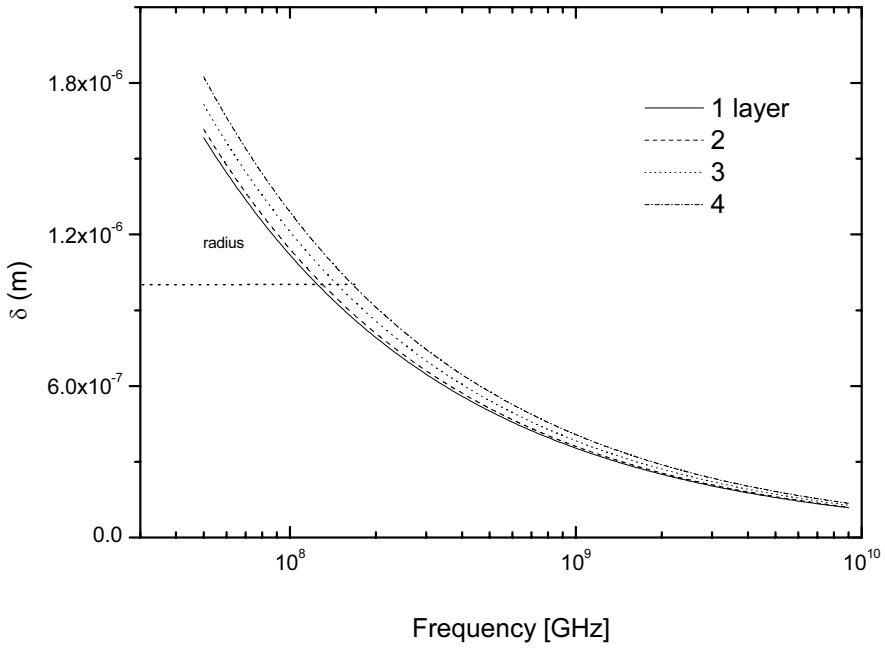


Fig. 5. Skin depth of several microwire layers.

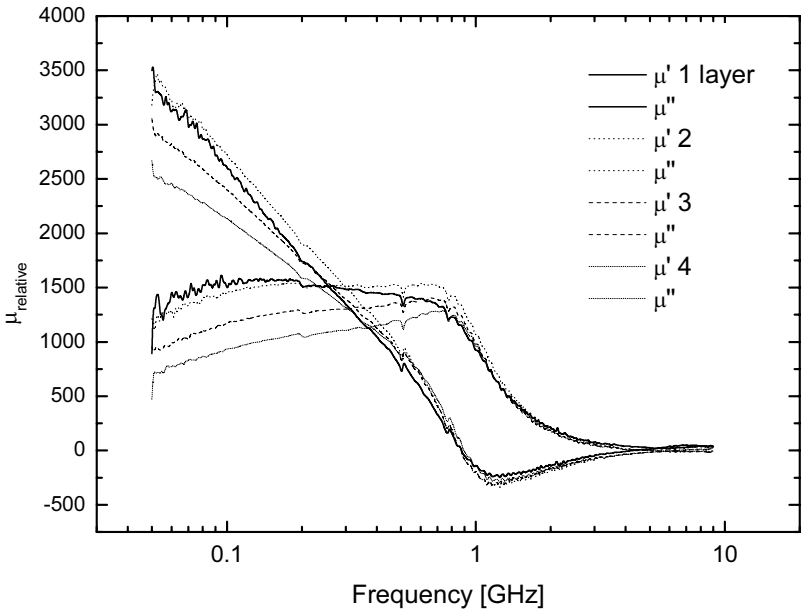


Fig. 6. Permeability profile for a multilayer configuration.

It is plausible that, if the high initial permeability region in the wire has the magnetization component along the wire axis, the resonance occurs at a frequency very close to the above estimation.

A rough estimation assuming H_k to be 10 or 20 Oe, as is seen in Fig. 3, results in negligible difference from the above mention frequency value, as $2\pi M + H_e \sim 2\pi M$. This is another support for the assumption that there is a specific region in the wire which gives rise to high permeability and its high frequency profile, with magnetization direction off axis.

The permeability profile for the cases of multiple layers is shown in Fig. 4. These curves were measured without H_y. The relative permeability values μ'' and μ' decrease as the layer number increases; also, the resonance frequency value shifts from 0.85 to 0.9 GHz. This may be due to the dipole-dipole interaction between the wires as is observed in Fig. 3. This slight change of frequency profiles is not influential in the noise suppression experiment studying the advantage of a multilayer.

4. Noise absorption experiments

In order to investigate the noise absorption performances, a microstrip line (with 50 Ω characteristic impedance) has been designed. The signal line has a width of 610 μm and a thickness of 35 μm and is set on a 300 thick Teflon base with permittivity $\varepsilon_r = 3.6$. The insertion losses are estimated below 0.08 at 1 GHz. The performance of the line is measured using an Agilent E8363B network analyzer from 0.1 to 20 GHz. The microwire array is placed on the top of the signal line.

Samples of a single and twofold array of microwires were prepared. The measured absorption coefficient S_{21} on the microstrip line is shown in Fig. 7. The S_{21} parameter at 1 GHz decreases from the background of -0.08 dB to -0.154 dB and -0.188 dB for the single layer and twofold layer (labeled as A to B). It is worth mentioning that larger decreases are observed at ~20 GHz. This is probably due to the metallic property of the wires.

The ratio of the power loss by the microwire layers to the input power calculated as $P_{loss}/P_{in} = 1-(|S_{11}|^2+|S_{21}|^2)$ is shown in Fig. 8. A loss peak nearby the resonance frequency is

observed around for both samples. It is noted that the twofold sample fails to be adequately effective. The reason is that the second layer is not close enough to interact with the near field from the microstrip line.

From the experimental results described above, the following should be considered for future improvements.

Although the amorphous wires exhibit superb frequency profiles and various advantages such as a high production rate and geometrical flexibility, the metal cores in the wires are coated with very thick glass layer resulting in a poor packing density of high permeability region in the wire array samples. If the packing density of 0.12 for the given samples were, for example, 0.8, the power loss would be about 0.3 for a single layer. This can be achieved by removing the glass coating. Furthermore, by designing a wire array with proper conductivity and permeability, high performance absorbers as the films reported in [2] may be developed.

These issues will be treated intensively as this research work progresses in the near future.

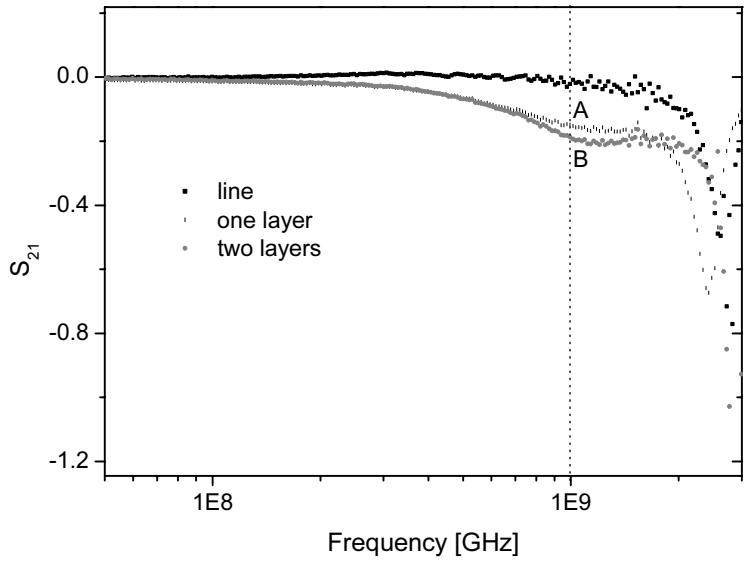


Fig. 7. S_{21} absorption profile of Co based microwires.

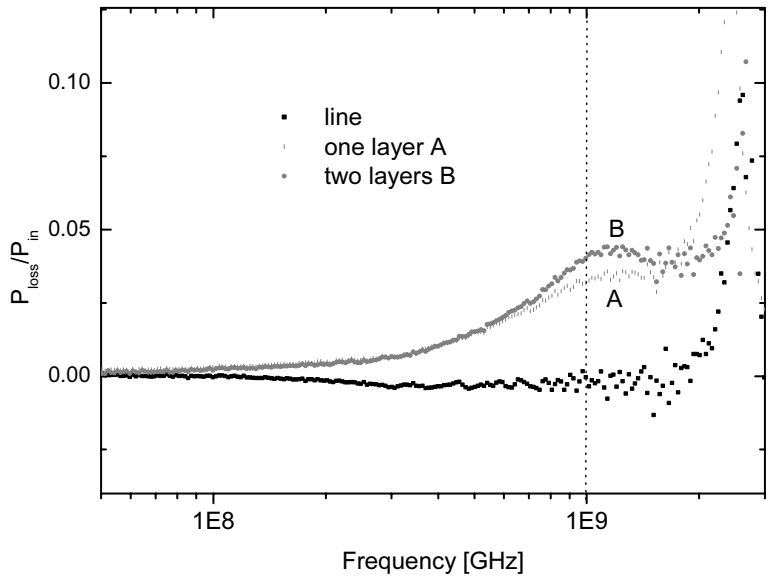


Fig. 8. Power loss profile of Co based microwire thin film.

5. Conclusions

The high permeability features for extremely thin Co based amorphous microwires were investigated. It is likely that a high permeability region exists in the wire which was not revealed by B-H loop measurements. The rapid decrease of high permeability at high frequencies was ascribed to eddy current losses. The electromagnetic noise absorption of sheet type samples with dimensions (5 x 4) mm^2 made of 223 aligned in parallel microwires was measured. The twofold layers exhibit S_{21} of -2.5 dB, with insertion losses below 0.2 at 1 GHz, which corresponds to the resonance of magnetization at 0.8 GHz. This low absorption rate was ascribed to the poor packing density of the metal core in the wires, subject to be investigated and discussed in further development.

Acknowledgements

K.L Garcia thanks the Japanese Society for the Promotion of Science (JSPS) for a Postdoctoral fellowship at Yamaguchi lab, Tohoku University.

References

[1] K.H. Kim and M. Yamaguchi, Phys. Met. Metallogr., 102, S74-S79, (2006).
[2] K. Maruta, S. Muroga, Y. Shimada, and M. Yamaguchi, Technical Report of IEICE, 106, 25, (2006).
[3] Y. Shimada, M. Yamaguchi, S. Ohnuma, T. Itoh, W.D. Li, S. Ikeda, K.H. Kim, and H. Nagura, IEEE Trans. Magn., 39, 5, (2003).
[4] S.A. Baranov, Tech. Phys. Lett., 24, 7, 21, (1998).
[5] S.A. Baranov, V.N. Berzhanski, S.K. Zotov, V.L. Kokoz, V.S. Larin, and A.V. Torkunov, Fiz. Met. Metalloved., 67, 73, (1989).
[6] S. Baranov, J. Magn. Magn. Mater., 266, 278, (2003).
[7] S.A. Baranov and M. Vazquez, Tech. Phys. Lett., 30, 989, (2004).
[8] M. Vazquez and A.-L. Adenot-Engelvin, J. Magn. Magn. Mater., 321, 2066, (2009).
[9] S. Chikazumi, Physics of Magnetism, John Wiley & Sons, 321, 1964.
[10] C. Kittel, Phys. Rev., 73, 155, (1948).

METROLOGICAL SIMULATORS OF ELECTRICAL PASSIVE QUANTITIES WITH ALGORITHMIC STRUCTURE

V. Nastas

Department of Telecommunications, Technical University of Moldova, 168, Stefan cel Mare ave., MD-2005, Chisinau, Republic of Moldova
E-mail: vitalienastas@gmail.com
(Received 22 January 2010)

Abstract

This Review contains the systematization of the material in the domain of metrological simulators of electrical passive quantities with an algorithmic structure (MS-A, for simplicity), used as reference elements in impedance measurement. The general aspects of these devices, the MS-A classification under the relevant criteria, the formal–structural method of synthesis of the simulators with necessary characteristics, and the example of synthesis are discussed. The analysis of basic types of MS-A: I-commanded and U-commanded, in Cartesian coordinates and in polar coordinates, with asymmetric and with symmetric connection is carried out. For each type of simulator, the conversion diagram of information, the simulator structure, and the internal circuits are presented. The paper also contains the analysis of stability of presented structures, in which the conditions of guaranteed functional stability for MS-A at its connection in external circuits are determined, and experimental confirmation of the results.

1. Introduction

The use of metrological simulators of passive electrical quantities (MSs, for simplicity) as reference elements (REs) in impedance and admittance measurement opens great prospects for the improvement of technical characteristics in this branch [1-5]. The most important advantages resulting therefrom are:

- exclusion of adjustable reactive elements and reactance boxes;
- exclusion of the switchings in the measurement circuits, determined by the variation of type and character of the measured quantity;
- simplification of measuring circuit equilibration algorithm up to two operations in the case of measuring the both components of passive quantity (PQ);
- measurement of PQ with any character (active, reactive or complex) and with any substitution equivalent circuit (series, parallel) without modification of the measuring circuit structure;
- measurement of PQ with character of negative resistance;
- digital control and complete automatization of measuring process;
- possibility of RE implementation in integrated circuits;
- reduction of the price, dimension, and weight of the devices.

From the functional point of view, a metrological simulator of passive electrical quantities may be defined as an elementary resource of measurement, which assures the reproduction of a virtual PQ with the necessary character and with known metrological characteristics [6]. The researches in this domain [7] determined a class of devices that can be potentially used as MSs. According to the proposed classification [7], there are different types of MS with various characteristics and possibilities of use.

MSs with an algorithmic structure (MS-A) [7] hold a particular position in this class of devices. These devices were specially synthesized for using as REs in the devices for PQ measurement (impedance meters and admittance meters); they possess technical characteristics optimized from this point of view. The MS-A structures were synthesized through the formal–structural method (FSM) according to the requirements on the RE characteristics determined by the measuring method and circuit [8]. The most important features of them are:

- possibility to obtain PQs represented in the desired coordinate system (Cartesian or polar) and with the desired character of components (active, reactive, or complex) reproduction;
- independent control of the reproduced PQ components;
- guaranteed stability at the variation of external impedance in the measuring circuit in the range of determined values;
- possibility of reproduction of the simulated PQs (SPQs) placed in the entire complex plan: (±R, ±**j**X) for impedances or (±G, ±**j**B) for admittances;
- guaranteed and determined systematic error.

2. MS-A. General overview

Functionally, a MS-A represents a device with two poles, which assures the reproduction of SPQ represented in the necessary coordinates system and the possibility of independent control of the components (Fig. 1) [9]. The PQ reproduced by MS-A (impedance $\mathbf{Z_i}$, or admittance $\mathbf{Y_i}$) may be represented in Cartesian coordinates [10]

$$\mathbf{Z_i} = R_i + \mathbf{j}X_i,\ \mathbf{Y_i} = G_i + \mathbf{j}B_i, \tag{1}$$

alternatively, in the polar coordinates [11, 12]

$$\mathbf{Z_i} = Z_i \exp(\mathbf{j}\varphi_i),\ \mathbf{Y_i} = Y_i \exp(\mathbf{j}\psi_i), \tag{2}$$

where R_i, X_i (G_i, B_i) are the active and the reactive component of the reproduced impedance (admittance), respectively; Z_i, φ_i (Y_i,ψ_i) are the module and the phase of these quantities, respectively.

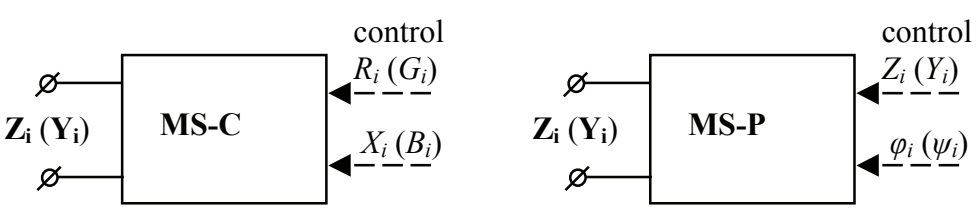


Fig. 1. Functional representation of MS-A in Cartesian coordinates (a) and in polar coordinates (b).

Therefore, the MS-A (MS-C) Cartesian coordinates must assure an independent control of the active R_i and the reactive X_i components (respectively, G_i, B_i for admittances); the polar coordinates MS-A (MS-P), an independent control of module Z_i (Y_i) and phase φ_i (ψ_i) of the PQs reproduced on the input poles.

For assurance of the MS universality in order to reproduce a PQ with any character, the domains of variation of the reproduced PQ components must be [7]

- for simulated impedances

$$\begin{gathered} R = \{-R_{\max} \div +R_{\max}\};\ X = \{-X_{\max} \div +X_{\max}\} \\ Z = \{0 \div Z_{\max}\};\ \varphi = \{0 \div 360^{\circ}\}, \end{gathered} \tag{3}$$

- for simulated admittances

$$\begin{gathered} G = \{-G_{\max} \div +G_{\max}\};\ B = \{-B_{\max} \div +B_{\max}\} \\ Y = \{0 \div Y_{\max}\};\ \psi = \{0 \div 360^{\circ}\}. \end{gathered} \tag{4}$$

In this case, the coverage of the entire complex plan is assured, and, due to this, there is a possibility to reproduce an SPQ with any character.

2.1. The MS-A classification

As is known, there are many types of MSs with various features, according to the classification proposed in [7]: classical MS [13], MS-A, MS with a ladder structure [14], MS type gyrator [15], etc.

MS-A forms a class of devices with the structure synthesized through the FSM according to the requirements determined by its concrete application in measuring circuits. The MS-A classification by relevant criteria is represented in Table 1.

Table 1. Classification of MS-A.

Character of poles	The primary input quantity	The type of coordinates for SPQ representation	
		Cartesian coordinates	**P**olar coordinates
Asymmetrical connection	**U** – comanded	**U-MS-C-As**	**U-MS-P-As**
	I – comanded	**I-MS-C-As**	**I-MS-P-As**
Symmetrical connection	**U** – comanded	**U-MS-C-S**	**U-MS-P-S**
	I – comanded	**I-MS-C-S**	**I-MS-P-S**

For designation of type of MS-A in the table, the qualifying abbreviations have been used. For example, for the simulator represented in the table as "I-MS-C-As", the signification is:

I – current commanded (U – voltage commanded),

C – Cartesian coordinates (P – polar coordinates),

As – with asymmetrical connection (S – symmetrical connection).

These devices correspond to the general classification proposed in [7]; further, for simplicity, three criteria are considered relevant.

1. The type of primary input quantity. One of the basic properties of SPQ determines this criterion: to be commanded by current (for I-MS) or by voltage (for U-MS) [9]. As is known [9], the type of primary input quantity determines some features of MS and the type of its stability.

The principle of SPQ reproduction in I-MS [8] (Fig. 2a) is based on the process of formation of the voltage $\mathbf{U_i}$ from the input current $\mathbf{I_i}$ by means of its conversion under necessary dependence and its application in the input circuit of the SPQ. Together with the current $\mathbf{I_i}$, the voltage $\mathbf{U_i}$ results in the reproduction of the virtual impedance $\mathbf{Z_i}$

$$\mathbf{Z_i} = \mathbf{U_i} / \mathbf{I_i} = (\mathbf{K_{conv}} \cdot \mathbf{I_i}) / \mathbf{I_i} = \mathbf{K_{conv}}, \quad (5)$$

where the conversion factor $\mathbf{K_{conv}}$ determines the mode of SPQ representation. For this, the structure of I-MS (Fig. 2b) contains a current-to-voltage converter I/U and a functional converter CF, which assures the desired form of $\mathbf{K_{conv}}$.

Therefore, it follows from (5) that I-MS are destined for reproduction of PQs with impedance character and possess the stability up to no-loaded regime [9].

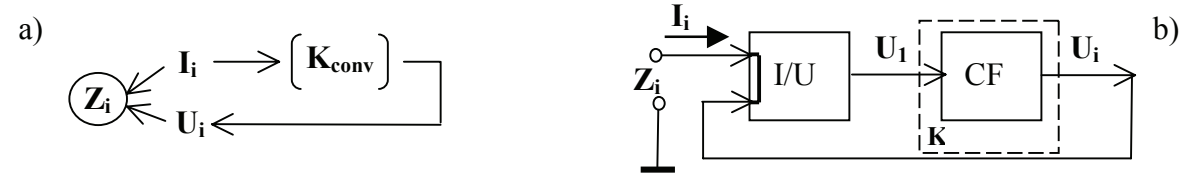


Fig. 2. The conversion algorithm of information (a) and the block-diagram of I-SPQ (b).

For U-SPQ [8] (Fig. 3), the voltage $\mathbf{U_i}$ is used as a primary input quantity. It is consecutively passed through the functional converter CF and through the voltage-to-current converter U/I. The reproduced PQ $\mathbf{Y_i}$ possesses the character of admittance and is obtained as a result of interaction of the entering voltage $\mathbf{U_i}$ and the current $\mathbf{I_i}$, produced by the converter U/I

$$\mathbf{Y_i} = \mathbf{I_i} / \mathbf{U_i} = (\mathbf{K_{conv}} \cdot \mathbf{U_i}) / \mathbf{U_i} = \mathbf{K_{conv}}, \quad (6)$$

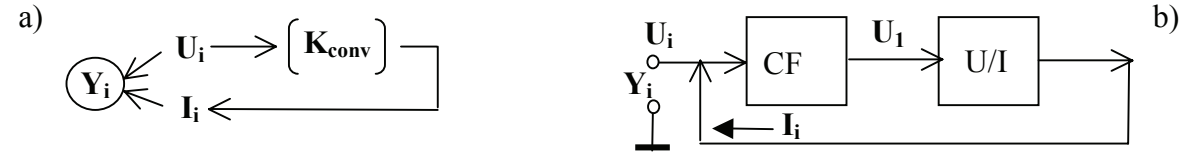


Fig. 3. The conversion algorithm of information (a) and the block-diagram of the U-SPQ (b).

As follows from (6), the U-MS are destined for reproduction of PQ with character of admittance and possess the stability up to short-circuit regime [9].

It is obvious that, under the principle of duality between the quantities with impedance and admittance character [16], the PQ reproduced by I-MS and U-MS also can be represented in the dual forms in comparison with those from (5) and (6), according to the known relation ($\mathbf{Z} = \mathbf{Y}^{-1}$). However, this affects the mutual influence of SPQ components control: functionally, in order to assure the independent control of the components, it is necessary to assure the directly proportional dependence between the SPQ components and the conversion factor $\mathbf{K_{conv}}$.

2. The type of coordinates for SPQ representation. The importance of this criterion is determined by the necessity of using MS as RE in two types of measurements of passive quantities: measurements in Cartesian coordinates and measurements in polar coordinates. So, for SPQ represented in Cartesian coordinates, the conversion factor $\mathbf{K_{conv}}$ must assure an independent control of active and reactive components, according to (1)

$$\mathbf{K_{conv.Z}} = N_R + \mathbf{j}N_X, \ \mathbf{K_{conv.Y}} = N_G + \mathbf{j}N_B; \quad (7)$$

for SPQ represented in polar coordinates, an independent control of the module and phase of SPQ, according to (2)

$$\mathbf{K_{conv.Z}} = N_Z \exp(\mathbf{j}\varphi), \ \mathbf{K_{conv.Y}} = N_Y \exp(\mathbf{j}\psi). \quad (8)$$

3. The type of poles for connection of SPQs in an external circuit. According to this criterion, MS can have one pole connected to ground in the internal mode (asymmetric connection, in MS-As) (Fig. 4a) or the both poles being free (symmetric connection, in MS-S) (Fig. 4b). MS-As have the simplest structures determined by using of the operational amplifiers with internal ground connection. In MS-S, the symmetric character of the PQ poles is obtained by means of precision differential amplifiers [17], which complicates the device structure. In spite of the more complicated structure, MS-S has found the practical application in some circuits and some measuring cases of PQ, where the RE with symmetric poles are necessary (for example, in the bridge circuits).

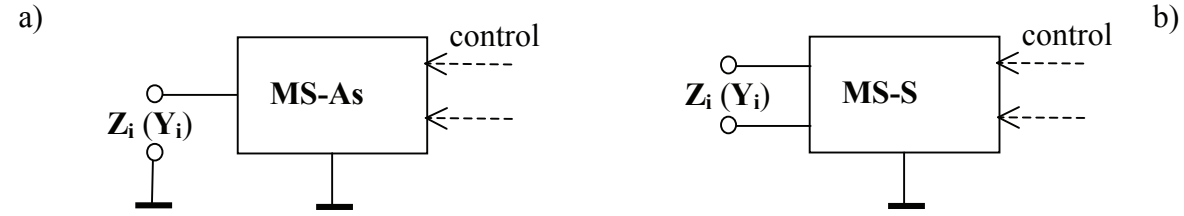


Fig. 4. The functional representation of MS-As (a) and MS-S (b).

In accordance with the mentioned above criteria, there has been defined a class that consists of eight types of MS-A (Table 1).

3. Synthesis of the MS-A structures

3.1. The formal–structural method of synthesis

For synthesis of the MS-A internal circuit, the FSM of synthesis [12] has been used. The method ensures obtaining of the internal structure of MS-A at the level of internal circuit according to its type from Table 1. The method includes the following steps:

1. *Determination of the initial requirements on MS-A.* This is made according to the use of features through designation of the type of MS from Table 1. It includes:
- determination of the primary input quantity (I-type or U-type);
- determination of coordinates type for representation of SPQ (C-type or P-type);
- existence of the ground connection of the input pole of MS-A (As-type or S-type).

2. *Elaboration of the Information's Conversion Diagram (ICD) inside the structure of MS-A.* At this step, the sequence of operations at the physical quantities in MS-A structure is represented in the form of a diagram. The used operations are the conversion of voltages and currents, which can be made with functional links, based on electronic circuits or on other devices (Table 2). The next types of information conversion operations are usually used at the synthesis of MS-A structures [18]:
- voltage into current, or current into voltage conversion,
- voltage into voltage conversion with constant or variable conversion factor,
- voltages sum or subtraction,
- introduction of the fixed or variable phase angle,
- multiplication or division of the quantities,
- functional conversion of the quantities.

3. *Elaboration of the MS-A functional structure.* The step includes the elaboration of the internal structure of MS-A at the level of functional units according to the presented in p. 2 ICD. The available functional units (Fig. 5) correspond to the information conversion operations from Table 2 and, for MS-A from Table 1; they represent the units based on operational amplifiers (OAs) [17].

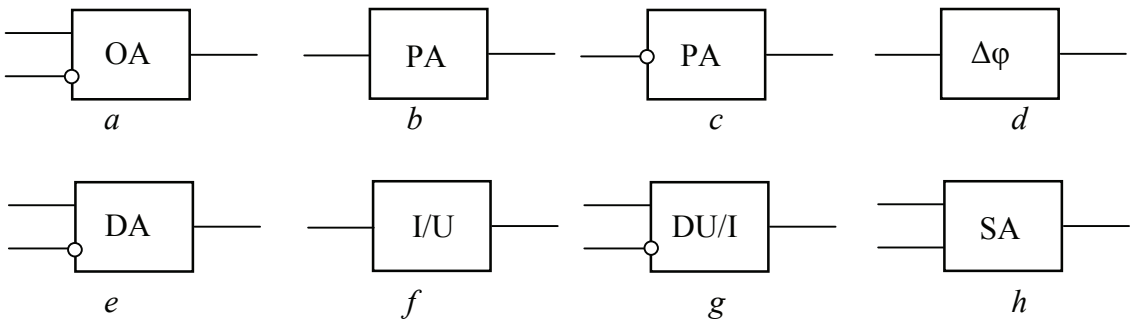


Fig. 5. The functional units used in the MS-A structures: *a* is operational amplifier; *b,* programmable non-inverting amplifier; *c,* inverting programmable amplifier; *d*, phase shifter; *e,* differential amplifier; *f,* current-to-voltage converter; *g,* differential current-to-voltage converter; *h,* summator amplifier.

The other types of functional units can be used, by necessity, for the synthesis of other types of MS-A.

4. *Elaboration of the internal circuit of MS-A*. It is the last step of the synthesis, in which the functional units from MS-A structure are substituted with circuits that implement the respec-

tive functions. The units based on OAs are used in the low frequencies MS-A for implementation of the above mentioned operations. They assure high accuracy of operations [17]. In the high-frequency circuits, where the OA accuracy decreases considerably, the functional units of MS-A can be implemented on the basis of transistors or other components.

Table 2. Typical conversion operations of the information in ICD.

Graphical representation	Essence of operation
A → **B**	Quantity **A** into quantity **B** conversion
A → **B** (**C** ↓)	Functional conversion of quantity **A** into quantity **B** under the influence of quantity **C**
N_A ⇢ **A**	Control of quantity A with numerical quantity N
A → ⊕ ← **B**; ⊕ ↓ **C**	Sum of quantities A and B with obtaining of the result C **C = A + B**
A → ⊖ ← **B**; ⊖ ↓ **C**	Subtraction of quantity B from quantity A with obtaining of the result C **C = A - B**
A → ⊗ ← **B**; ⊗ ↓ **C**	Multiplication of quantities A and B with obtaining of the result C **C = A · B**
A → ⊘ ← **B**; ⊘ ↓ **C**	Division of quantities A and B with obtaining of the result C **C = A / B**
A, **B** → [**F**] → **C**	Obtaining of quantity C functionally dependent on the quantities **A**, **B** **C = F (A, B)**

3.2. Synthesis of U-MS-P-As

In order to confirm the possibility of the presented above FSM for MS-A synthesis, we shall examine as example the process of synthesis of the internal circuit of voltage-commanded (U-) MS for reproduction the SPQ with asymmetrical connection (-As) in the polar coordinates (-P) [19]. This process is executed in the following order of actions.

1. *Determination of the initial conditions*. Since the voltage $\mathbf{U_i}$ is the primary input quantity, SPQ will have the character of admittance $\mathbf{Y_i}$ [9]. The necessity of the SPQ representation in polar coordinates defines the expression for the reproduced admittance $\mathbf{Y_i}$ as type (2). So, MS must assure the independent control of the components Z_i, ψ_i and possess one pole with the internal connection to the ground.

2. *Elaboration of ICD*. Reproduction of the required type of simulated admittance $\mathbf{Y_i}$ may be obtained in the following order of operations:

- variation of the voltage $\mathbf{U_i}$ module in order to control the SPQ module Y_i (op. 2, Table 2),
- variation of the voltage $\mathbf{U_i}$ phase in order to control the phase of SPQ ψ_i (op. 2, Table 2),
- conversion of the obtained voltage into input current $\mathbf{I_i}$ (op. 2, Table 2),

- introduction of the current $\mathbf{I_i}$ into input circuit of MS, where it, together with the voltage $\mathbf{U_i}$, reproduces the simulated admittance $\mathbf{Y_i}$ (op. 7, Table 2).

The information conversion diagram (Fig. 6a) assures the execution of these operations in a desired order.

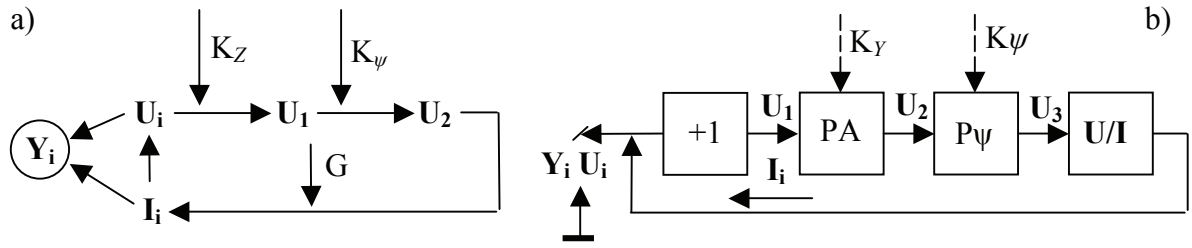


Fig. 6. ICD for the structure synthesis (a) and the internal structure (b) of U-MS-P-As.

3. *Elaboration of the MS functional structure*. The internal structure for implementation of the above determined operation from ICD contains the functional units (Fig. 6b):

- The programmable amplifier PA with the transfer factor K_{PA} for control of the module Y_i. The voltage at its output constitutes: $\mathbf{U_1} = K_{PA} \cdot \mathbf{U_i}$;
- The programmable phase shifter PΦ with transfer factor $\mathbf{K_{P\Phi}}=1\cdot exp\ (\mathbf{j}\psi_i)$, which produces the voltage $\mathbf{U_2} = \mathbf{K_{P\Phi}} \cdot \mathbf{U_1} = K_{PA}\ exp\ (\mathbf{j}\psi_i) \cdot \mathbf{U_i}$ at its output;
- The voltage-to-current converter U/I with conversion factor G for conversion of the voltage $\mathbf{U_2}$ in the current $\mathbf{I_i}$. The current on its output: $\mathbf{I_I} = G\cdot\mathbf{U_2}=G\cdot K_{PA}\ exp\ (\mathbf{j}\psi_i)\cdot\mathbf{U_i}$.

With evidence of the unit transfer functions, the virtual admittance reproduced by MS at its poles constitutes [19]

$$\mathbf{Y_i} = \mathbf{I_i} / \mathbf{U_i} = G \cdot K_{PA}\ exp\ (\mathbf{j}\psi_i) \equiv Y_i\ exp\ (\mathbf{j}\psi_i). \quad (9)$$

As follows from (9), the reproduced admittance $\mathbf{Y_i}$ is represented in polar coordinates and assures an independent control of module Y_i through the variation in the transfer factor K_{PA} of the programmable amplifier PA and of the phase ψ_i through the variation in the phase angle ψ_i introduced by the programmable phase shifter PΦ.

4**.** *Elaboration of the internal circuit of U-MS-P-As*. The functional units based on OAs [17] were used for the practical implementation of the structure from Fig. 6b (Fig. 7).

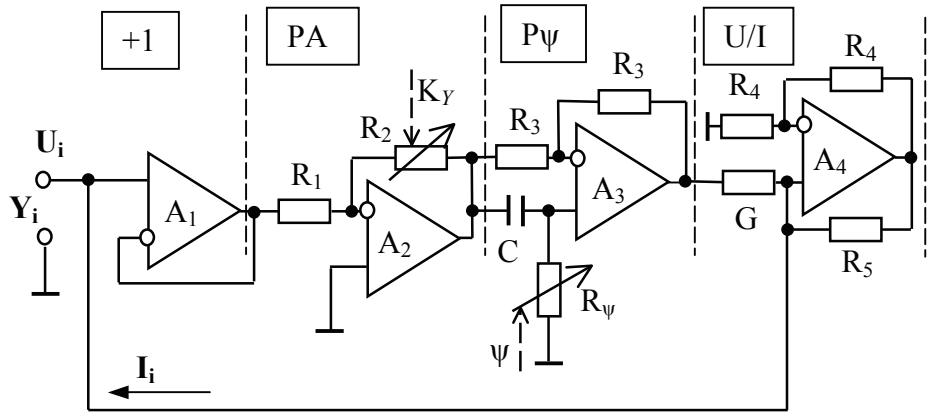


Fig. 7. The internal circuit of U-MS-P-As.

The voltage repeater A_1 was used in order to exclude the influence of the units input impedance to the reproduced admittance. The PA and PΦ units from structural scheme are im-

plemented on the basis of OAs A_2, A_3; the current-to-voltage converter U/I, on the basis of OA A_4. With regard for the known expressions [17] for the transfer factors of the MS units, the expression for reproduced admittance $\mathbf{Y_i}$ becomes

$$\mathbf{Y_i} = 2G \cdot {R_2}/{R_1}\, exp\left[\mathbf{j} arctg\left(\omega R_\psi C\right)\right]. \qquad (10)$$

As follows from (10), the variation of R_2 results in the control of module Z_i; the variation of R_ψ, in the control of phase ψ_i of the reproduced admittance without mutual influence. According to (10), the variation domains of module Y_i and phase ψ_i of the reproduced admittance are

$$Y_i = 0 \div 2G \cdot R_{2max}/R_1,\ \psi_i = 0 \div 180°. \qquad (11)$$

In order to assure the range of control of admittance phase $\psi_i = 0 \div 360°$, what is essential for reproducing the admittance with any character, in the simulator contains a programmable phase shifter with the phase angle control range $0 \div 360°$ [7]. It can be formed of two consecutive links of the phase shifter used in the presented simulator [20] or of the next type, [21], for example.

3.3. Other types of MS-A with asymmetrical connection

The other types of simulators from Table 1 were synthesized in the same sequence of actions used at the U-MS-P-As synthesis. For these ones, we shall present only the synthesis results with the most important elements of this process.

The simulator I-MS-P-As [22]**.** The device presents a current-commanded MS (**I-**) for reproduction of simulated impedances in polar coordinates (-P) with asymmetrical connection (-As). I-MS-P-As can be used as a RE in the polar-coordinate impedance meters, which assure a direct measurement of the impedance module and phase [23]. The information conversion diagram (Fig. 8a), the functional structure (Fig. 8b), and the internal circuit (Fig. 8c) of this simulator were obtained similarly to the procedure used at the U-MS-P-As synthesis.

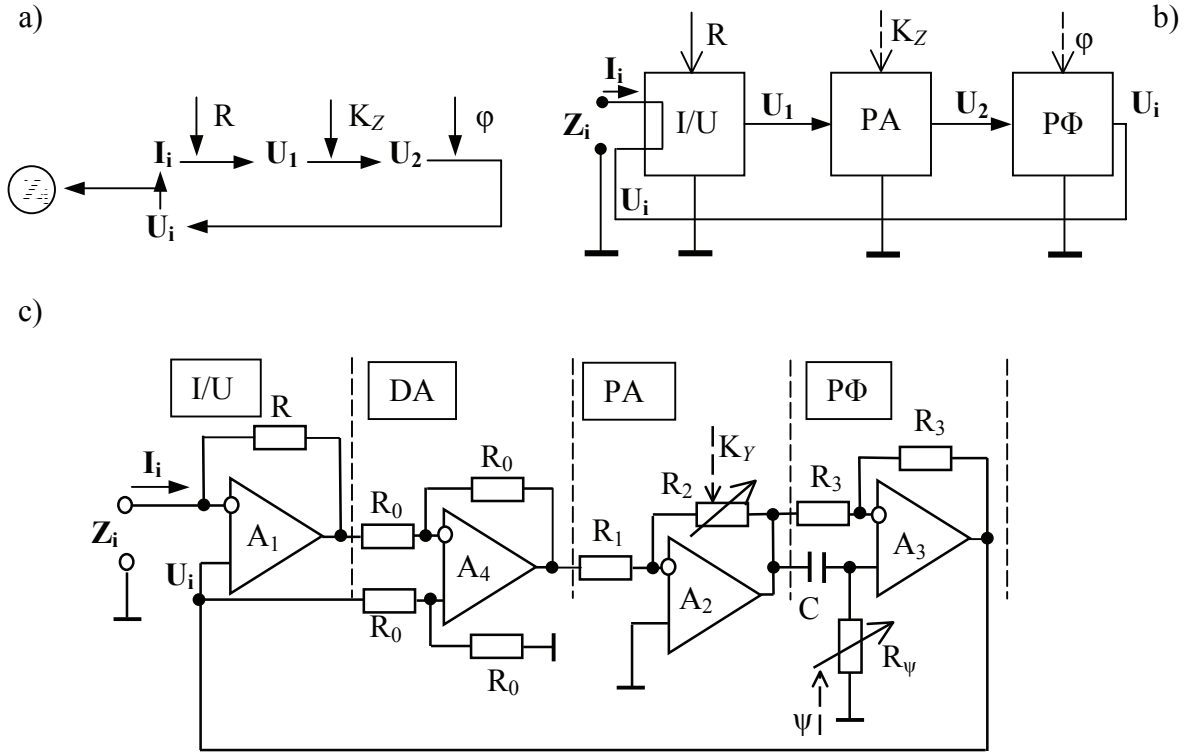


Fig. 8. ICD (a), the functional structure (b), and the internal circuit of I-MS-P-As (c).

The units known from the U-MS-P-As synthesis (a programmable amplifier PA and a programmable phase shifter PΦ) and the other units based on OAs, such as a current-to-voltage converter IU and a differential amplifier DA [8], were used in the process of the functional structure and internal circuit elaboration of I-MS-P-As. The function of DA unit is the neutralization of the common feedback loop action, which appears upon introduction of the voltage $\mathbf{U_i}$ into input circuit and leads to the expression for simulated impedance $\mathbf{Z_i}$ distortion.

The impedance $\mathbf{Z_i}$ reproduced by the simulator can be obtained by the following sequence of conversions (Fig. 8a):

- conversion of the input current $\mathbf{I_i}$ into the intermediate voltage $\mathbf{U_1}$ by means of the converter IU;
- control of its module and phase by means of the programmable amplifier PA and the programmable phase shifter PΦ, respectively;
- application of the voltage $\mathbf{U_i}$ into the input circuit of the IU converter.

The reproduced impedance $\mathbf{Z_i}$ consists of [22]

$$\mathbf{Z_i} = 2R \cdot {R_2}/{R_1}\, exp\left[\mathbf{j} arctg\left(\omega R_\varphi C\right)\right]. \tag{12}$$

Expression (12) completely confirms the correspondence of the characteristics of the synthesized I-MS-P-As to the initial requirements. The control of reproduced impedance components is assured similarly to the control in U-MS-P-As: the variation of R_2 assures the module Z_i control; the variation of R_φ, the phase φ_i control. The domains of value variation are

$$Z_i = 0 \div 2R \cdot R_{2max}/R_1,\ \varphi_i = 0 \div 180° \tag{13}$$

and can be enlarged in the same way as in U-MS-P-As.

The simulators U-MS-C-As [24] *and I-MS-C-As* [25]. They are destined for reproduction of SPQ in Cartesian coordinates with asymmetrical connection and differ from previous by the type of input quantities: in U-MS-C-As the input quantity is the voltage $\mathbf{U_i}$, and in I-MS-C-As – the current $\mathbf{I_i}$. Thus, in accordance with (1), (3), and (4), U-MS-C-As assures the reproduction of the simulated admittances, and I-MS-C-As – reproduction of the simulated impedances with independent control of the active and reactive components. The same method of information conversion (Fig. 9a) and the same circuit for creating the active and reactive components (Fig. 9b) were used in either type of these simulators. The intermediate voltage $\mathbf{U_1}$ is transmitted through two channels: the first not involving phase change for the active component; the second, with a phase angle of 90^o to create the reactive component. The control of the voltage values is assured by the programmable amplifiers PA1 and PA2 for the components control in the two channels. The signal introduced later into the input circuit is obtained from the sum of the voltages $\mathbf{U_R}$, $\mathbf{U_X}$ (Fig. 9). The control of SPQ active component in the domain of positive and negative values is assured by regulation of the transfer factor K_R of amplifier PA1 in the values range ($-K_{Rmax} \div 0 \div +K_{Rmax}$) and the same control for the reactive component is achieved by regulation of the transfer factor K_X of amplifier PA2 in the values range ($-K_{Xmax} \div 0 \div +K_{Xmax}$).

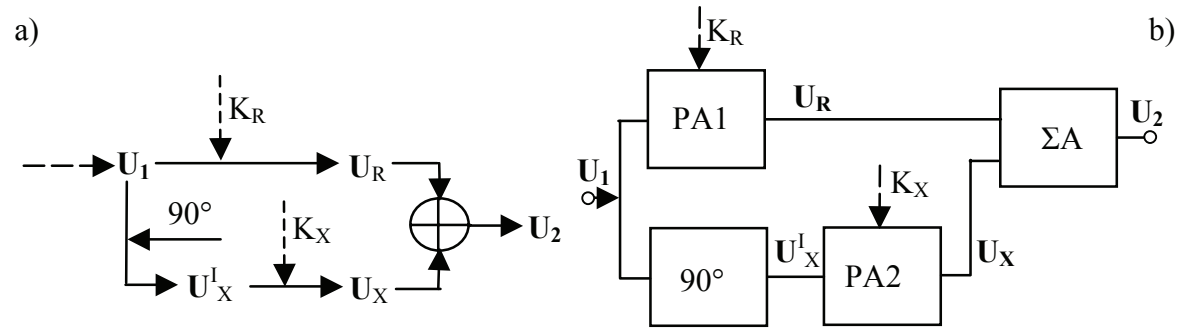


Fig. 9. ICD (a) and the circuit portion (b) for creating the SPQ components in Cartesian coordinates MS.

The ICD (Fig. 10a), the functional structure (Fig. 10b) and the internal circuit (Fig. 10c) for U-MS-C-As [24] were elaborated by the method used for the synthesis of previous simulators. The quantity reproduced by the simulator at its poles has an admittance character and represents

$$\mathbf{Y_i} = G(K_G + \mathbf{j}K_B), \tag{14}$$

where G is the conversion factor of UI converter.

The ICD (Fig. 11a), the functional structure (Fig. 11b), and the internal circuit (Fig 11c) were obtained for I-MS-C-As in a similar manner. Its reproduced quantity possesses a character of impedance and represents

$$\mathbf{Z_i} = R\,(K_R + \mathbf{j}K_X). \tag{15}$$

As follows from (14), the reproduced impedance $\mathbf{Z_i}$ is represented in Cartesian coordinates and possesses the separate control of components: the active component is controlled by variation of K_R; the reactive component, by variation of K_X.

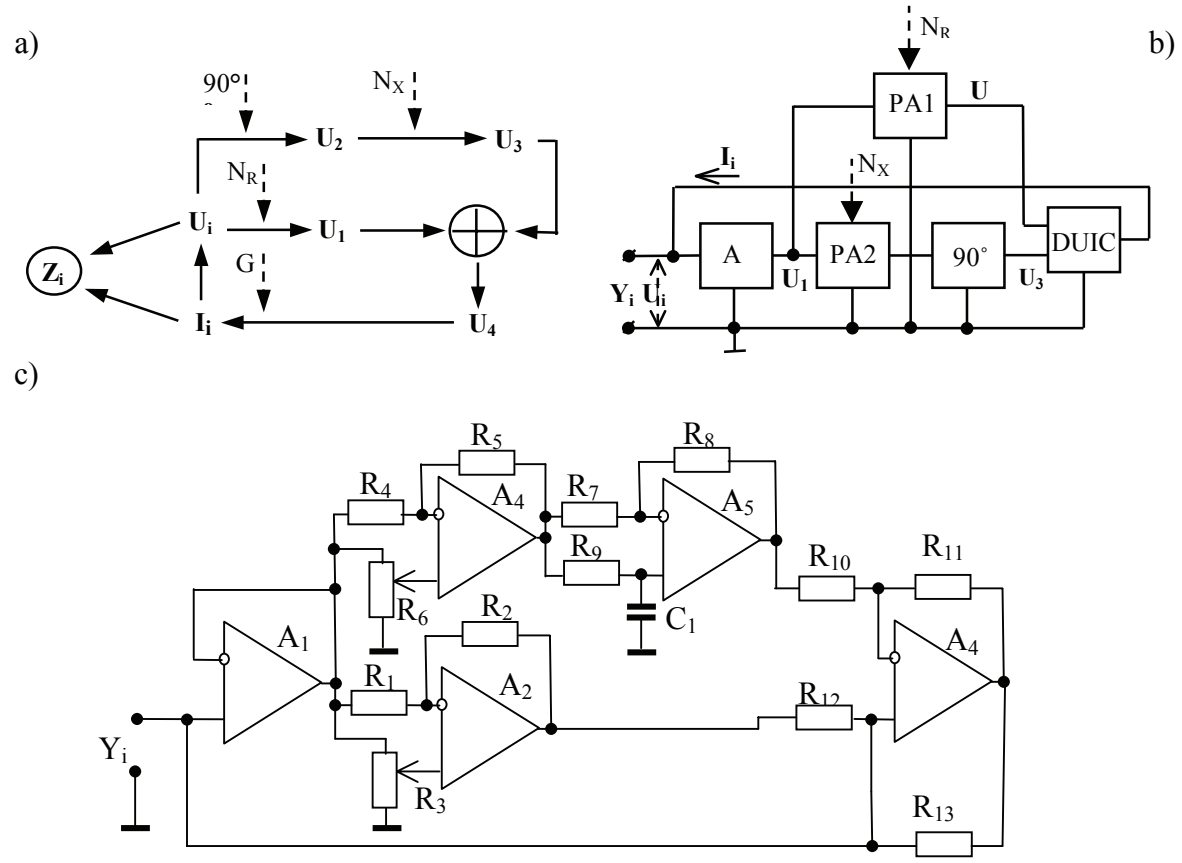


Fig. 10. The ICD (a), the functional structure (b), and the internal circuit of U-MS-C-As (c).

3.4. MS-A for reproduction of PQs with symmetrical connection

In some cases of using of MS as RE, it is necessary to reproduce SPQ with the both poles free for connection in the external circuit (Fig. 4b).

It refers, for example, to using the MS in measurement circuits with serial simulated resonance for grounded PQ measurement, in DC or AC bridges, etc. MS-A for reproduction of SPQ with symmetrical connection of the poles (MS-S) for this purpose was elaborated (Table 1).

The main aim of the elaboration of the MS-S structures was the use of OAs with traditional power supply by the grounded sources [17]. In order to obtain the symmetrical nature of SPQ poles, the circuits SM-S assure compensation of the common mode input impedance

of OAs by creating the reproduced quantity depending only on the differential input quantity. The intermediary differential quantity obtained in the result of subtraction of the primary common mode input quantities is converted into Cartesian (Fig. 6) or into polar coordinates (Fig. 8) and, being introduced into the entry circuit of SM-S, assures the reproduction of the respective SPQ.

The U-MS-C-S simulator [26, 27]. This simulator (Fig. 12) assures the reproduction of the voltage-commanded SPQ represented in Cartesian coordinates with symmetrical poles. A potentiometric differential amplifier [17] based on OAs A1, A2, and A3 was used in order to create an intermediate voltage proportional to the difference of the common mode input voltages and to eliminate the influence of OA input impedance on SPQ. In order to obtain the active and reactive components and in order to assure an independent control in Cartesian coordinates, the intermediate voltage $\mathbf{U_d}$ was converted as shown in Fig. 9.

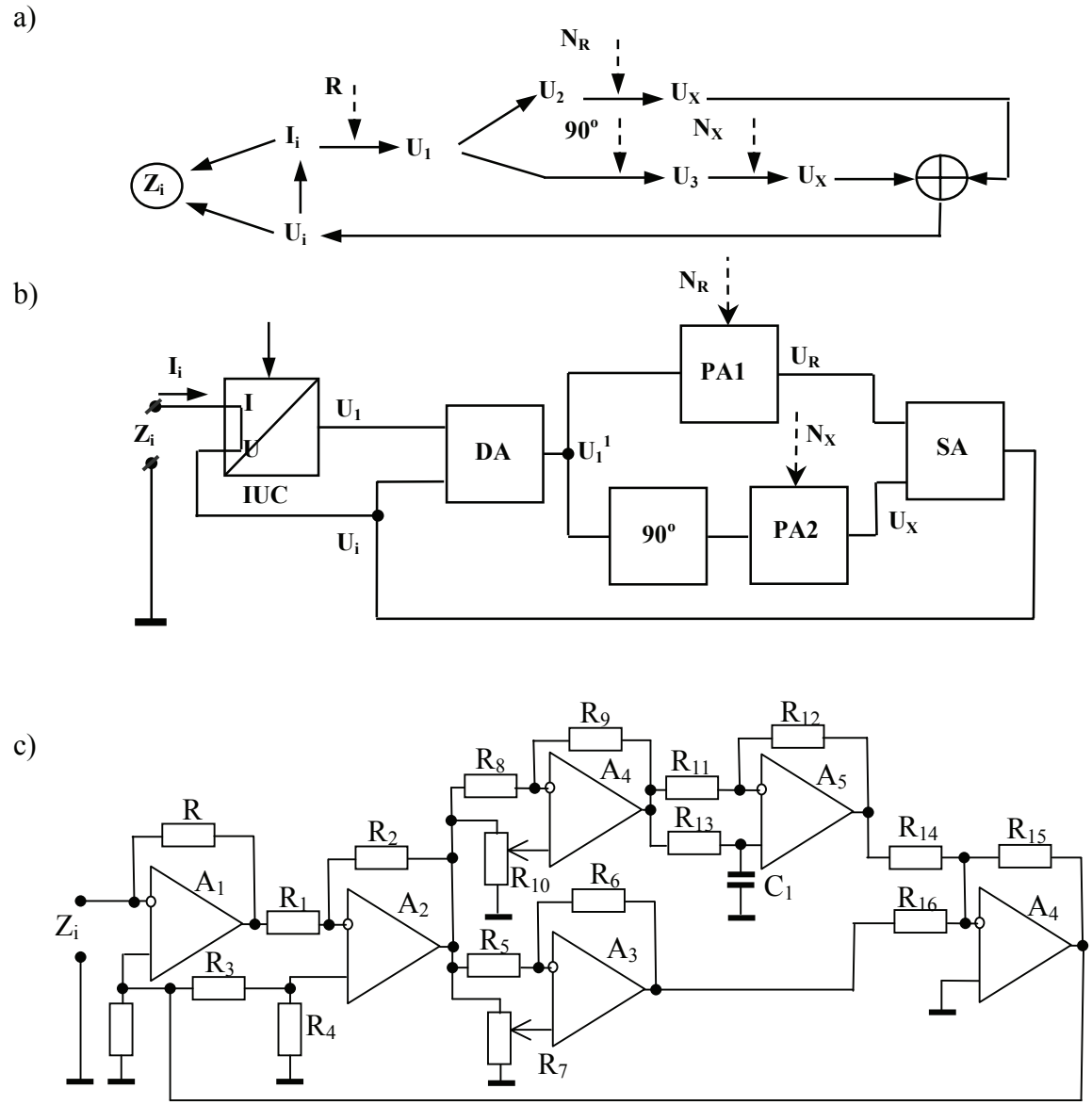


Fig. 11. The information conversion diagram (a), the functional structure (b), and the internal circuit of I-MS-C-As (c).

The currents $\mathbf{I_i}$, $\mathbf{-I_i}$ introduced in the input circuit of MS, are obtained by two voltage-to-current converters UI1, UI2 with equal and reverse transfer factors. As a result of ultrahigh input impedance of differential amplifier [17], the reproduced admittance is

$$\mathbf{Y_i} = \mathbf{I_i} / \mathbf{U_{id}} = (R_{conv})^{-1}(K_1 + \mathbf{j}K_2). \quad (16)$$

The symmetrical character of the admittance reproduced by the simulator is assured by exclusion of the influence of the common mode input voltages on the entry current $\mathbf{I_i}$.

The U-MS-P-S simulator [28] (Fig. 13). The input circuit is similar to the one from U-MS-C-S. In order to assure an independent control of admittance components reproduced by the simulator in polar coordinates, the intermediate voltage $\mathbf{U_d}$ is consecutively passed through the programmable amplifier PA and programmable phase shifter PΦ, similarly to U-MS-P-As, and it is used for generation of the input currents $\mathbf{I_i}$, $\mathbf{-I_i}$ by the voltage-to-current converters UI1, UI2. The admittance reproduced by U-MS-P-S is [28]

$$\mathbf{Y_i} = 2G \cdot K_Z \, exp\left[\mathbf{j} arctg\left(\omega R_\psi C\right)\right]. \quad (17)$$

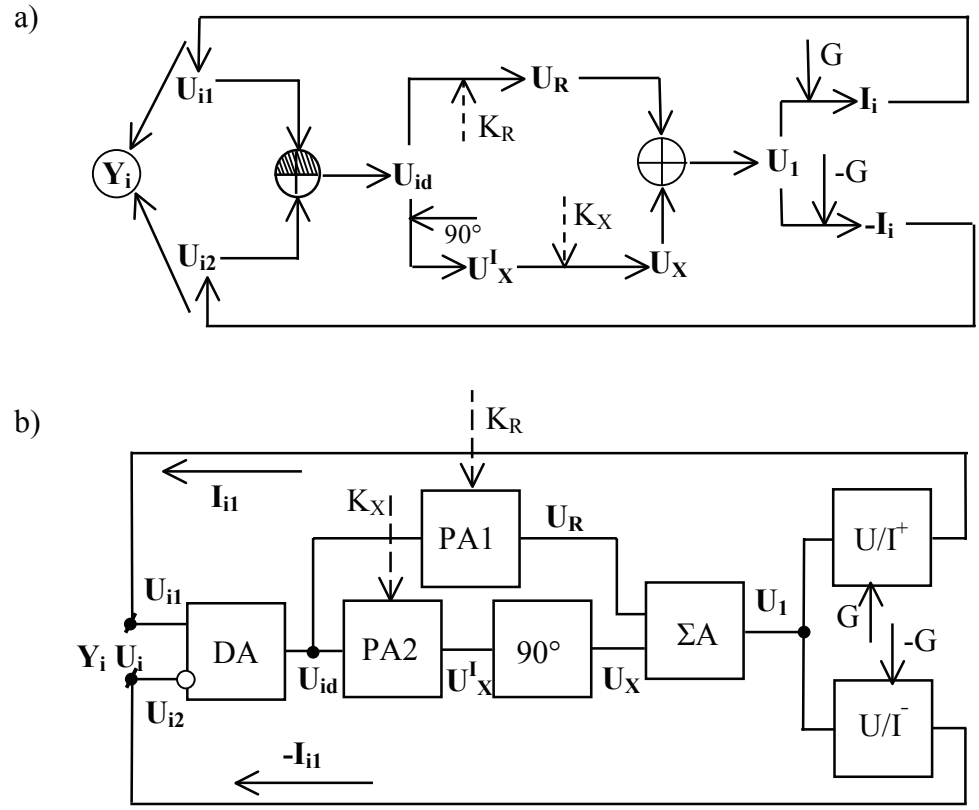


Fig. 12. The information conversion diagram (a) and the functional structure (b) for U-MS-C-S.

The current-commanded MS-S simulators. As follows from Table 1, there are two types of current-commanded impedance simulators with symmetrical poles: for reproduction of the impedances represented in Cartesian coordinates (I-MS-C-S) [29] and for reproduction of the impedances represented in polar coordinates (I-MS-P-S) [30].

Synthesis of these MS-As was carried out by the same method which was used for the synthesis of previous simulators. In order to obtain an intermediate quantity convenient for conversion in polar or Cartesian coordinates (with voltage character), a current-to-voltage converter with differential inputs DIU was used. In order to express the reproduced impedance in needed coordinates, the intermediate voltage $\mathbf{U_d}$ was converted similarly to conversion

in I-MS-P-As and I-MS-C-As structures by using the circuit with transfer factor **K** expressed in Cartesian (Fig. 8) or polar coordinates (Fig. 11).

The voltage $\mathbf{U_d}^1$ destined to be applied in the input circuit was changed into differential voltage by using an inversing amplifier A3 with unit transfer factor.

The information conversion diagrams, the functional structures and the internal circuits of I-MS-P-S and I-MS-C-S are represented in Figs. 14 and 15, respectively.

The impedances reproduced by the simulators are

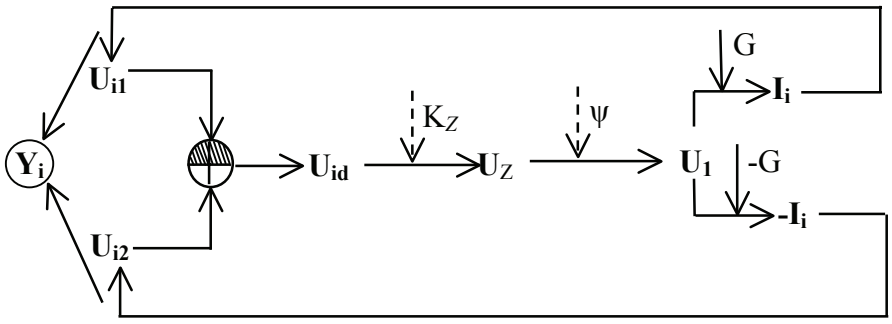
- for I-MS-C-S [29]

$$\mathbf{Z_i} = \mathbf{2}R_{conv}\,(K_1 + \mathbf{j}K_2); \tag{18}$$

- for I-MS-P-S [30]

$$\mathbf{Z_i} = 4R \cdot K_Z \, exp\left[\mathbf{j}arctg\left(\omega R_\varphi C\right)\right]. \tag{19}$$

a)



b)

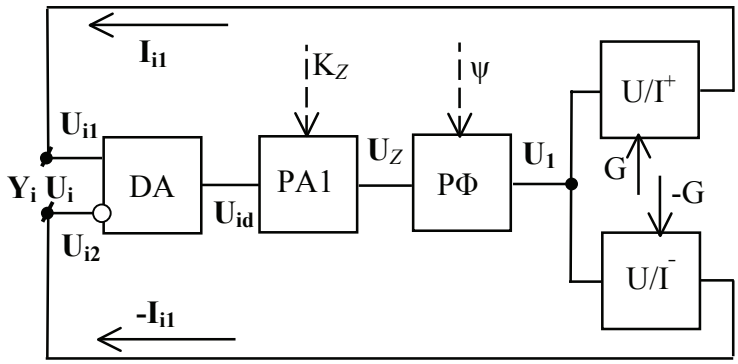


Fig. 13. The information conversion diagram (a) and the internal structure (b) for U-MS-P-S.

As follows from (18), for I-MS-C-S, $\mathbf{Z_i}$ is expressed in Cartesian coordinates and assures an independent control of the active component (by regulation of the factor K_1) and of the reactive component (by regulation of K_2). It follows from (19) that the impedance reproduced by I-MS-P-S is expressed in polar coordinates and assures an independent control of the module (by regulation of the resistor R_2) and of the phase (by regulation of R_φ).

In order to obtain the expressions for the reproduced impedances correct from algorithmic point of view, at the implementation of circuits based on OAs (Figs. 14b, 15b), the differential amplifiers which neutralize the general feedback were used [8].

4. Experimental confirmation of the results

The authenticity of results of synthesis of the simulators from Table 2 was determined by the simulation in the program PSPICE and by experimental modelling (simulation) of devices in the laboratory.

The studied simulators were connected as components of the circuit with simulated resonance (CSR) [2] together with a complex PQ with the known value. The balancing of the CRS was carried out by regulation of MPS components to achieve the state of total resonance, which corresponds to the zero value of the unbalance signal. In order to study the I-MS, the series CSR was used; to study U-MS, the parallel CSR, which is determined by the type of stability of these simulators [2].

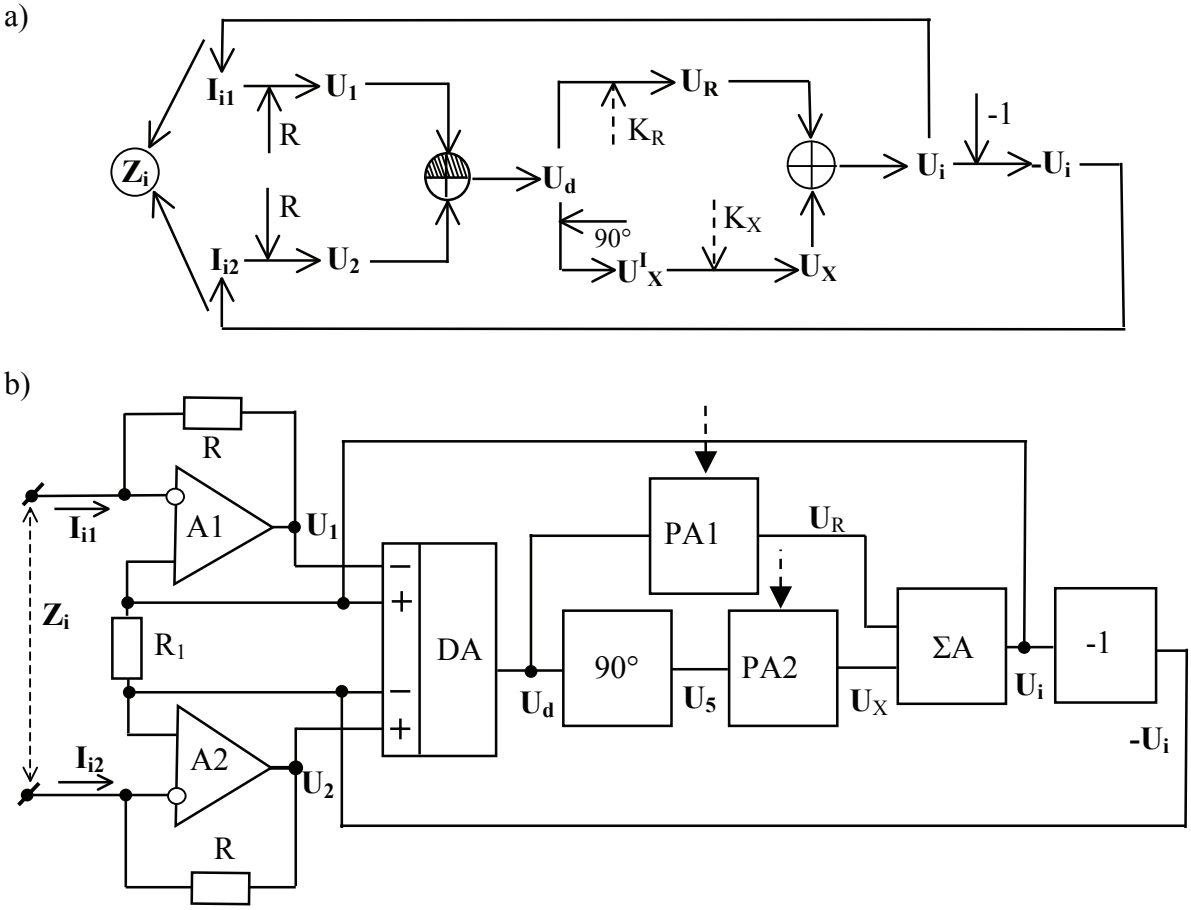


Fig. 14. The information conversion diagram (a) and the functional structure (b) for I-MS-C-S.

The studying of I-SM-C-As [31] as a component of the series CSR (Fig. 16a) is an example of modelling in the program SPICE of the simulators.

The simulation was carried out for two cases: for the measured impedance with an active character and with a complex character. The harmonic generator V1 with frequency F_S = 1 KHz and voltage U_S = 10 V for supplying CSR with measuring signal and the external resistor R_2 = 1 kΩ as a model of measured resistance were used.

The state of resonant circuit was determined by comparing the phase of the unbalance signal $\mathbf{U_{de}}$ (on p. 2 of the circuit) with the phase of the reference signal from the generator $\mathbf{U_{ref}}$ (on p. 1) and of the signal $\mathbf{U_{de}}$ amplitude by means of an XSC1 oscilloscope. In the case of modelling the complex MPS, the balancing of the CSR was made in two stages: the balancing on the active component by variation of the resistor R_9 and the balancing on the reactive component by variation the resistor R_{21} until the total resonance condition ($\mathbf{U_{de}}$ = 0) is achieved. The end of the first stage is the achievement of a phase angle of 90° between the signals $\mathbf{U_{de}}$ and $\mathbf{U_{ref}}$ (Fig. 18b); the end of the second phase is the obtaining of the state $\mathbf{U_{de}}$ = 0 (Fig. 18c).

The modelling results have demonstrated full compliance of the SPQ component values calculated from expression (15) to those obtained experimentally [31]. The same result was obtained by experimental laboratory modelling of the device.

The same methods were used for experimental testing of the other simulators presented in Table 1. The experimental results obtained for SPQ values in the state of total CSR resonance fully correspond to the theoretical values determined by the method of admittances matrix [7].

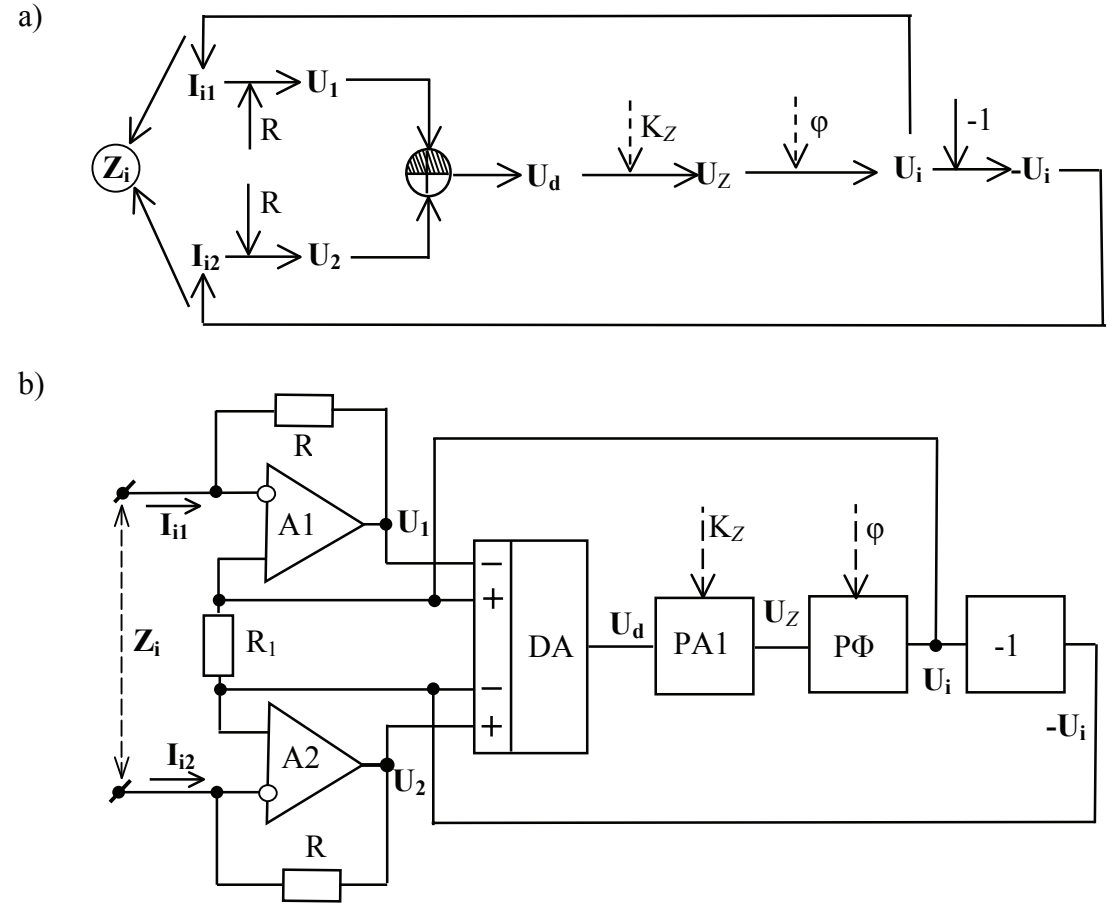
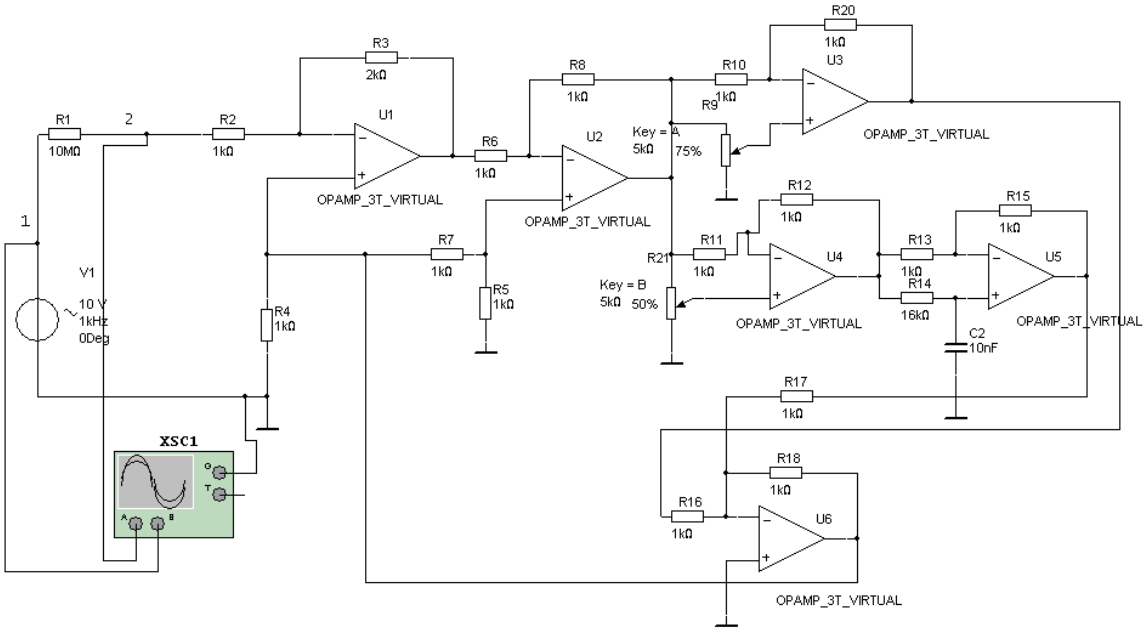


Fig. 15. The information conversion diagram (a) and the functional structure (b) for I-MS-P-S.

A result interesting in view of the possible practical application was obtained for simulators with symmetric connection poles (MS-S, Table 1). The experimental results confirmed **the floating character of the admittances reproduced by U-MS-CS and U-SM-SP**. This was demonstrated by the lack of influence on the state of resonance of any point-to-mass connection for the parallel CRS. The obtained result confirms the possibility to use these simulators as REs for reproduction of the floating measures of PQswith components commanded in Cartesian (U-MS-C-S) or polar (U-MS-P-S) coordinates.

The results of research of I-commanded (I-MS-CS, I-SM-SP) simulators reproduced by the presented circuits (Figs. 14b, 15b) showed that the MPQs reproduced by these simulators, though possessing symmetric character, are not floating. This is accounted by the usage of OAs supplied from grounded sources [17], which leads to the asymmetric nature of voltage $\mathbf{U_{id}}$ introduced in the entry circuit of MS and therefore of the reproduced impedance. To obtain the floating simulated impedances in these simulators, it is recommended to apply the methods of OA feeding from floating energy sources [17], or the internal compensation of the common mode component of entry impedance by means of the feedback loops.

a)



b) c)

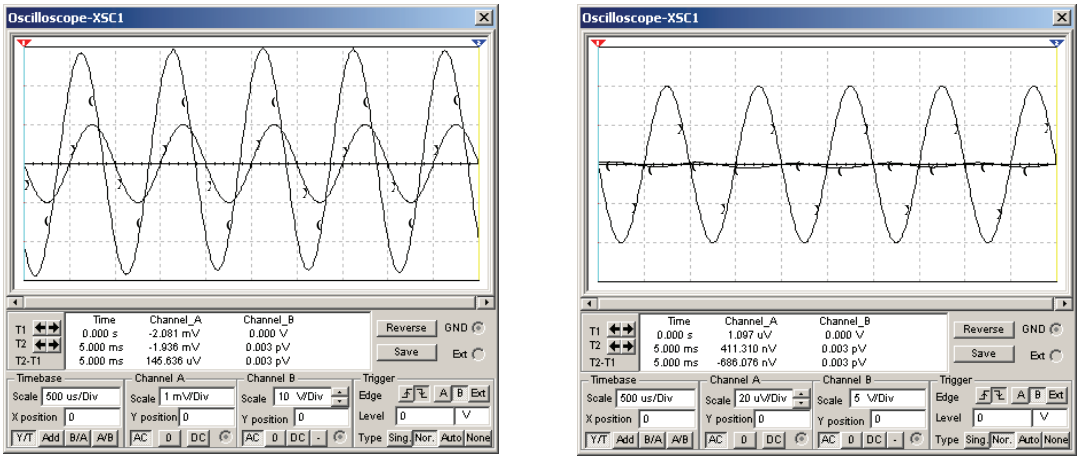


Fig. 16. The series CSR to study I-SM-C-As modelled in the program SPICE (a), the CSR balancing results at the active component (b) and at the reactive component (c) (° is the signal $\mathbf{U_{de}}$, ^{x} is the signal $\mathbf{U_{ref}}$).

5. Stability of MS-A

Practical application of MS-A is possible only upon fulfilment of the condition of guaranteed assurance of stability of devices at the variation of internal and external factors of the measuring circuits in the entire allowable range of values. The following factors which can potentially influence the MS-A stability were selected from the analysis of the real conditions of measuring circuits:

- the MS-A type according to Table 1,
- the character and the values of components of SPQ,
- the character and the value of equivalent external impedance (Fig. 4),
- the characteristics of OAs.

So, depending on the factor which causes the loss of device's stability, the problem of MS-A stability must be examined from three points of view:

1. Assurance of stability on direct current (DC).
2. Assurance of stability at high frequencies.
3. Assurance of the functional stability.

Stability on DC. The loss of stability on DC in the circuits based on OAs is caused by prevalence of positive feedback over negative one at direct current [17]. In order to maintain the guaranteed stability on DC, it is necessary to assure the summary negative character of DC feedback in each link (local feedback) and in the entire circuit (general loop of feedback).

Stability at the high frequencies. The absence of this type of stability is manifested by excitation of some MS links or of the entire circuit because of non-ideality of frequency characteristics of OAs [17]. In order to prevent the excitation at high frequency of the links or of the whole circuit, the optimal correction of OAs through known methods and the correct distribution of the unit frequency ranges at their cascade connection with common feedback loops are required.

The functional stability. It is determined by the summary character of the general feedback at the operation frequency on connection of the MS in an external circuit. It depends on MS type (Table 1), on the character and values of the reproduced impedance, and on the external equivalent impedance connected to MS poles [8]. The loss of this type of stability is manifested through circuit excitation at working or near-working frequency, a regime similar to the one which appears in a generator. If, in order to ensure the first two types of stability, we must observe only of the general principles of OA use [17], the assurance of the functional stability is conditioned by imposing some restrictions on the type of the used simulator, on the type of the measurement circuit, and on the values of the above impedances. Since the assurance of stability presents one of the obligatory conditions of practical use of MS [32], the determination of the functional stability conditions for every type of MS-A is absolutely indispensable.

As it is known [9], the MS stability depends firstly on the primary input quantity (**I** or **U**) (p. 2.1): I-MS maintains stability to the no-load regime; U-MS, to the short-circuit regime. However, this condition is too general and it does not adequately determine the conditions of MS usage as components of measuring circuits. In order to obtain these conditions, we shall make the analysis of MS-containing circuits based on the classical theory of stability of the circuits based on OAs with feedback loops.

In order to estimate the conditions of the functional stability, we shall take advantage of the Nyquist criterion in application to the OA-containing circuits [17]. As is known, the circuits with feedbacks save the stability if the critical point (-1, +j0) is located to left of the hodograph of the transfer characteristics on the feedback loop, when the frequency changes from f = 0 up to f $\rightarrow\infty$ (Fig. 17).

Thus, in order to estimate the stability conditions for the circuits in accordance to the Nyquist criterion, it is necessary to determine its loop gain factor $\boldsymbol{\beta A}$ and to examine it in the neighbourhood of the critical point (-1, + j0) in coordinates Re ($\boldsymbol{\beta A}$), Im ($\boldsymbol{\beta A}$)**.** The condition of functional stability for the MS circuit looks like [17]

$$\boldsymbol{Re}\,[H_{io}] > -1, \tag{20}$$

where H_{io} – loop gain factor of the MS circuit.

The model of I-MSI [8] to determine the condition of stability is represented in Fig. 18a. Through the equivalent differential amplifier with the gain factor **K,** the portion of the circuit of I-MS limited by the dotted line (Fig. 2b) is designated.

The model for U-MS (Fig. 18b) contains the amplification link **K** and the current-to-voltage converter based on OA A1 connected into the feedback loop. $\mathbf{Z_e}$ stands for the equivalent external impedance connected to poles of MS.

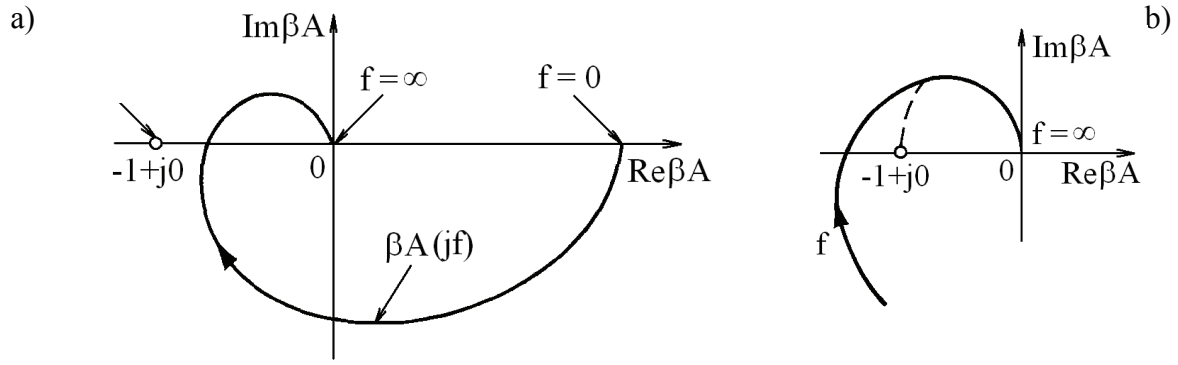


Fig. 17. Hodograph of the loop gain factor for stable (a) and unstable (b) circuits.

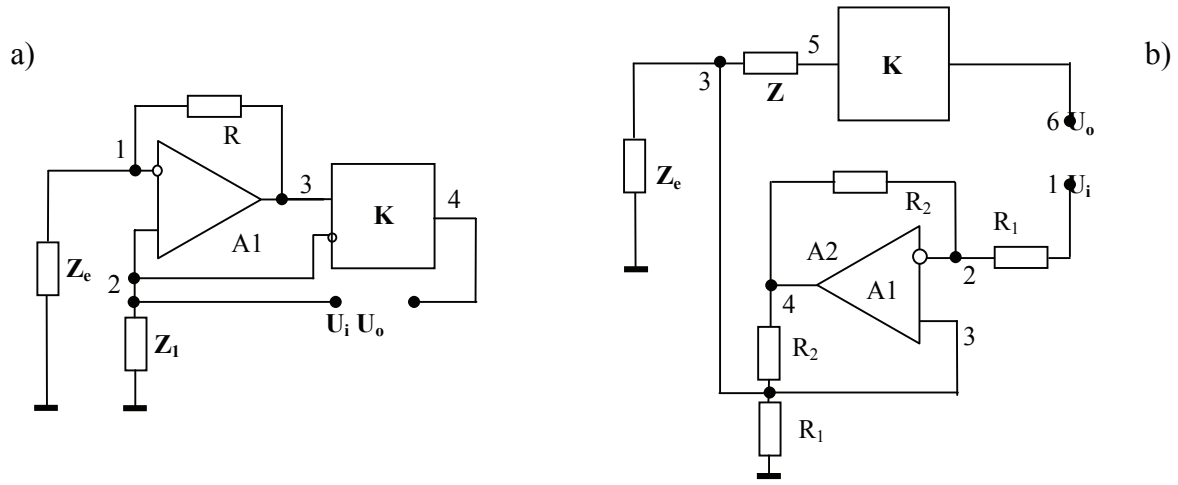


Fig. 18. The models of I – MS (a) and U-MS (b) for stability analysis.

In order to determine the stability conditions, in either case, the expressions for the transfer factor in loop $\mathbf{H_{io}}$ of the circuits were obtained [33]. They are

- for I-MS (Fig. 18a)

$$\mathbf{H_{io}} = \mathbf{U_o}/\mathbf{U_i} = \mathbf{K}\,R / \mathbf{Z_e};\ (21)$$

- for U-MS (Fig. 18b)

$$\mathbf{H_{io}} = \mathbf{U_o}/\mathbf{U_i} = -\,\mathbf{K}\,\mathbf{Z_e} / R_1.\ (22)$$

The transfer factors for I-MS-C and I-MS-P can be obtained by substituting the expressions for **K** in Cartesian (7) and in polar (8) coordinates in (21). In the same way, the transfer factors for U-MS-C and U-MS-P can be obtained (Table 3).

Table 3. The expressions for stability analysis of MS-As.

Type of MS	The transfer factor $\mathbf{H_{io}}$	
	The stability condition	
	for MS-C: $\mathbf{K} = K_1 + \mathbf{j}K_2$	for MS-P: $\mathbf{K} = K_Z \exp(\mathbf{j}\varphi)$
I-MS: $\mathbf{H_{io}} = \mathbf{K}\cdot R / \mathbf{Z_e}$	$\mathbf{H_{io}}=R(K_1+\mathbf{j}K_2)/(R_e+\mathbf{j}X_e)$ (3.1)	$\mathbf{H_{io}}=(R/Z_e)K_Z \exp(\mathbf{j}(\varphi-\varphi_e))$ (3.3)
	$\lvert\mathbf{Z_e}\rvert^2 > R_e R_{X\,max} + X_e X_{X\,max}$ (3.2)	$Z_e > R\,K_Z$ (3.4)
U-MS: $\mathbf{H_{io}} = -\,\mathbf{K}\,\mathbf{Z_e} / R_1$	$\mathbf{H_{io}}=-(K_1+\mathbf{j}K_2)(R_e+\mathbf{j}X_e)/R_1$ (3.5)	$\mathbf{H_{io}}=-K_Z \exp(\mathrm{j}(\varphi-\varphi_e))Z_e/R_1$ (3.6)
	$K_{1max} R_e + K_{2max} X_e < R_1$ (3.7)	$Z_e < R_1/K_Z$ (3.8)

The problems of the error analysis for impedance simulators on the basis of operational amplifiers were investigated in [6]. As follows from [6], the greatest influence on the error of the given value of the reproduced impedance is exerted by the factors of non-ideality of the operational amplifier and, in particular, by the limited value of gain factor and its frequency dependence.

For minimization of this component of the error, the operating frequency of measuring signal must not strongly exceed the frequency of pole of the OA characteristic [17]. The acceptable error $\delta \leq 0.1\%$ was obtained at the operating frequency F = 100 Hz.

6. Conclusions

1. The MS-A forms a class of devices that can be used as REs to measure the impedance using the method of simultaneous comparison. Compared to the classical RE, they have such advantages as the possibility of reproduction of reference quantities with any character, the possibility of expression of the reproduced quantities in the Cartesian or polar coordinates, and the possibility of independent control of the reproduced quantity components.
2. The MS-A in Cartesian coordinates ensure reproduction of SPQs **located throughout the entire complex plane** ±R, ±**j**X (±G, ±**j**B for admittances) **with possibility of independent control of active and reactive components and with smooth variation of the character**. The MS-A in polar coordinates ensures reproduction of the same type PQs in coordinates system Z, φ (Y, ψ for admittances) with independent control of the module and phase values in bands Z, $Y = (0 \div \max)$, φ, $\psi = (0 \div 360°)$.
3. The MS-A classification by relevant criteria established a class of eight devices, which are distinguished by the possibilities of application in various devices for the impedance measurement. Each type of MS-A possesses relevant criteria, such as the primary input quantity (voltage or current), type of the coordinates for the reproduced quantity expression (Cartesian or polar), and type of the connection poles for external circuit (asymmetric or symmetric poles).
4. For synthesis of the MS-A internal structure, the formal–structural method was used. The method ensures the synthesis of the simulator's internal structure in accordance to the requirements determined by the conditions of use. The structures for all eight types of MS-A were synthesized using this method. The results of experimental study of these devices fully confirm their characteristics obtained in theoretical mode.
5. The practical application of MS-A is conditioned by the stability of devices. Three aspects of the stability were determined: the high frequency stability, the direct current stability, and the functional stability. The first two types of stability are ensured through compliance with standard principles for the use of OAs. The functional stability is dependent on the type of MS-A, on the character and value of the reproduced quantity, and on the character and value of the external equivalent impedance connected to the poles of MS.
6. The analysis of functional stability through the Nyquist criterion has determined the restrictions in use for each type of MS-A. Thus, the current-commanded simulators exhibit stability up to the no-load regime and are recommended for use in series circuits with simulated resonance; the voltage-commanded simulators possess stability up to the short-circuit regime and are recommended for use in parallel circuits with simulated resonance. The analysis has determined the feasible conditions of absolute stability for each type of MS-A.

References

[1] V. Nastas and M. Scinteianu, Proc. of the 8th Int. Conf. OPTIM 2002, Brasov, 3, 683, (2002).
[2] V. Nastas and M. Scinteianu, Meridian ingineresc, Chisinau, Tehnica-Info, 2, 70, (2001).
[3] A. Nordeng, Impedance synthesizer. European Patent 0656542 A2, (1995).
[4] Joe C.P. Liu, Chi K. Tse, Franki N.K. Poon, M.H. Pong, and Y.M. Lai, Proc. ECCTD, Cork Ireland, 1, 173, (2005).
[5] V. Nastas, Proc. of the 6th Int. Conf. ICMCS 2009, Chisinau, 1, 61, (2009).
[6] V. Nastas and A. Cazac, Meridian ingineresc, Chisinau, Tehnica-Info, 3, 49, (2003).
[7] V. Nastas, Meridian ingineresc, Chisinau, Tehnica-Info, 3, 23, (2009).
[8] V. Nastas, Proc. of the 9th Int. Conf. OPTIM 2004, Brasov, 4, 89, (2004).
[9] F. Bening, Negative viderstande in elektronischen schaltungen, Veb Verlag, Berlin, 288 p., 1971.
[10] V. Nastas and P. Nicolaev, Proc. of the 6th Int. Conf. ICMCS 2009, Chisinau, 65, (2009).
[11] V. Nastas, M. Scinteianu, and A. Cazac, Proc. of the 3rd Int. Conf. on Microelectronics and Computer Science, Chisinau, 2, 308, (2002).
[12] V. Nastas, I Conferinta Nationala de Telecomunicatii, Electronica si Informatica, Chisinau, 218, (2006).
[13] V. Nastas, Proc. of the 5th Int. Conf. on Microelectronics and Computer Science, Chisinau, 2, 55, (2007).
[14] V. Nastas, Mold. J. Phys. Sci., 1, 3, 28, (2002).
[15] L. Hulsman, Teoriya i raschet activnykh RC-tsepei, Moscow, 240 p., 1973.
[16] V. Sigorsky and A. Petrenco, Algoritmy analiza elektronnykh skhem, Moscow, 608 p., 1976.
[17] J. Dostal, Operational amplifiers, Elsevier SPC, New York, 512 p., 1981.
[18] V. Nastas, Proc. of the 5th Int. Conf. ICMCS 2009, Chisinau, 1, 277, (2007).
[19] V. Nastas and A. Cazac, Admittance converter. Patent MD 2462, (2004).
[20] V. Nastas and A. Cazac, Defazor. Patent MD 2844, Chisinau, (2006).
[21] V. Nastas, Impedance converter. Patent MD 3136, Chisinau, (2007).
[22] V. Nastas and M. Scanteianu, Impedance converter. Patent MD 2130, (2003).
[23] V. Nastas, Impedantmetru. Patent MD 2866, (2006).
[24] V. Nastas, Admitance converter. Patent MD 3111, (2008).
[25] V. Nastas, Impedance converter. Patent MD 3154, (2008).
[26] V. Nastas, Admitance converter. Patent MD 3461, (2008).
[27] V. Nastas and P. Nicolaev, Admitance converter. Patent MD s90, (2009).
[28] V. Nastas, Admitance converter. Patent MD 3173, (2008).
[29] V. Nastas, Analele Universitatii din Craiova, seria: Inginerie Electrica, Craiova, 31, 2, 14, (2007).
[30] V. Nastas, Impedance converter. Patent MD 3689, (2009).
[31] V. Nastas and P. Nicolaev, Proc. of the 2nd Int. Conf. "Telecommunications, Electronics and Informatics", Chisinau, 2, 211, (2008).
[32] V. Nastas, Mold. J. Phys. Sci., 7, 4, 481, (2008).
[33] V. Nastas, V. Moraru, and N. Torita, Proc. of the 2nd Int. Conf. "Telecommunications, Electronics and Informatics", Chisinau, 2, 217, (2008).

TEMPORAL AND SPATIAL VARIABILITY OF TOTAL OZONE CONTENT IN MOLDOVA: SATELLITE RETRIEVALS AND GROUND OBSERVATIONS

A. Aculinin and V. Smicov

Institute of Applied Physics, Academy of Sciences of Moldova, 5, Academiei str., MD-2028, Chisinau, Republic of Moldova; e-mail: akulinin@phys.asm.md
(Received 9 December 2009)

Abstract

The total ozone content (TOC) values retrieved from measurements of spectral backscattered radiation with Total Ozone Mapping Spectrometer (TOMS) instruments onboard of Nimbus-7, Meteor-3 and Earth Probe platforms and with Ozone Monitoring Instrument (OMI) onboard of Aura platform are used to evaluate temporal and spatial distribution of statistical means of TOC in Moldova from 1978 to 2008. Longitudinal and latitudinal variations of monthly, seasonally and yearly means of total ozone content across the territory of Moldova are observed. Differences between statistically mean values of TOC retrieved for two pairs of sites situated at the borders both in latitude and longitude directions are used for analysis. Ranges of differences between retrieved TOC for respective pair of sites in the North-South and for pair of sites in the West-East directions amounted to 25 DU and 10 DU (for monthly means of TOC), and to 4.8 DU and 2.2 DU (for yearly means of TOC). Monthly means of total ozone content for all sites reveal distinctive seasonal variation with maximum on the order of 378 DU in spring and minimum of the order of 289 DU in fall. The climatic norm for yearly means of TOC amounts to 335 DU. The trend of TOC over Moldova is -2% per decade.

1. Introduction

Ozone in atmosphere represents a natural shield against the harmful solar UV radiation which reaches the Earth's surface and exerts unfavorable influence on the human health, on the other biological life, and ecosystems. Height distribution of the atmospheric ozone and total ozone content on regional and global scales are changing as a result of chemical destroying of ozone molecules due to an increase in the ozone-depleting substances generated from natural processes and growing of anthropogenic activities. Another reason of ozone variability consists in the complex dynamical processes taking place in the atmosphere. These reasons are closely connected with each other through the complex processes, such as heat and mass transfer and chemical reactions in atmosphere. It was well established from numerous ground observations and from satellite platforms retrievals that total column ozone content at the middle latitudes (35°N-60°N) has been decreasing for several decades over 1978-2005 [1, 2]. During these decades of observations, a specific trend with ozone depletion was clearly observed. In the period from pre-1980 to 1997-2001, the total ozone content has depleted on an average by about 3% in the northern middle latitudes (35°N-60°N) [1] and averaged for the period 2002-2005 total ozone was about 3% lower than their 1964-1980 values and similar to their 1998-2001 values [2]. Value of ozone depletion also depends on season of the year. In the Northern Hemisphere ozone depletion amounted to 4% and to 2% in winter-spring and in summer-autumn, respectively [1].

In this paper, we study variation of the total ozone content (TOC) in column of atmosphere or total column ozone derived for the particular sites in Moldova by using linear interpolation of gridded multi-annual datasets from satellite observations. Gridded datasets were

retrieved from TOMS measurements onboard of Nimbus-7 (N7), Meteor-3 (M3), and Earth Probe (EP) satellite platforms and from OMI measurements onboard of Aura satellite platform during the period from 1978 to 2008. TOC mean values retrieved from TOMS and OMI measurements are compared with the ground observations carried out with hand-held photometer at the solar radiation monitoring station at the Institute of Applied Physics, Chisinau (Kishinev) from 2004 to 2008. Latitudinal and longitudinal variations of monthly, seasonally, and yearly means of TOC derived by using interpolation of gridded data from satellite observations for particular sites on the territory of Moldova will be presented.

2. Measurement approach

Datasets with daily means of TOC retrieved from observations of the Earth's spectral backscattered radiation by using of TOMS instruments onboard of the N7, M3, and EP satellite platforms during the period from 1978 to 2005 and OMI instrument onboard of the Aura satellite platform during a period from 2004 to 2008 were joined together to create long-term time series of TOC. TOMS is a nadir looking instrument that measures albedo and derives total column ozone from the differential albedo in three pairs of spectral bands in the ultraviolet region. The EP TOMS instrument provides instantaneous field of view or the size of the "footprint" of each measurement of 39 km x 39 km at nadir. A more detailed description of TOMS instrument and about it final product can be found elsewhere [3]. The Ozone Monitoring Instrument is a Dutch-Finnish ozone monitoring instrument that is flying on NASA's Aura Mission, part of the Earth Observation System (EOS). Data are obtained from retrieving procedures of satellite images of the Earth with high spectral and highest spatial resolution with small pixel size which is of 13 km x 24 km. More detailed description of the instrument, data processing, calibration and characterization can be found elsewhere [4]. Theoretical basis of the OMI ozone product algorithm for deriving the total column ozone from spectral radiances and irradiances is based on the TOMS ver.8 algorithm applied to OMI data, and it is described in detail elsewhere [4-7]. Total column ozone values at specific sites across the territory of Moldova are derived from linear interpolation of gridded multi-year TOMS and OMI data. Interpolated TOC values to these site's coordinates are used to analyze latitudinal, φ and longitudinal, λ_o variation of monthly, seasonally, and yearly mean values of TOC on the territory of Moldova.

Ground-based observations of total ozone content in a column of atmosphere are regularly carried out with hand-held narrowband filter photometer MICROTOPS II ozonometer (Solar Light Co) at the solar radiation monitoring station (φ=47.00°N, λ_o=28.82°E, h=205 m a.s.l.). The station is situated in an urban environment and installed at the building roof at the Institute of Applied Physics, Chisinau, Moldova. The ground station is in operation since July 2003. TOC values were derived from the ratios of direct solar ultraviolet radiances simultaneously measured at 3 discrete wavelengths 305.5, 312.5, and 320-nm within the UV-B spectral range. Column ozone data are retrieved by using optical differential absorption technique. Detailed description of the instrument and measurement algorithm can be found elsewhere [8]. As a rule, measurements of the total ozone content are carried out during midday hours, when the small values of air mass *m* are the case. Observations were carried out for air masses *m* up to values *m* = 3–3.5 (for AM and PM). MICROTOPS II ozonometer allows for making measurements of the total ozone content with an accuracy of ~2% relative to the Dobson and Brewer spectrophotometers [8-10].

3. Data analysis

Regular observations of the total column ozone at the ground station at the Institute of Applied Physics in Chisinau have been started by Atmospheric Research Group (ARG) since

Table 1. List of sites under examination on the territory of Moldova.

Site	(φ, λ_o)-coordinates	
Briceni	48.35°N	27.7°E
Chisinau	47.0°N	28.82°E
Vulcanesti	45.68°N	28.4°E
North border (N)	48.12°N	28.82°E
South border (S)	45.95°N	28.82°E
West border (W)	47.00°N	28.06°E
East border (E)	47.00°N	29.61°E

July 2003 [11]. Respective results of TOC ground observations were processed and compiled into the datasets, and hereafter referred to as ARG data. Multi-year series of TOC values acquired from TOMS measurements at the N7, M3, and EP satellite platforms and OMI measurements at the Aura satellite platform are utilized. TOMS and OMI gridded datasets data are available on-line through the NASA Goddard Space Flight Center archive in ASCII form.

Typical grid resolution (latitude x longitude) of datasets is of 1° x 1.25° - for TOMS data [12] and of 1° x 1° for OMI data [13]. Suitable column ozone values specific to pre-selected sites in Moldova are computed by using linear interpolation of gridded TOMS and OMI data. These datasets are used to assess variability of TOC across the Moldova. Analysis of TOC variability is carried out by taking into account interpolated data for two pairs of sites: situated at the borders in the North–South (N-S) and West-East (W-E) directions. Interpolated values of TOC for intermediate sites (see Table 1) will be also used. For Moldova, the distances between the north and the south borders, and the west and the east borders are ~340 km and ~260 km, respectively. In turn, this corresponds to respective angular spans of $\delta\varphi$ ~3° for latitude and $\delta\lambda_o$ ~3.5° for longitude. Statistics of monthly $\langle X\rangle_m$, seasonally $\langle X\rangle_s$, and yearly $\langle X\rangle_Y$ means of TOC are computed by using daily means $\langle X\rangle_d$ as basic values from combined TOMS and OMI time-series and from ARG ground datasets. In spite of the short range distances between borders of Moldova, amounting to a few degrees, daily means $\langle X\rangle_d$ revealed significant variability along (φ, λ_o)-coordinates. As an example of daily TOC variability, retrieved values of $\langle X\rangle_d$ from the series of satellite observations at specific coordinates for three days are shown in Figs. 1 and 2. There were used observations from TOMS and OMI ozone datasets for set of following days: August 20, 2008 was characterized as clear cloudless day with low column ozone content $\langle X\rangle_d$= 291 DU; December 1, 1999 was characterized as a day with the extremely low value of TOC with $\langle X\rangle_d$=209 DU ever observed in the long-time series and February 17, 2008 was characterized as a clear day with high value of TOC $\langle X\rangle_d$= 426 DU. The step of interpolation of satellite data along each of (φ, λ_o)-coordinates amounts to 0.5°. It can be clearly seen that there are noticeable contrasts in TOC values retrieved from satellite observations along the latitude and longitude directions. These contrasts must be taken into account, i.e., in modeling of daily UV doses and UV Index at the sites on the territory of Moldova.

In order to study the spatial variability of total column ozone in Moldova, we selected two pairs of sites situated just at the borders in the N-S and, respectively, W-E directions with (φ, λ_o)-coordinates presented in Table 1. Daily means of TOC $\langle X\rangle_d$ as basic values were retrieved from long-time series of TOMS and OMI observations and then linearly interpolated at coordinates of these particular sites. Adequate monthly, seasonally, and yearly means were computed on the basis of daily means $\langle X\rangle_d$. To make assessments of TOC variability across the territory of Moldova, the equation of difference Δ is introduced. It is defined as a difference between respective daily (d), monthly (m), seasonal (s), and yearly (Y) means of TOC observed at pre-selected pairs of border's sites in the N-S and W-E directions as follows (with adequate indexes): $\langle X(N)\rangle_k$ - $\langle X(S)\rangle_k$ and $\langle X(W)\rangle_k$ - $\langle X(E)\rangle_k$, and these differences are designated as Δ_k(NS) and Δ_k(WE), where index k= {d,m,s,Y}.

The spread of differences Δ_d between daily means of TOC $\langle X\rangle_d$ observed at respective border's sites in the N-S and W-E directions shows large variability. For OMI observations during a period of 2005-2008, the differences Δ_d showed variability ranged from – 50 DU to

50 DU for N-S and from -40 DU to 25 DU for the W-E directions. These margins of difference Δ_d may be considered as rare extreme values because of high nonuniformity of ozone layer occurring along the N-S and W-E directions for some daily TOC observations. The frequency of occurrence of differences Δ_d between daily means of TOC values retrieved at pairs of sites in these directions is shown in Fig. 3. It can be clearly seen that broad distribution function of differences Δ is most of all specific to the N-S direction.

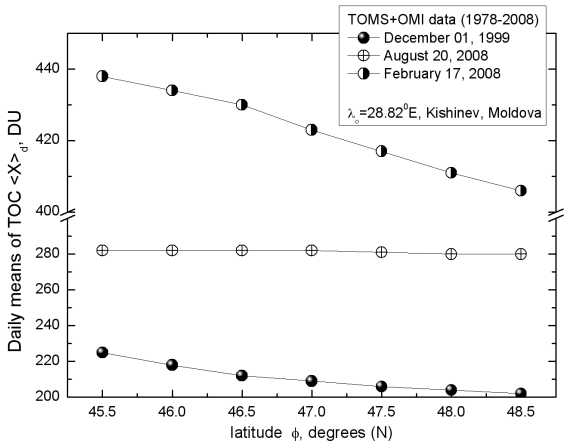


Fig. 1. Latitudinal variation of daily means TOC $<X>_d$ retrieved from TOMS and OMI observations for set of days: 20/08/2008, $<X>_d$=291 DU (clear day); 01/12/1999, $<X>_d$=209 DU (extremely low TOC); 17/02/2008 $<X>_d$=426 DU (clear day, high TOC).

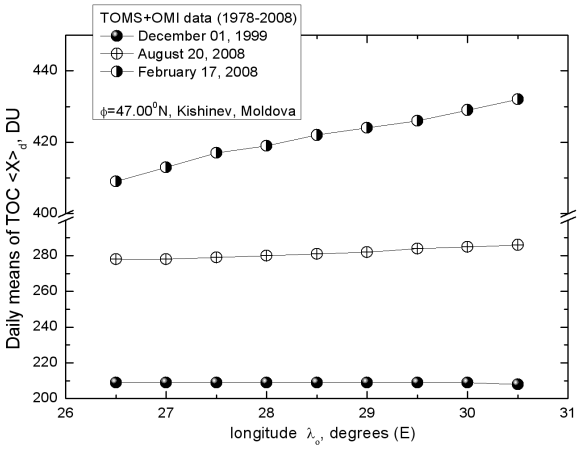


Fig. 2. Longitudinal variation of daily means TOC $<X>_d$ retrieved from TOMS and OMI observations for set of days: 20/08/2008, $<X>_d$=291 DU (clear day); 01/12/1999, $<X>_d$=209 DU (extremely low TOC); 17/02/2008 $<X>_d$=426 DU (clear day, high TOC).

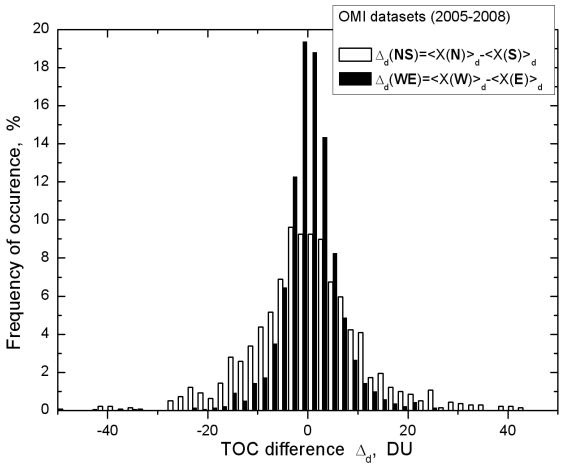


Fig. 3. Frequency of occurrence (in %) of differences Δ_d (in DU) between daily means of TOC, $<X>_d$ retrieved from OMI datasets for each of two pairs of border sites in the N-S and W-E directions (see Table 1). Period of observations: 2005-2008.

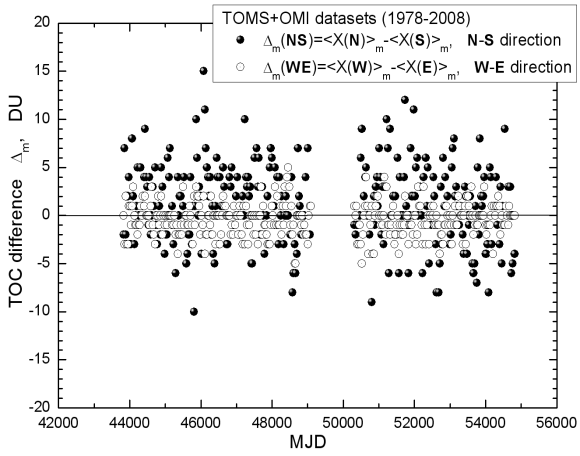


Fig. 4. Latitudinal and longitudinal variability of difference Δ_m (in DU) between monthly means of TOC, $<X>_m$ retrieved from TOMS and OMI datasets for each of two pairs of border sites in the N-S and W-E directions, Δ_m(NS) and Δ_m(WE) (see Table 1). Period of observations: 1978-2008.

Latitudinal and longitudinal variability of difference Δ_m between monthly means of TOC, $<X>_m$ retrieved from combined TOMS and OMI datasets for two pairs of sites in the

N-S and W-E directions is shown in Fig. 4. Data sets were joined together from TOMS observations from 1978 to 2004 and OMI observations from 2005 to 2008 to generate extended series of observations. Differences Δ_m for these directions are noted as follows: Δ_m(NS) and Δ_m(WE). It can be clearly seen that value of spread of differences Δ_m in the N-S direction is larger than in the W-E direction for synchronous observations; these values obey the inequality $|\Delta_m(NS)|<15$ DU and $|\Delta_m(WE)|< 5$ DU, respectively. Figure 4 shows that range of variability of monthly means of TOC in the N-S direction prevails over that in the W-E direction.

The seasonal character of latitudinal (in the N-S direction) and longitudinal (in the W-E direction) variability of derived differences Δ_{MY} between multiyear monthly means of TOC, $\langle X\rangle_{MY}$ at two pairs of sites is shown in Fig. 5. Values of $\langle X\rangle_{MY}$ were obtained from multiyear averaging of monthly means of TOC $\langle X\rangle_{m;}$ $\langle X\rangle_m$ were derived from combined TOMS and OMI multiyear datasets of daily means of TOC during the period of observations from 1978 to 2008. One can clearly see the presence of seasonal variability of Δ_{MY}, in particular, TOC difference for pairs of border sites in the N-S direction is positive with $\Delta_{MY}(NS) < 3$ DU for January-August period and negative one for September-December period with $\Delta_{MY}(NS) \sim -1$ DU. The difference of TOC for pair of border sites in the W-E direction is positive with $\Delta_{MY}(WE) \sim 1$ DU for June-August; for other months, it is negative with $\Delta_{MY}(WE) > -1.5$ DU.

The variability of multiyear monthly means of TOC $\langle X\rangle_{MY}$ at the Briceni, Chisinau and Vulcanesti sites situated along the N-S direction is shown in Fig. 6. The ground observations carried out at the ARG station in Chisinau from 2004 to 2008 are also shown in Fig. 6. It is clearly seen that the seasonal variability has a distinct maximum of TOC in March-April with $\langle X\rangle_{MY} \sim 378$ DU and a minimum of TOC in October-November with $\langle X\rangle_{MY} \sim 289$ DU.

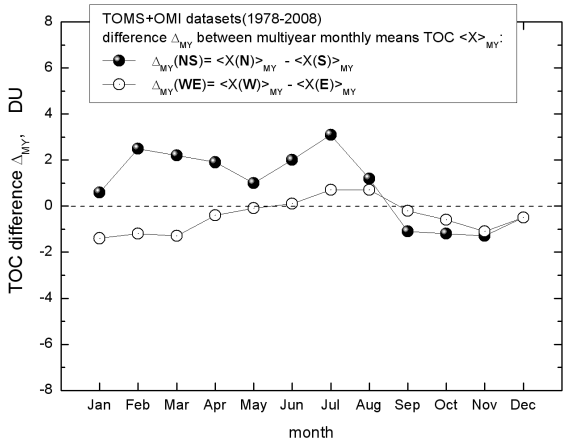


Fig. 5. Latitudinal and longitudinal variability of difference Δ_{MY} (in DU) between multiyear monthly means of TOC, $\langle X\rangle_{MY}$ derived from TOMS and OMI multiyear datasets for two pairs of border sites in the N-S and W-E directions. Period of observations: 1978-2008.

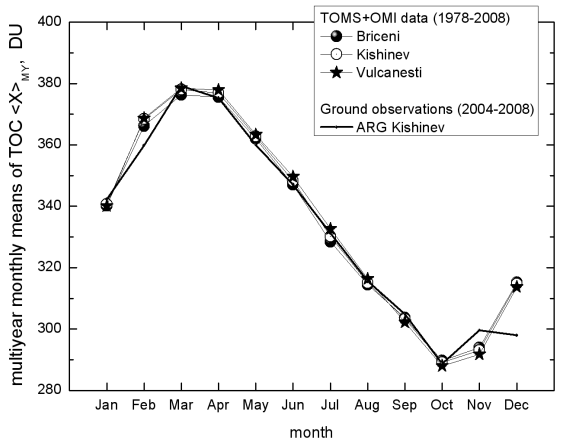


Fig. 6. Variability of multiyear monthly means of TOC $\langle X\rangle_{MY}$ retrieved from TOMS and OMI datasets at the sites Briceni, Chisinau, and Vulcanesti during the period of observations from 1978 to 2008 and from ground observations at the ARG station in Chisinau from 2004 to 2008.

The variability of zonal monthly means of TOC $\langle X\rangle_m$ derived from multiyear OMI observations in the northern hemisphere at belt for latitudes from 45°N to 50°N and ARG ground observations in Chisinau are shown in Fig. 7. One can clearly see a fairly good resemblance between belt data from OMI observations and local measurements at ARG ground station. Meanwhile, there is an appreciable discrepancy with value on the order of ~15-30 DU between OMI belt and ARG observations for data regarding the winter season.

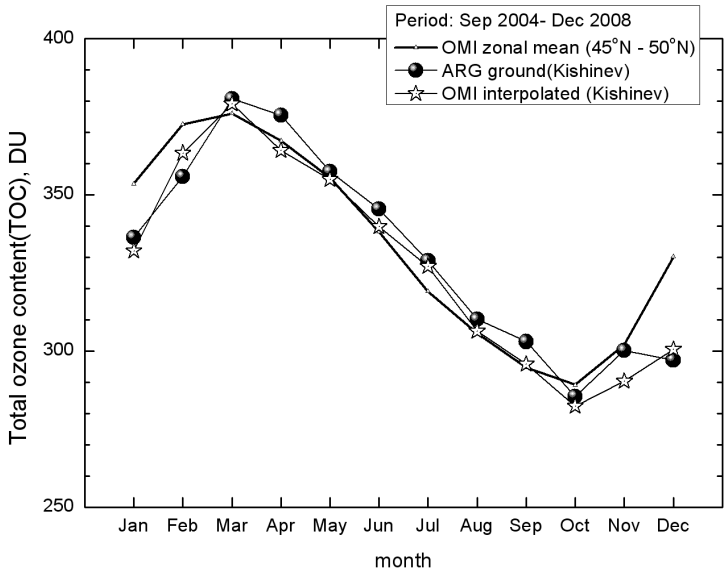


Fig. 7. Variation of zonal monthly means $\langle X \rangle_m$ derived from multiyear satellite OMI observations in the northern hemisphere at belt of latitudes from 45°N to 50°N and from ground observations at the ARG station in Chisinau during the period from 2004 to 2008.

The multiyear monthly means of TOC retrieved from TOMS and OMI observations (1978-2008) and ground observations ARG (2004-2008) at the Chisinau site are presented in Table 2. One can clearly see a good coincidence of monthly means retrieved from satellite (platforms N7, M3, EP, and Aura) observations with those acquired from direct observations at the ground ARG station. Some discrepancies between monthly mean statistics for satellite observations and ground measurements (see Fig. 6 and Table 2) may be attributed to low representativeness of statistical set of direct ground observations of TOC in the course of some periods, mainly in November and December of 2004, 2006, and 2007. It was due to the presence a large number of days with cloudy or overcast conditions. During these days, direct ground observations of TOC were impossible.

Table 2. Multiyear monthly means of TOC retrieved from TOMS and OMI observations (1978-2008) and direct ground ARG observations (2004-2008) at the Chisinau site.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
TOMS+OMI	341	368	378	377	363	348	330	315	303	289	293	315
ARG	342	360	379	375	360	347	331	316	305	289	300	297

To make assessments of variability of seasonal means of TOC, $\langle X \rangle_s$ respective datasets of daily means TOC were retrieved from combined multiyear series of TOMS and OMI observations during the period from 1978 to 2008 and interpolated at specific sites (Table 1) across the territory of Moldova. Interpolated daily means of TOC for selected sites at the borders were grouped into the seasons as follows: winter (December, January, February), spring (March, April, May), summer (June, July, August), and fall (September, October, November). It was found that the variability of seasonal means $\langle X \rangle_s$ in latitudinal and longitudinal directions for border's set of sites is small. The differences Δ_s between seasonally mean values of TOC retrieved from satellite observations at the respective pairs of sites at the borders in the N-S and W-E directions are presented in Table 3. Differences Δ_s(NS) and Δ_s(WE) range from -1.2 DU to 2.1 DU and from -1.0 DU to 0.5 DU, respectively. It should be noted that the change in the sign of the difference Δ_s both to the N-S and W-E directions occurs in summer.

Table 4 presents the seasonally means TOC values averaged over Moldova from TOMS and OMI observations from 1978 to 2008. A distinct seasonal dependence is observed. The fall season exhibits a low mean value of TOC $\langle X \rangle_s$~295 DU, and the spring season reveals high means of TOC $\langle X \rangle_s$ ~373 DU.

Table 3. Differences Δ_s between seasonal means of TOC $<X>_s$ retrieved from combined TOMS and OMI satellite observations at the pairs of sites at borders of Moldova in the N-S and W-E directions from 1978 to 2008.

season	Δ_s (NS), in DU	Δ_s(WE), in DU
winter	0.9	-1.0
spring	1.8	-0.7
summer	+2.1	+0.5
fall	-1.2	-0.6

Table 4. Seasonal means of TOC $<X>_s$ retrieved from combined TOMS and OMI observations over the territory of Moldova from 1978 to 2008.

season	TOC $<X>_s$, in DU
winter	340
spring	373
summer	331
fall	295

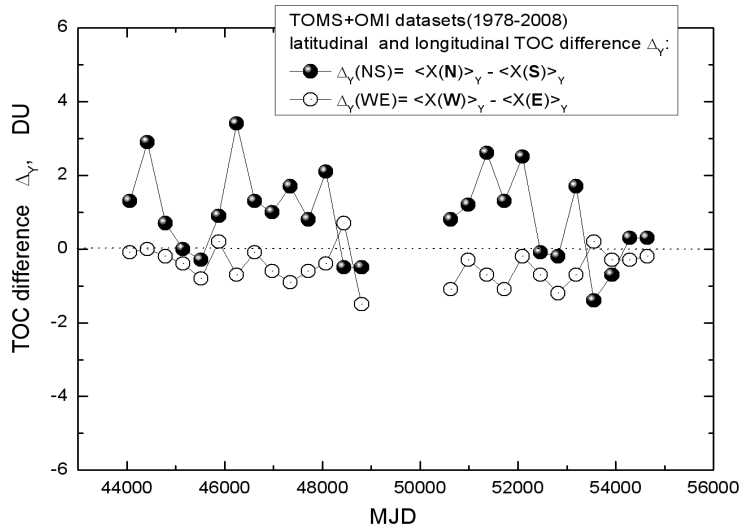


Fig. 8. Latitudinal and longitudinal variability of difference Δ_Y (in DU) between respective yearly means of TOC $<X>_Y$ derived from TOMS and OMI multiyear observations for two pairs of sites near the borders along the N-S and W-E directions. Period of observations: 1978-2008.

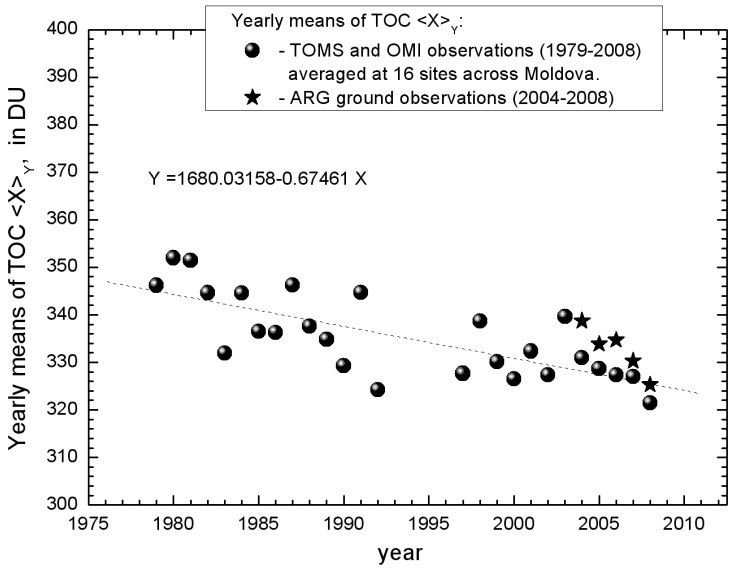


Fig. 9. Time-series of yearly means of TOC averaged at 16 sites across Moldova from TOMS and OMI multiyear observations. Period of observations: from 1979 to 2008.

Time series of latitudinal and longitudinal variability of difference Δ_Y derived between yearly means of TOC $<X>_Y$ from combined TOMS and OMI multiyear observations for two pairs of sites at borders in the N-S and W-E directions is shown in Fig. 8. The difference Δ_Y(NS) reveals a stronger irregularity with large amplitude in comparison with the Δ_Y(WE) due to the significant non-uniform distribution of TOC in the N-S direction. The total ranges of yearly differences Δ_Y(NS) and Δ_Y(WE) are ~4.8DU and ~2.2 DU, respectively. It can be clearly seen from Fig. 7 that the difference Δ_Y(WE) is negative nearly in all cases in the course of time-series of observations, thereby indicating the predominance of yearly means of TOC in the east region of Moldova. Meanwhile, yearly mean value of TOC $<X>_Y$ observed at the north region of Moldova, in most cases, prevails over the south yearly mean $<X>_Y$ data in the course of time-series. The range of difference Δ_Y(NS) is 2.2 times as large as Δ_Y(WE).

The time-series of yearly means of TOC data retrieved from TOMS and OMI combined observations over the territory of Moldova is shown in Fig. 9. Respective means of TOC were derived from averaging the observations acquired at 16 sites over the territory of Moldova. The results of five years of observations of TOC carried out at ARG ground station are also shown in Fig. 9. The trend of TOC over Moldova was derived relative to the climatic norm, and it was estimated to be -2% per decade. The linear regression equation is shown in Fig. 9. The climatic norm for yearly means of TOC $<X>_Y$ amounts to ~ 335 DU during the period of observations from 1979 to 2008.

4. Conclusions

Analysis of longitudinal and latitudinal variations of statistical means of TOC observed at the sites situated at borders across the territory of Moldova was carried out. It was based on using retrievals of daily means of TOC from TOMS and OMI observations during the period from 1978 to 2008. TOC values at specific coordinates of border sites in the N-S and W-E directions were determined by using linear interpolation from parent databases. Statistical means, such as monthly, seasonally and yearly means, were derived directly from interpolated daily means of TOC. Derived TOC values revealed both spatial and temporal variability. Longitudinal and latitudinal variability was found to depend upon the applied statistical means. Differences between statistically mean values of TOC retrieved for pairs of sites situated at the borders both in latitude (N-S) and longitude (W-E) were used for analysis. It was established that the ranges of differences between statistically derived TOC for respective pair of sites in the N-S and for pair of sites in the W-E directions amounted to ~25 DU and ~10 DU (for monthly means of TOC), to ~3.3 DU and 1.5 DU (for seasonally means of TOC), and to ~4.8 DU and ~2.2 DU (for yearly means of TOC). The range of difference Δ_Y(NS) is 2.2 times as large as for Δ_Y(WE). An analogous relation is specific to ratio of differences between the seasonally means Δ_s(NS) and Δ_s(WE).

The monthly means of total ozone content for all sites revealed a distinctive seasonal variation with maximum ~378 DU (in spring) and minimum ~ 289 DU (in fall). The seasonal means of TOC $<X>_s$ retrieved from combined TOMS and OMI observations, averaged over the territory of Moldova, ranged from ~295 DU (in fall) to ~373 DU (in spring) during the period from 1978 to 2008. The climatic norm for yearly means of TOC $<X>_Y$, derived from the averaging of respective values of TOC at 16 sites over the territory of Moldova from 1978 to 2008, amounts to ~335 DU; the trend of TOC over Moldova is -2% per decade.

References

[1] M.P. Chipperfield et al., Global ozone: past and future, Chapter 4 in Scientific Assessment of Ozone Depletion: 2002, Global Ozone Research and Monitoring Project–Report No. 47, 498, World Meteorological Organization, Geneva, Switzerland, (2003).

[2] M.P. Chipperfield et al., Global Ozone: Past and Present, Chapter 3 in Scientific Assessment of Ozone Depletion: 2006, Global Ozone Research and Monitoring Project–Report No. 50, 572, World Meteorological Organization, Geneva, Switzerland, (2007).

[3] R.D. McPeters, P.K. Bhartia, A.J. Krueger, J.R. Herman, C.G. Wellemeyer, C.J. Seftor, G. Jaross, O. Torres, L. Moy, G. Labow, W. Byerly, S.L. Taylor, T. Swissler, and R.P. Cebula, Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide, NASA Technical Publication 1998-206895, 64, National Aeronautics and Space Administration, Goddard Space Flight Center Greenbelt, Maryland 20771, (1998), ftp://toms.gsfc.nasa.gov/pub/eptoms/EARTHPROBE_USERGUIDE.PDF

[4] J.P. Veefkind, J.F. de Haan, E.J. Brinksma, M. Kroon, and P.F. Levelt, Geoscience and Remote Sensing, IEEE Transactions, 44 (5), 1239-1244, DOI 10.1109/TGRS.2006. 871204, (2006).

[5] OMI Algorithm Theoretical Basis Document (ATBD), vol. I, OMI Instrument, Level 0-1b processor, Calibration & Operations, Ed. by P.F. Levelt, ATBD-OMI-01, ver. 1.1, 50, August, 2002, http://www.knmi.nl/omi/documents/data/OMI_ATBD_Volume_1_V1d1.pdf

[6] OMI Algorithm Theoretical Basis Document (ATBD), vol. II, OMI Ozone Products, Ed. by P.K. Bhartia, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, ATBD-OMI-02, ver. 2.0, 91, August, 2002, http://www.knmi.nl/omi/documents/data/OMI_ATBD_Volume_2_V2.pdf

[7] P.K. Bhartia and C.W. Wellemeyer, TOMS-V8 Total O3 Algorithm: Algorithm Theoretical Basis Document (ATBD), http://toms.gsfc.nasa.gov/version8/v8toms_atbd.pdf

[8] M. Morys, F.M. Mims III, S. Hagerup, S.E. Anderson, A. Baker, J. Kia, and T. Wallkup, J. Geophys. Res., 106, 14573-14582, (2001).

[9] U. Kohler, Geophys. Res. Lett., 26 (10), 1385-1388, (1999).

[10] D.H. Holdren, R.O. Olsen, and F.J. Schmidlin, Geophys. Res. Lett., 28 (20), 3859-3862, (2000).

[11] A. Aculinin, A. Smirnov, V. Smicov, T. Eck, and A. Policarpov, Mold. J. Phys. Sci., 3, 2, 204-213, (2004).

[12] Total Ozone Mapping Spectrometer (TOMS), Ozone Processing Team – NASA/GSFC code 613.3, Greenbelt, Maryland, USA; ftp://toms.gsfc.nasa.gov/pub/eptoms/data/ozone/

[13] Total Ozone Mapping Spectrometer (TOMS): Ozone Monitoring Instrument (OMI) Maps and Data, Ozone Processing Team – NASA/GSFC code 613.3, Greenbelt, Maryland, USA; ftp://toms.gsfc.nasa.gov/pub/omi/data/ozone/

THE CHOICE OF METHODS FOR COMPRESSING THE INPUT SIGNALS OF AN ADAPTIVE TELEMETERING SYSTEM

M. Vladov

Institute of Electronic Engineering and Industrial Technologies, Academy of Sciences of Moldova, 3/3, Academiei str., MD-2028, Chisinau, Republic of Moldova
(Received 9 December 2009)

Abstract

Methods for compressing quasideterministic and stochastic signals are proposed. For compressing quasideterministic signals, it is proposed to transfer the extremes of signals; thus, the factor of compression for five groups of quasideterministic signals is gained more than 5. For compressing stochastic signals by irreversible methods, the irreversible aperture compression is recommended. The irreversible compression allows reducing data flow, and the aperture method of compression allows obtaining a compression up to 1.7 times.

Introduction

A gain in the amount of measuring data during development and flight tests of aircraft engines results in necessity to expand communication channels advancing bend, to increase the process time of measuring data as well as to increase the memory capacity of recording devices. In addition, sometimes a certain part of measuring data is redundant and sometimes pick-off signals change insignificantly. Therefore, the design of adaptive telemetry systems with data compression is of current concern. It offers savings in soft and hardware resources and information processing.

This work describes the methods of quasideterministic and stochastic signals from detectors of various dimensions of an aircraft telemetry system.

1. Quasideterministic signals

To compress quasideterministic signals containing a mixture of harmonic vibration and low frequency component (trend), it is proposed to use a method at which values of extremum of signals should be transferred (Fig. 1).

Different means may be used to recover such a signal.

If the oscillation amplitude is substantially lower than the measurement range of a signal, the linear interpolation is a simplest mean. In this case, an imprecision of approximation will depend on change of signal and oscillation amplitude ratio. For example, if the oscillation amplitude does not exceed 5% of the measurement range, the relative error of approximation will be 0.5%. At large amplitude values, we may either add intermediate signal values or recover the signals using polynomials of higher degree. The last method is more preferable. The accuracy in the last case will be determined by the accuracy of transferred extremum values.

On the basis of data presented in [1], we carried out a calculation of compression ratio of quasideterministic signals for constant value of uniform time sampling equal to 300 Hz (table).

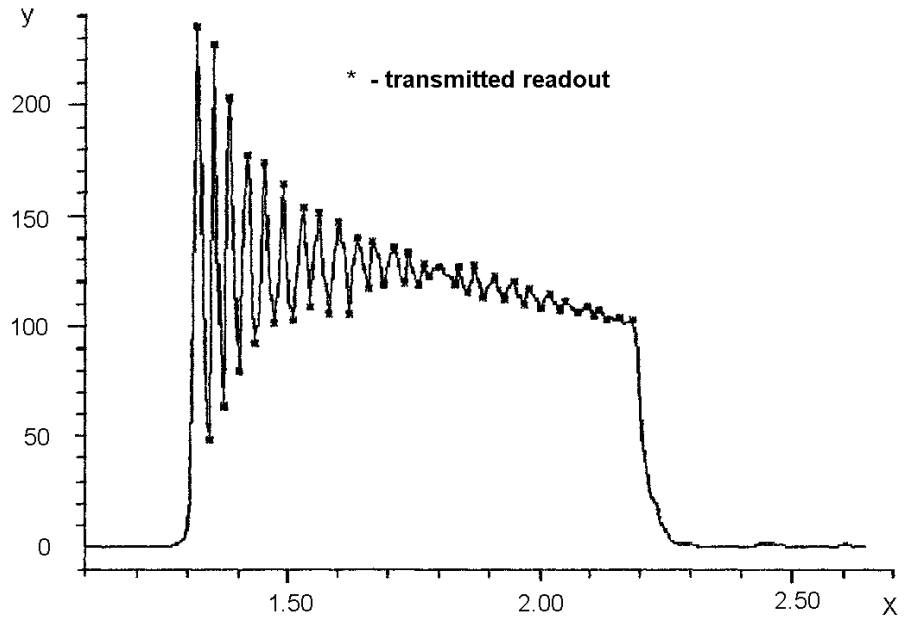


Fig. 1. Extrema of signal that are subject to transfer.

Compression ratio calculation of quasideterministic signals.

Set of signal	Central frequency of the signal sampling (Hz)	Compression ratio
1	52	5.7
1	40	5.2
2	58	5.9
3	61	4.9
3	56	5.4

In order to make calculations at preset section of a certain length **T**, the number of extremum of a signal **V** is found. In this case, the medium frequency of signal sampling is formed by the relation $\overline{f} = \mathbf{V}/\mathbf{T}$. The compression ratio calculation was made for the most active test sections (initial and final). Subject to a real nonstationarity of signals, an integral compression ratio will be considerably higher.

2. Stochastic signals

Some methods of irreversible and aperture compression are available to compress stochastic signals (Fig. 2) describing fixed stochastic processes (signals of groups 4, 5, and 6).

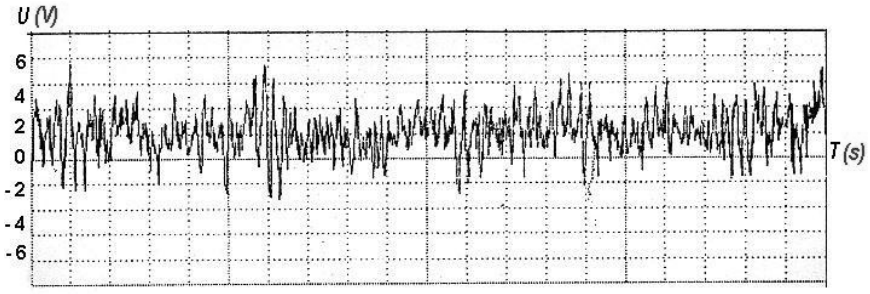


Fig. 2. Stochastic signals.

Irreversible compression is recommended in case when characteristic wavelength is of interest but not significant of measurable values. This compression technique may be used, for example, for measuring accelerations and some other quantities. Particularly in [1], it is proposed to process high-frequency component of flight data Space Shuttle by means of autoregressive (AR) algorithms and autoregressive algorithms with moving average (ARMA). These procedures enable to appreciate spectral concentrations of data set on changing of acceleration in the start time. The result of processing is rating of a current spectrum whose dynamics is lower than dynamics of the signal proper at least by an order. Signal characteristic is passed to a receiving station: 8 discharges (stages) – spectrum value and 8 discharges (stages) – frequency value. 8 spectrum discharges enable to transmit the values of spectrum components with inaccuracy not more than 0.5%. 8 frequency discharges enable to transmit the values of spectrum components for 256 frequency values. Irreversible compression enables to reduce data flow at least by an order.

Aperture compression is ordered to be used for the last signal groups on basis of interpolation polynomials of 1-2 levels. Suppose the signals are represented by a model of fixed stochastic process with a correlation function

$$\mathbf{R}_1(\tau) = \exp(-\alpha|\tau|)\cdot(1+\alpha|\tau|+\alpha^2\tau^2/3), \tag{1}$$

where α is the coefficient, has different values for various stochastic signals; $\tau = t_2 - t_1$ is the time difference, for correlation function $k(t_1, t_2)$, or

$$\mathbf{R}_2(\tau) = \exp(-\alpha^2\tau^2). \tag{2}$$

In this case, the analytical form can be obtained for average frequency of the adaptive time digitization opportune to make calculations. Signals with correlation function $\mathbf{R}_2(\tau)$ are "smooth enough", and belong to the infinitely differentiable class. It is possible to use compression algorithms of higher degrees for such signals. We shall restrict our consideration to the second-degree interpolation taking into account that, as the degree of an approximating polynomial increases, the realization complexity of compression algorithm increases sharply.

The digitization average frequency is defined by the formula according to [2] with the use of the extrapolation algorithms of compression

$$\overline{f}_{en} = \frac{2(n+1)\Gamma[(3n+4)/(2n+2)]}{(n+2)\sqrt{p}}\cdot\left[\frac{\sigma_{n+1}\sqrt{2}}{\varepsilon ad(n+1)}\right]\cdot\frac{1}{n+1}, \tag{3}$$

where Γ is the gamma function; n is the derivative number; σ_{n+1} is the mean square value $(n+1)$ derivative; εad is the admissible error of sampling.

The signal dispersion for n derivative

$$\sigma_n^2 = (-1)^n\cdot\mathbf{R}^{2n}(0)\sigma^2. \tag{4}$$

where σ^2 is the signal dispersion.

For signals with correlation function, the signal dispersion for the first, second, and third derivatives is as follows

$$\begin{aligned}\sigma_1^2 &= \alpha^2\cdot\sigma^2;\\ \sigma_2^2 &= 12\alpha^4\cdot\sigma^2;\\ \sigma_3^2 &= 120\alpha^6\cdot\sigma^2.\end{aligned} \tag{5}$$

For n = 2, (3) will be

$$\overline{f}_{en} = \frac{0.58\alpha}{\sqrt[3]{\varepsilon_0}}, \tag{6}$$

where $\varepsilon_0 = \varepsilon_{ad} / 6\sigma$ is the adjusted value of the admissible error.

We shall define the average frequency of signal digitization at its description by the second-degree interpolation polynomial. During the extrapolation restoration by the second-degree polynomial, the step of the temporal uniform sampling according to [3] is

$$\Delta T_e = \sqrt[3]{6\varepsilon_{ad} / M_3}. \tag{7}$$

where M_3 is the module maximum of the third derivative.

During the interpolation restoration by the second-degree polynomial and uniform precision factor, the step of the temporal uniform sampling is as follows

$$\Delta T_i = \sqrt[3]{15.6\varepsilon_{ad} / M_3}. \tag{8}$$

Ratio $\Delta T_i / \Delta T_e = 1.38$. The average interval increases by 1.38 as in the adaptive digitization; as a result, the average frequency of signal digitization will be

$$\overline{f}_{n2} = \frac{0.42\alpha}{\sqrt[3]{\varepsilon_0}}. \tag{9}$$

Thereupon, the discretization medium frequencies for the signals of 4-6 groups described by correlation function $\mathbf{R}_2(\tau)$ will be, as it follows from (9), with values $\alpha = 175$ and $\varepsilon_0 = 0.005$:

- for the signals of group 4, $\overline{f}_{n2} = 430$ Hz;
- for the signals of group 5, $\overline{f}_{n2} = 860$ Hz;
- for the signals of group 6, $\overline{f}_{n2} = 1280$ Hz.

Then the compression factor for the signals of groups 4-6 described by the correlation function $\mathbf{R}_2(\tau)$ will be 2.3 for frequencies with uniform time digitization of 1000, 2000, and 3000 Hz for signals of groups 4-6.

A signal with the correlation function $\mathbf{R}_1(\tau)$ is subject to double differentiation. For such signals, the approximation by a polynomial of a degree above the first slightly increases a digitization step; therefore, in this case, the linear interpolation is applicable. We have for a signal with the correlation function $\mathbf{R}_1(\tau)$

$$\begin{aligned} \sigma_1^2 &= \alpha^2 \cdot \sigma^2 / 3; \\ \sigma_2^2 &= \alpha^4 \cdot \sigma^2. \end{aligned} \tag{10}$$

Then, on the basis of expression (3), n =1, and taking into account the fact, that the digitization step at linear interpolation ΔT_i is more than twice as much as the digitization step of linear extrapolation ΔT_e, we obtain

$$\overline{f}_{n1} = 0.29\sqrt{\frac{\sigma_2}{\varepsilon_{ad}}} = \frac{0.12\alpha}{\sqrt{\varepsilon_0}}. \tag{11}$$

Then, for the signals of groups 4-6 described by the correlation function $\mathbf{R}_1(\tau)$, as it follows from (11), at values $\alpha = 340$ and $\varepsilon_0 = 0,005$ average frequencies of digitization will be equal:

- for the signals of group 4, $\overline{f}_{n1} = 577$ Hz;
- for the signals of group 5, $\overline{f}_{n1} = 1154$ Hz;
- for the signals of group 6, $\overline{f}_{n1} = 1731$ Hz.

Then the compression factor for the signals of groups 4-6 described by the correlation function $\mathbf{R}_1(\tau)$, will be 1.7 for frequencies with uniform time discretization of 1000, 2000, and 3000 Hz for the signals of groups 4-6.

Conclusions

It is determined that signal compression with the ratio exceeding 5 is possible for quasideterministic signals of groups 1-3 during transfer of the extremum values of these signals. Compression of the stochastic signal of groups 4-6 can be made with the ratio reaching a value of 1.7.

Taking into account the derived ratio compression (additional gauges of telemetering parameters can be connected), an increase in the transfer speed at least by a factor of 1.7 and a two-fold decrease in the channel bandwidth, which leads to improvement of system noise immunity, can take place.

References

[1] P.D. Spanos and L.J. Mushung, AIAA/ASME/ASCE/AHS/ASC 31st Struct., Struct. Dyn. and Mater. Conf., Long Beach, Caliph., April 2-4, 1990: Collect. Tech. Pap. Pt. 4., Washington (D.C.), 2173, (1990).

[2] S.N. Dolinov and I.A. Karabanov, Transformation of measuring signals, L.: Edited LETI, 76 p., 1986.

[3] B.I. Avdeev et al., Adaptive telemetry gage systems, Edited by A.V. Fremke, L.: Energoizdat, Leningrad branch, 248 p., 1981.

EXTERNAL Sq CURRENTS IN THE WEST AFRICAN REGION

T.N. Obiekezie[1] and F.N. Okeke[2]

[1]Department of Physics and Industrial Physics, Nnamdi Azikiwe University, Awka, P.M.B. 5025, Anambra state, Nigeria
E-mail: as27ro@yahoo.com
[2]Department of Physics and Astronomy University of Nigeria Nsukka
E-mail: franciscaokeke@yahoo.com
(Received 3 February 2009)

Abstract

The external Sq currents for quiet days in 1993 were studied for the West African region using the magnetic data obtained in the course of the French participation in the International Equatorial Electrojet Year (IEEY). Spherical Harmonic Analysis technique was used to separate the internal and external contributions of the quiet field variations. The external current range provided a simple way of viewing a full year's change in the external Sq current system. The maximum range was found in March with a value of 34.4×10^6 A and the minimum found in December with a value of 1.7×10^6 A. The seasonal variation was found highest in March (equinox) with a value of 16.6×10^6 A followed by June (summer solstice) 9.42×10^6 A; the least was found in December (Winter solstice) 7.76×10^6 A.

1. Introduction

The current sources of Sq are known to arise primarily from thermal-tidal motions in the ionosphere and they induce a secondary current that flows in the conducting earth. The depth of this induction depends upon how rapidly the source varies and the conductivity patterns within the earth. At the earth's surface, magnetometers measure the sum of the fields from the source and induced currents.

By using Gauss Spherical Harmonic Analysis (SHA) method or other integral methods one can separate this current into its component parts. The authors of [1, 2] used this SHA method to show that the daily quiet-time geomagnetic field variations came mostly from sources of current external to the Earth; this finding led to the discovery of the ionosphere. Recent investigations of the external source currents for the India-Siberia region [3-5] and Australian region [6] are the latest example of this Gauss SHA method.

The aim of this work is to analyze the quiet-time geomagnetic fields for the year 1993 in the West African region using the SHA method to separate the Sq ionospheric source currents from the induced currents within the Earth and examine how these external currents change through the months and seasons in the year.

2. Data and analysis

The geomagnetic data set used in this study consists of hourly mean values of H, D, and Z Fields obtained from the International Equatorial Electrojet Year (IEEY) record in the West African sub-sector. The data were collected during the intensive, coordinated, and pluridisci-

plinary investigation of ionospheric dynamics in the equatorial region of the African sector [7, 8]. The network consists of ten stations which operated from November 1992 to November 1994 along a 1200 km long North-South profile across the magnetic dip equator. The sites are located between Ivory Coast in the south and Mali in the North with approximately 150 km space between sites. Table 1 shows the geographic location of the stations.

Table 1. Geographic positions of the magneto telluric stations operating during the IEEY West African experiment.

	stations	symbols of stations	latitudes (°N)	Dip-latitudes (°N)	distances (km) from dip-equator	Longitudes (°W)
1	Tombouctou	TOM	16.733	5.513	611.98	3.000
2	Mopti	MOP	14.508	3.288	365.00	4.087
3	San	SAN	13.237	2.017	223.91	4.879
4	Koutiala	KOU	12.356	1.136	126.11	5.448
5	Sikass	SIK	11,344	0.124	13.75	5.706
6	Nielle	NIE	10.203	-1.017	-112.85	5.636
7	Korhogo	KOR	9.336	-1.884	-209.17	5.427
8	Katiola	KAT	8.183	-3.037	-337.1	5.044
9	Tiebissou	TIE	7.218	-4.003	-444.48	5.241
10	Lamto	LAM	6.233	-4.988	-553.61	5.01

The method of analysis for this work involves the spherical harmonic analysis (SHA) devised by [9] in solving the magnetic potential function V, in which it was shown that the potential has two parts—the external (source) and internal (induced) parts of the potential function. The magnetic potential V at geocentric distance r is expressed as

$$V = C + R\sum_{n}\sum_{m}\left\{\left[a_n^{me} + a_n^{mi}\right]Cosm\phi + \left(b_n^{me} + b_n^{mi}\right)Sinm\phi\right\}P_n^m(\theta), \tag{1}$$

where $C, \theta, R, and \phi$ denote the constant of integration, the geomagnetic colatitude, the earths radius, and the local time of the observatory. The $a_n^{me}, a_n^{mi}, b_n^{me}$ and b_n^{mi} are the Legendre polynomial coefficients, where e and i represent the external and internal values, respectively. p_n^m are the Legendre polynomials and functions of colatitude θ only. The integers n and m are called degree and order, respectively; n has a value of 1 or greater, and m is always less than or equal to n.

On the other hand, the surface magnetic variation fields, H, D, and Z are analyzed in the form

$$H = \sum_{n=m}\sum_{m=1}\left[a_n^m Cos(m\phi) + b_n^m Sin(m\phi)\right]\frac{\partial P_n^m(\theta)}{\partial\theta}, \tag{2}$$

$$D = \sum_{n=m}\sum_{m=1}\left[-b_n^m Cos(m\phi) + a_n^m Sin(m\phi)\right]m\frac{P_n^m(\theta)}{Sin\,\theta}, \tag{3}$$

$$Z = \sum_{n=m}\sum_{m=1}\left[a_n^m Cos(m\phi) + b_n^m Sin(m\phi)\right]P_n^m(\theta), \tag{4}$$

where

$$a_n^m = \left(a_n^{me} + a_n^{mi}\right) \text{ and } b_n^m = \left(b_n^{me} + b_n^{mi}\right). \tag{5}$$

The equivalent current function $J(\phi)$, in amperes, for an hour of the day is obtained from

$$J(\phi) = \sum_{m=1}^{4}\sum_{n=1}^{12}\left(U_n^m Cos(m\phi) + V_n^m Sin(m\phi)\right)P_n^m(\theta), \tag{6}$$

$$U_n^m = -\left(\frac{5R}{2\pi}\right)\left(\frac{2n+1}{n+1}\right)a_n^{me}\left(\frac{a}{R}\right)^n, \tag{7}$$

$$V_n^m = -\left(\frac{5R}{2\pi}\right)\left(\frac{2n+1}{n+1}\right)b_n^{me}\left(\frac{a}{R}\right)^n. \tag{8}$$

For the internal current representation

$$U_n^m = \left(\frac{5R}{2\pi}\right)\left(\frac{2n+1}{n}\right)a_n^{mi}\left(\frac{R}{a}\right)^{n+1}, \tag{9}$$

$$V_n^m = \left(\frac{5R}{2\pi}\right)\left(\frac{2n+1}{n}\right)b_n^{mi}\left(\frac{R}{a}\right)^{n+1}, \tag{10}$$

where a is the geocentric distance of the current sheets [10].

The analysis was started with the selection of magnetically quite days from the five internationally quite days (IQDS) in each month for the year 1993. The base line for this analysis is taken as the average for the mean of the five quite days for the month. The base line removal was followed by the Fourier analysis of the three components of the magnetic field, then came the polynomial fitting of the latitudinal variation of the station field Fourier components and the computation of the Legendre polynomial coefficients. To avoid noise that is common in high harmonics, the SHA coefficients were computed to degree 12 and order 4.

3. Results and discussion

Figure 1 illustrates the separated currents; external, internal, and total (external + internal) currents in the West African region for the month of March in 1993. The station Mopti is used to represent the northern stations while Katiola represents the southern stations. The internal currents are seen to be different from the external currents both in amplitude and in phase. These phase and amplitude differences are a function of the Earth conductivity (4).

The external current pattern is seen to be different in the northern and southern stations. The external current pattern in the northern stations is seen to be exactly opposite that in the southern stations.

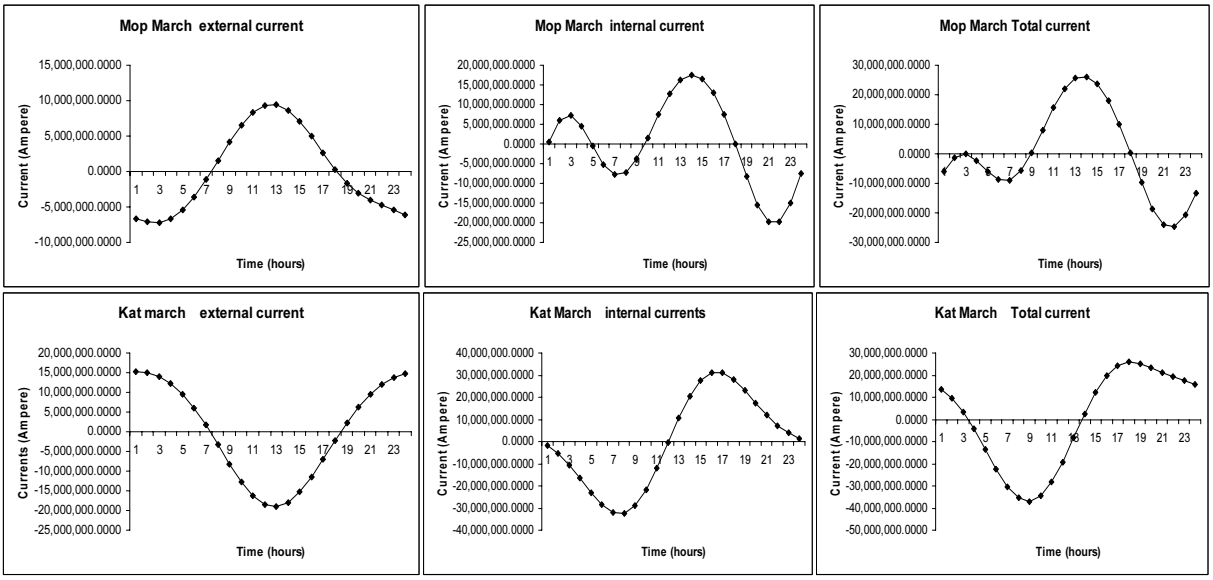


Fig 1. Separated external, internal, and total currents for Mopti (MOP) and Katiola (KAT) for the month of March.

For the northernmost stations (Tombouctou, Mopti, and San), it is seen to be negative in the early morning hours and positive in the afternoon; then turns negative again in the late noon hours with a shape of an anticline, while for the southernmost stations (Katiola, Lamto, and Tiebissou), it is seen to be positive in the early morning hours, negative in the afternoon; then it turns positive again in the late noon hours, having a shape of a syncline. The midday increase at the northernmost stations is matched by a midday decrease in the southernmost stations.

The external current range for each month and station is identified as the difference between the maximum and the minimum current levels for the individual month and station. This range is a single value for a given month and station and provides a simple way to view a full year's change in the external Sq current system. Table 2 shows the calculated monthly ranges for the different stations. The maximum range is found in March with a value of 34.3×10^6 A and the minimum range is found in December with a value of 1.7×10^6. The yearly averages of March, June, and December are used to represent the equinoxial, summer and winter solsticial variation of these external currents. We found from Table 2 the maximum value in March (equinox) with a value of 16.6 x 10^6 A followed by June (summer solstice) 9.42×10^6 A; the least is found in December (Winter solstice) 7.76×10^6 A. Our representation of the quite external currents can be compared to several regional Sq studies. The two most comprehensive ones are that of [11, 12].

The authors of [11] analyzed the international Geophysical year (IGY) 1958 records by dividing the seasons into three month groups of D, E, and J months. Their computed current vortices from the Euro-African sector (they called it zone 1) showed maximum amplitudes of 12.9×10^4, 22.5×10^4 and 26.4×10^4 for the D, E, and J months, thus having a maximum occurring at the J months. If we followed [11] three months grouping we would get 11.13×10^6, 13.27×10^6, and 8.58×10^6 A for D, E, and J months having a minimum in the J months, a maximum for the E months. This is in disagreement with [11], who reported their zone 1, northern hemisphere (Euro-African region) J month currents to be larger than their E month current. A part of the discrepancy may be due to this decision to average together recordings of two different continents. The stations in this study are located very close to the equator; in effect, we are dealing with equatorial stations, so it is not surprising that we got our maximum at the equinox, when the sun is overhead at the equator. This is understood since one of the critical features affecting the Sq field pattern is solar ionization.

Table 2. Monthly range of external current intensity in Ampere.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
		CURR	ENT		INTE	NSI	TY	X	10^6			
Tom	3.8	7.8	11.3	13.2	10.4	8.0	3.6	8.9	7.8	6.3	4.6	3.6
Mop	4.3	3.9	16.5	18.8	5.3	5.6	6.0	6.2	6.0	4.7	3.3	2.5
San	9.5	2.6	3.0	3.4	3.7	4.1	4.6	4.6	4.1	3.1	1.9	1.7
Kou	4.9	6.1	8.6	17.0	2.8	3.2	3.6	3.4	3.0	2.1	2.4	2.5
Sik	7.9	9.5	5.5	5.6	4.2	2.7	2.3	2.2	3.7	5.5	5.9	5.4
Nie	17.5	26.0	25.8	19.8	15.7	12.6	4.2	5.2	7.8	10.3	10.3	9.3
Kor	20.2	23.0	12.6	21.4	18.1	16.8	6.1	7.4	10.6	13.4	13.0	11.3
Kat	23.7	13.5	34.3	27.8	25.5	17.3	8.9	10.5	21.2	28.6	22.3	13.2
Tie	20.9	29.0	29.2	16.7	14.1	11.7	10.7	12.3	15.8	18.0	16.5	14.0
Lam	16.3	20.7	19.2	16.2	14.2	12.2	11.4	12.9	15.3	17.8	16.3	14.1
Yearly Average	12.9	14.21	16.6	16	11.4	9.42	6.14	7.36	9.53	10.98	9.65	7.76

The authors of [12] analyzed the International Quite Sun Year (IQSY) 1965 quite days and found their African external current intensities in winter, equinox, and summer to be 3.0×10^4, 10.2×10^4, (8.6×10^4 for spring and 11.8×10^4 for fall equinox), and 13.7×10^4, thus having a maximum value in summer and a minimum in winter. This is also in disagreement with our result of a maximum in equinox. The discrepancy here could be due to the fact that their stations were located in the southern hemisphere, so that at their summer time when the sun is over head they had their maximum.

Our result may be compared with the one obtained in [13], where it is found that the temperature in the upper atmosphere estimated from satellite drag on geomagnetically quite days is higher during equinox than during summer, although this claim needs to be re-verified before the relationship between the two effects can be definitely determined.

4. Conclusions

The External Sq current for the West African region for 1993 has been obtained from magnetometer data obtained during the IEEY experiment in Africa by the application of the method of Spherical Harmonic Analysis. From the results obtained the following deductions can be made.

(1) The stations located on the Northern side of the magnetic dip equator have the external current pattern exactly opposite than in the southernmost stations.

(2) The external current intensity is maximum during the March Equinox and minimum during the June solstice. This is in contrast with other regional results but could be ascribed to our station being an equatorial station and the sun being over head at the equator during equinox.

(3) The external currents varied both in amplitude and in phase with the internal currents; the variation is seen to be a function of the induction medium (the Earth).

We therefore suggest that more work should be done on this variation of external and internal currents to better elucidate the relationship between the source current and the induction medium (the Earth).

Acknowledgments

The IEEY experiment carried out in the African sector was possible because of the funds provided by: Ministère de la Coopération, Département de la Recherche et des Formations, ORSTOM, Département TOA (Terre Océan Atmosphère), CNET Centre Lannion, Ministère de la Recherche et de la Technologie, Centre National de la Recherche Scientifique, Département SDU (Sciencesde l'univers), CEA, Commissariat à 1Énergie Atomique the Université Paris-Sud; Abidjan University, Ivory Coast; Dakar University, Senegal. The efforts of the different individuals and groups who participated in the IEEY studies are greatly acknowledged.

References

[1] A. Shuster, Phil. Trans. R. Soc. Lond. A, 180, 467, (1889).
[2] A. Schutser, Phil. Trans. R. Soc. Lond. A, 280, 463, (1908).
[3] W.H. Campbell, E.R. Schiffmacher, and B.R. Arora, J. Geomg. Geoelectr., 44, 459, (1992).
[4] W.H. Campbell, B.R. Arora, and E.R. Schiffmacher, J. Geophys. Res., 98, 3741, (1993).
[5] B.R. Arora, W.H. Campbell, and E.R. Schiffmacher, Geophys. J. R. Ast. Soc., 17, 457, (1995).
[6] W.H. Campbell, C.E. Barton, and W. Welsh, Earth Planets Space, 50, 347, (1998).

[7] J. Vassal, M. Menvielle, Y. Cogen, M. Dukhan, V. Doumaya, K. Boka and O. Fambitakoye, Ann. Geophysicae, 16, 677, (1998).
[8] Y. Cohen, Ann. Geophysicae, 16, 637, (1998).
[9] C.F. Gauss, Allgemeine Theories des Erdmagnetismus, Resultate aus den Beobachtunyen des magnetischen Vereins im Jahre 1838, edited by Carl Friedrich Gauss and Wilhelm Weber, Leipzig, 1-57, 1839.
[10] W.H. Campbell, Introduction to geomagnetic fields, Cambridge University Press, New York, 304 p., 1997.
[11] S. Matsushita, and H. Maeda, J. Geophys. Res., 70, 2535, (1965).
[12] W.H. Campbell and E.R. Schiffmacher, J. Geomag Geoelectr., 40, 1387, (1988).
[13] L.G. Jacchia, Rev. Mod. Phys., 35, 973, (1963).

SIMULATION OF ROCKET SEEDING OF CONVECTIVE CLOUDS WITH COARSE HYGROSCOPIC AEROSOL
2. CONDENSATION AND COAGULATION PROCESSES IN CLOUD ZONES SEEDED WITH HYGROSCOPIC PARTICLES

M.T. Abshaev, A.M. Abshaev, M.K. Zhekamukhov, E.I. Potapov*, I.A. Garaba*, and E.A. Zasavitsky*

High Mountain Geophysical Institute of ROSHYDROMET, 2, Lenin blvd., 360020, Nalchik, Kabardino-Balkaria, Russian Federation
E-mail: abshaev@mail.ru
**Institute of Electronic Engineering and Industrial Technologies, Academy of Sciences of Moldova, 3/3, Academiei str., MD-2028, Chisinau, Republic of Moldova*
E-mail: efim@lises.asm.md
(Received 18 January 2010)

The closed system of equations is derived to describe the variation in dispersion of cloud environment due to the processes of condensation and gravity coagulation in the zone of salt particle propagation at hygroscopic seeding of convective clouds with antihail rockets. The competition for moisture between natural cloud droplets entrained into this zone from ambient cloud air due to turbulent diffusion and by artificial salt *NaCl* particles is considered. The problems of optimal methods for hygroscopic seeding of convective clouds with a view to enhance forced precipitation are discussed.

In the course of hygroscopic seeding of convective clouds by means of antihail rockets, there is such a stage of expansion of a cloud of reactive gases at which the salt crystals are completely dissolved in cloud droplets formed on salt particles, and further growth of these droplets occurs in conditions of the competition for moisture between "salt" and natural cloud droplets. In the given work, the condensation and coagulation growth of cloud droplets is considered taking into account this competition.

In the nature of things, artificial "salt" droplets will grow more intensively than natural cloud ones. The "salt" droplets will grow into large cloud droplets, which will grow due to coagulation with smaller cloud droplets. The problem reduces to the elucidation of the role of salt crystals in the enlargement of droplets and the intensification of the precipitation formation process.

Below, we present a derivation of the system of equations describing the condensation and coagulation processes in the zone of propagation of salt particles brought into a cloud by the rocket method.

The total liquid-water content in unit volume of a cloud in the salt-particle propagation zone is $\rho_{v\infty}+\omega^*+\omega$, and in the cylinder radius R – $\pi R^2(\rho_{v\infty}+\omega^*+\omega)$, where $\rho_{v\infty}$ is the density of the water vapor in the cloud air; ω^* is the liquid-water content of "salt" cloud water, i.e., the total weight of the droplets formed on the salt crystals in unit volume; ω is the total weight of the droplets formed on the nonhydroscopic nuclei of crystallization and of the droplets entrained from the cloud environment.

A change in this value occurs due to turbulent entrainment of cloud environment, where the liquid-droplet water content of a cloud is ω_e, and the water vapor content in unit volume is ρ_{ve}. Thus, we can write

$$\frac{d}{dt}\left[\pi R^2\left(\rho_{v\infty}+\omega^*+\omega\right)\right]=\frac{d}{dt}\left[\pi R^2\right]\times\left(\omega_e+\rho_{ve}\right).$$

Let ω_e be constant and time-invariant. Then, we obtain

$$\frac{d\rho_{v\infty}}{dt}+\frac{d\omega^*}{dt}+\frac{d\omega}{dt}=\frac{2}{R}\frac{dR}{dt}\left(\rho_{ve}+\omega_e-\rho_{v\infty}-\omega^*-\omega\right). \quad (1)$$

Here the multiplier $\frac{2}{R}\frac{dR}{dt}$ is determined by the law of expansion of the zone of propagation of the salt brought into the cloud. Hereinafter, we will assume that salt particles brought into the cloud by the rocket method propagate similar to the propagation of the admixture of an instantaneous linear source and that the radius of the crystal propagation zone varies by the law [1]

$$R=\alpha\, t^{3/2},$$

where $\alpha=0.58\sqrt{\varepsilon}$; ε is the rate of dissipation of the turbulent kinetic energy.

We consider the stage of expansion of a jet of reactive gases from the moment when the temperature of the gases decreases to the temperature value in the cloud environment; that is, when the jet radius achieves the value R_k, which is determined by the formula $R_k=\alpha t_k^{3/2}$. Then, $R=\alpha\left(t_k+t\right)^{3/2}$, where t_k is the time in the course of which the cloud expands to the value R_k. Herein, we obtain

$$\frac{dR}{dt}=\frac{3}{2}\alpha\left(t_k+t\right)^{1/2},\quad \frac{2}{R}\frac{dR}{dt}=\frac{2}{\alpha\left(t_k+t\right)^{3/2}}\cdot\frac{3}{2}\alpha\left(t_k+t\right)^{1/2}=\frac{3}{t_k+t},$$

and equation (1) takes the form

$$\frac{d\rho_{v\infty}}{dt}+\frac{d\omega^*}{dt}+\frac{d\omega}{dt}=\frac{3}{t_k+t}\left(\rho_{ve}+\omega_e-\rho_{v\infty}-\omega^*-\omega\right). \quad (2)$$

Let us divide the spectrum of cloud droplets into gradations with respect to their radii. Thereupon, the liquid-water contents of both types of droplets can be written in the form

$$\omega^*=\sum_{i=1}^{i_{\max}}\frac{4}{3}\pi\rho_l r_i^{*3}n_i^*,\ i=1,2,...,i_{\max} \quad (3)$$

$$\omega=\sum_{j=1}^{j_{\max}}\frac{4}{3}\pi\rho_l r_j^3 n_j,\ j=1,2,...,j_{\max} \quad (4)$$

where $i_{\max}$ and $j_{\max}$ are the number of gradations of "salt" and natural cloud droplets; r_i^* and r_j, n_i^* and n_j are the radii and concentrations of the droplets that fall into the i-th and j-th gradations.

Hence, if we differentiate, we obtain

$$\frac{d\omega^*}{dt}=\frac{4}{3}\pi\rho_l\sum_{i=1}^{i_{\max}}\left[r_i^{*2}\left(3n_i^*\frac{dr_i^*}{dt}+r_i^*\frac{dn_i^*}{dt}\right)\right]. \quad (5)$$

$$\frac{d\omega}{dt}=\frac{4}{3}\pi\rho_l\sum_{j=1}^{j_{\max}}\left[r_j^2\left(3n_j\frac{dr_j}{dt}+r_j\frac{dn_j}{dt}\right)\right]. \quad (6)$$

The equation of the condensation growth of a "salt" droplet falling into the i-th gradation with regard for a wind factor can be written as follows

$$\frac{dr_i^*}{dt}=\frac{D}{\rho_l r_i^*}\left(1+\beta R_{ei}^{1/2}\right)\frac{\rho_{v\infty}-\rho_{vs}(1-c)\left(1+\frac{2\sigma}{\rho_l R_n T r_i^*}\right)}{1+\frac{\rho_{vs}DL^2}{\lambda R_n T^2}\left(1+\frac{2\sigma}{\rho_l R_n T r_i^*}\right)(1-c)}, \tag{7}$$

where $R_e=\frac{2r\,v_r(r)}{\nu}$ is the Reynolds number, $\beta=0.21$, $c=c_s\left(\frac{r_{i0}}{r_i^*}\right)^3$.

In a similar fashion, we can write the law of the condensation growth of cloud droplets, assuming $c = 0$ in equation (7)

$$\frac{dr_j}{dt}=\frac{D}{\rho_l r_j}\left(1+\beta R_{ej}^{1/2}\right)\frac{\rho_{v\infty}-\rho_{vs}\left(1+\frac{2\sigma}{\rho_l R_n T r_j}\right)}{1+\frac{\rho_{vs}DL^2}{\lambda R_n T^2}\left(1+\frac{2\sigma}{\rho_l R_n T r_j}\right)}. \tag{8}$$

In order to determine $\frac{dn_i}{dt}$, let us compose the balance equation for the number of droplets of the *j*-gradation, taking into account the droplets entrained from the cloud environment $\frac{d}{dt}\left(\pi R^2 n_j\right)=\frac{d}{dt}\left(\pi R^2 n_{ej}\right)$, or

$$\frac{d}{dt}\left(R^2 n_j\right)=\frac{d}{dt}\left(R^2 n_{ej}\right).$$

Hence, on the hypothesis that n_e does not vary in time, we obtain

$$\frac{dn_j}{dt}=\frac{2}{R}\frac{dR}{dt}\left(n_e-n_j\right).$$

Substituting the values *R* and $\frac{dR}{dt}$ here, we have

$$\frac{dn_j}{dt}=\frac{3\left(n_{ej}-n_j\right)}{t_k+t}. \tag{9}$$

Equation (9) describes the decrease in the concentration of natural cloud droplets due to the expansion of the reactive-gas propagation zone.

In order to determine $\frac{dn_i^*}{dt}$, let us compose the balance equation for the number of "salt" droplets falling into the *i*-gradation. Taking into account that the total number of droplets grown on the salt crystals does not vary, we can write

$$\frac{d}{dt}\left(\pi R^2 n_i^*\right)=0 \text{ or } \frac{d}{dt}\left(R^2 n_i^*\right)=0.$$

Hence, we obtain $R^2\frac{dn_i^*}{dt}+2Rn_i^*\frac{dR}{dt}=0$. Substituting the above given values *R* and $\frac{dR}{dt}$ here, we have

$$\frac{dn_i^*}{dt}=-\frac{3n_i^*}{t_k+t}. \tag{10}$$

Equation (10) describes the decrease in the concentration of "salt" droplets due to the expansion of the reactive-gas propagation zone.

Now, let us consider equations that describe the coagulation growth of cloud droplets in an expanding cloud of reactive gases.

The gravity coagulation of "salt" droplets between each other and with natural cloud droplets is also one of the possible mechanisms of variation in their concentration.

Let us denote by $E_{ij} = E(r_i, r_j)$ the coefficient of coagulation between particles with radii r_i and r_j ($r_i > r_j$). Herein, we use the Langmuir interpolation formula for E_{ij} that has the form

$$E_{ij} = \frac{E_{vis} + E_{pot}\dfrac{\mathrm{Re}_{ij}}{60}}{1 + \dfrac{\mathrm{Re}_{ij}}{60}}, \tag{11}$$

where $\mathrm{Re}_{ij} = \dfrac{r_j\left(V_i - V_j\right)}{\nu}$ are the Reynolds numbers for particles with the radius r_j under air flow around a particle with the radius r_i; ν is the kinematic air viscosity; V_i and V_j are the velocities of fall of particles with the radius r_i and r_j with respect to the updraft; E_{pot} is the coagulation coefficient under potential air flow around droplets

$$\begin{cases} E_{pot} = \dfrac{K_{ij}^2}{\left(K_{ij} + 0{,}5\right)^2}, \; at \; K_{ij} > \dfrac{1}{12}, \\ E_{pot} = 0, \; at \; K_{ij} < \dfrac{1}{12}, \end{cases} \tag{12}$$

where $K_{ij} = \dfrac{4V_{ij}r_j^2\rho_k}{18\eta r_i} = \dfrac{2}{9}\dfrac{\rho_k}{\eta}r_j^2\dfrac{V_i - V_j}{r_i}$ is the Stokes number; V_{ij} is the velocity of mutual approach of the particles; E_{vis} is the coagulation coefficient at the viscous regime of flow around droplets

$$\begin{cases} E_{vis} = \left(1 + 0{,}75\dfrac{\ln 2 \cdot K_{ij}}{K_{ij} - 1{,}214}\right)^{-2}, \text{ at } K_{ij} > 1{,}214, \\ E_{vis} = 0, \text{ at } K_{ij} \le 1{,}214. \end{cases} \tag{13}$$

For the steady velocity of particle fall v_r we apply the Schmidt formula

$$v_r = \frac{ar^2}{1 + br^{3/2}}, \tag{14}$$

where, according to the data of V.I. Smirnov, $a = 1.43 \times 10^6$ (cm × s)$^{-1}$, $b = 8.5 \times 10^2$ cm$^{-3/2}$; taking into account the variation in the air density with altitude, the right member of formula (14) must be multiplied by $\sqrt{\rho_{e0}/\rho_e}$, where ρ_{e0} and ρ_e is the air density on the altitude $z = z_0$ and z, respectively, where z is the altitude of droplet location.

We will approximate the air density variation ρ with altitude by the following formulae

$$\rho_e = \rho_{e0}(1 - 0{,}08z), \text{ at } z \le 8 \text{ km},$$
$$\rho_e = 0{,}52\, e^{-0{,}11(z - 8)}, \text{ at } z > 8 \text{ km},$$

where $\rho_{e0} = 1.27$ g/cm^3; z is measured in kilometers above mean sea level.

We will assume that in the reactive-gas propagation zone, "salt" and natural cloud droplets will coagulate with each other due to the difference in gravitational velocities. Herein, owning to low concentration of "salt" droplets, we can ignore the coagulation between them.

The equation that describes the variation in the weight of a "salt" droplet with the radius r_i^* due to the coagulation with smaller natural droplets with the radius r_j can be written as follows

$$\frac{dm_i^*}{dt} = \sum_{j=1}^{j_{\max}} \pi\left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| \frac{4}{3}\pi\rho_l r_j^3 n_j \,.$$

Hence, we obtain

$$\left(\frac{dr_i^*}{dt}\right)_{coag} = \frac{\pi}{3r_i^{*2}} \sum_{j=1}^{j_{\max}} \left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| r_j^3 n_j \,. \tag{15}$$

If we unite formulae (7) and (15), we obtain the formula for individual growth of "salt" cloud droplets

$$\frac{dr_i^*}{dt} = \frac{D}{\rho_l r_i^*}\left(1+\beta R_{ei}^{1/2}\right)\frac{S-\rho(1-c)\left(1+\dfrac{2\delta}{\rho_l R_n T r_i^*}\right)}{1+\dfrac{\rho D L^2}{\lambda R_n T^2}\left(1+\dfrac{2\delta}{\rho_l R_n T r_i^*}\right)(1-c)} + \frac{\pi}{3r_i^{*2}}\sum_{j=1}^{j_{\max}}\left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| r_j^3 n_j \,. \tag{16}$$

The concentration of natural cloud droplets will decrease as a result of their capture by larger "salt" droplets. Taking into account equation (16), equation (9) will be written in the form

$$\frac{dn_j}{dt} = \frac{3\left(n_{ej} - n_j\right)}{t_k + t} - \pi n_j \sum_{i=1}^{i_{\max}} \left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| n_i^* \,. \tag{17}$$

Thus, assembling equations (2), (5), (6) (8), (10), (16), and (17), we arrive at the following system of equations that describe the condensation and coagulation growth of cloud particles in the zone of hygroscopic-particle propagation in a convective cloud

$$\frac{d\rho_{v\infty}}{dt} = -\frac{d\omega^*}{dt} - \frac{d\omega}{dt} + \frac{3}{t_k + t}\left(\rho_{ve} - \rho_{v\infty} + \omega_e - \omega^* - \omega\right),$$

$$\frac{d\omega^*}{dt} = \frac{4}{3}\pi\rho_l \sum_{i=1}^{i_{\max}}\left[r_i^{*2}\left(3n_i^*\frac{dr_i^*}{dt} + r_i^*\frac{dn_i^*}{dt}\right)\right],$$

$$\frac{d\omega}{dt} = \frac{4}{3}\pi\rho_l \sum_{j=1}^{j_{\max}}\left[r_j^{2}\left(3n_j\frac{dr_j}{dt} + r_j\frac{dn_j}{dt}\right)\right],$$

$$\frac{dr_i^*}{dt} = \frac{D}{\rho_l r_i^*}\left(1+\beta R_{ei}^{1/2}\right)\frac{\rho_{v\infty} - \rho_{vs}(1-c)\left(1+\dfrac{2\sigma}{\rho_l R_n T r_i^*}\right)}{1+\dfrac{\rho_{vs} D L^2}{\lambda R_n T^2}\left(1+\dfrac{2\sigma}{\rho_l R_n T r_i^*}\right)(1-c)} + \frac{\pi}{3r_i^{*2}}\sum_{j=1}^{j_{\max}}\left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| r_j^3 n_j \,,$$

$$\frac{dr_j}{dt} = \frac{D}{\rho_l r_j}\left(1+\beta R_{ej}^{1/2}\right)\frac{\rho_{v\infty} - \rho_{vs}\left(1+\dfrac{2\sigma}{\rho_l R_n T r_j}\right)}{1+\dfrac{\rho_{vs} D L^2}{\lambda R_n T^2}\left(1+\dfrac{2\sigma}{\rho_l R_n T r_j}\right)},$$

$$\frac{dn_i^*}{dt} = -\frac{3n_i^*}{t_k + t},$$

$$\frac{dn_j}{dt} = \frac{3\left(n_{ej} - n_j\right)}{t_k + t} - \pi n_j \sum_{i=1}^{i_{\max}} \left(r_i^* + r_j\right)^2 E_{ij}\left|V_i^* - V_j\right| n_i^* \,.$$

We should add the following relations to the system of equations obtained

$$\rho_{vs}(T) = \rho_{vs0}\frac{T_0}{T}\exp\left[-\frac{L}{R_n}\left(\frac{1}{T}-\frac{1}{T_0}\right)\right],\ T = T_0 - \gamma_w W t\ ,\ t_k = \left(\frac{R_k}{\alpha}\right)^{2/3}.$$

The system of equations obtained must be solved under the following initial conditions: at $t = 0$,

$$\omega^*(0) = \sum_{i=1}^{i_{\max}}\frac{4}{3}\pi\rho_l r_i^3 n_i = \omega_0^*\ ;\ \omega(0) = \sum_{j=1}^{j_{\max}}\pi\rho_l r_j^3 n_j = \omega_{0e}\ ;\ r_{i0}^* = 1.92 r_{ki};\ c_s = 0.357\ \text{g/cm}^3;\ \sigma = 72\ \text{erg/cm}^2.$$

Numerical calculations

Let us study the seeding of a deep cumulus cloud of the *Cu Cong* type with the *NaCl* aerosol brought into the cloud by an "As" new-generation antihail rocket [2] which is capable of seeding of 1 kg of reagent on a way of 12 km.

Let us carry out the simulated seeding at a height of 400-500 m above the *Cu Cong* cloud base at a temperature on the order of 10°C.

The distributions of hygroscopic aerosol and cloud droplets over size were set in the form of the Hrgian-Mazin distribution $n(r_i) = 0.98\frac{N}{2}\frac{r_i^2}{r_m^3}\exp\left(-\frac{3.92 r_i}{r_m}\right)$, where $N = \sum_{i=1}^{i\max} n$, here r_m is the modal particle radius.

For cloud droplets, the values r_m = 8 μm and N_j = 1000 cm^{-3} were taken; at these values the liquid-droplet water content is on the order of 2 g/m^3, which is in agreement with the tabular data for deep cumulus clouds.

It is assumed that the hygroscopic aerosol is distributed uniformly in a jet of reactive gases, its initial radius for an "As" antihail rocket is approximately 3 m. Herein, the initial density of the aerosol distribution in a cloud is on the order of $3\cdot10^{-3}$ g/m^3.

For the hygroscopic particles, we will set in sequence r_m = 1.0, 2.5, 5.0, 7.5, and 10 μm. Since the *NaCl* density is 2.17 g/cm^3, we obtain that N_i = 45, 105, 350, 2800, and 42000 in the reverse sequence. Larger particles lead to a sharp decrease in the hygroscopic particle concentration formed by an "As" rocket; therefore, we do not discuss them within the bounds of this work.

The initial problem is considered in the region $D = (0, R] \times [0, T]$, where R is the upper boundary of the particle radii; T is the time range set. Let us divide the range of the particle radii $(0, r]$ in increments of $h_r = R/N_R$; the time range $[0, T]$, in increments of $\tau = T/N_T$. Then, the region D of solutions of the initial problem will be approximated by the region, where

$$(0, R]^{h_r} = \{r_e = l h_r,\ l = \overline{1, N_R}\},\ [0, T]^{\tau} = \{t_m = m\tau,\ m = \overline{0, N_T}\}.$$

The results of the model calculations are given in Figs. 1-3.

One can see from Figs. 1 and 2 that the spectrum of "salt" cloud droplets is significantly transformed in the course of all 20 min of the calculation. The spectrum shifts to the region of larger particles. During 1-3 min after the seeding, the condensation growth of "salt" droplets takes place due to the difference in the saturated vapor pressure above the surface of "salt" and natural cloud droplets. When a radius of 50-100 μm is achieved, the droplet growth occurs mainly due to the gravity coagulation with smaller cloud droplets.

At a modal radius of salt particles of 2.5 μm (Fig. 1) in 10 min drops with a radius of 50-100 μm are formed; in 14 min, 100-250 μm; in 16 min, 200-450 μm; in 18-20 min, 400-800 μm with a concentration of 0.01-1 m^{-3}.

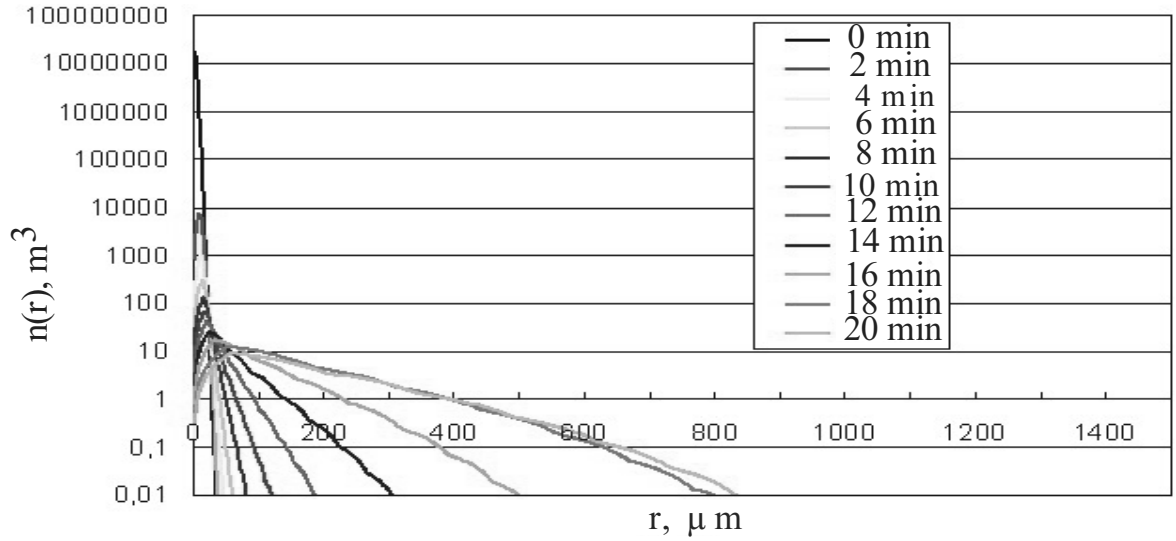


Fig. 1. Concentration of "salt" droplets in 1 m^3 in the logarithmic scale at a modal radius of hygroscopic aerosol of 2.5 μm.

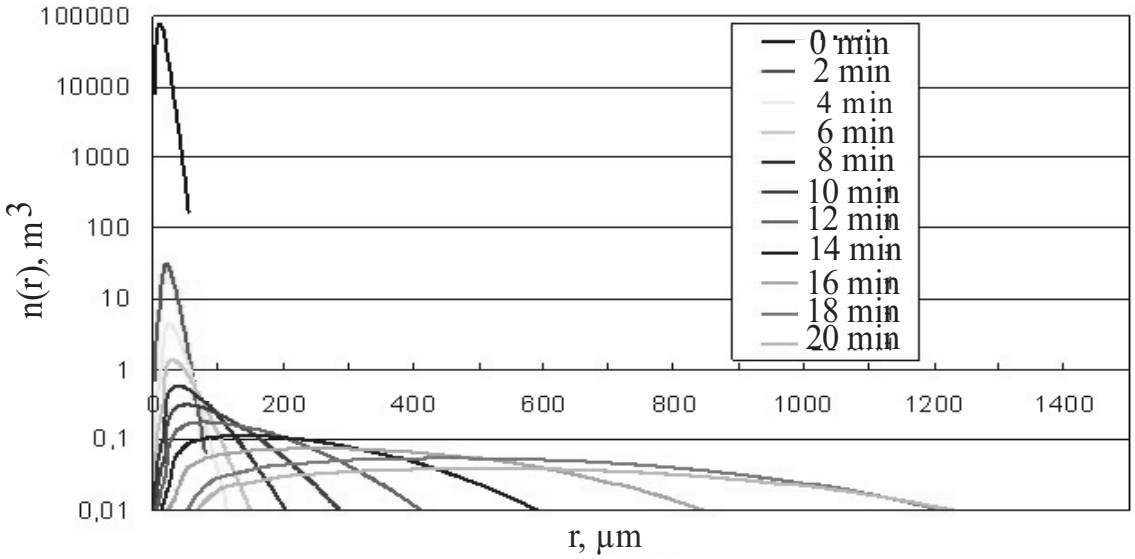


Fig. 2. Concentration of "salt" droplets in 1 m^3 in the logarithmic scale at a modal radius of hygroscopic aerosol of 10.0 μm.

At a modal radius of salt particles of 10.0 μm (Fig. 2), as early as in 1-3 min after the seeding drops with a radius of 50-100 μm are formed; in 8-10 min they grow to a radius of 150-250 μm; in 12-14 min, to 200-600 μm; in 18-20 min, to 500-1200 μm with a concentration of 0.01-0.08 m^{-3}.

It follows from the series of calculations carried out with different values of the modal radius of salt particles that as the latter increases there occurs an appreciable increase in the sizes of precipitation particles formed in the cloud with simultaneous decrease in their concentration. It is worth noting that in this case the continuous transition of small "salt" cloud droplets into large precipitation particles is achieved.

Figure 3 depicts the time evolution of the total liquid-water content Q of "salt" droplets in the seeding volume for spectra of different dispersion. One can see from the figure that in

the course of first 5 min after the seeding, the value of Q does not exceed 1-5 tons of water; by the end of the 10th-11th min it does not exceed 10 tons. Later, a sharp increase of Q for particles of any type takes place. The maximum increase is found for coarse particles. In 14 min the Q value amounts to 20-30 tons of water for the salt particles with an initial modal radius of 1 μm and 50 tons for the particles with a radius of 10 μm. If before 11-13 min the Q value of fine particles exceeded the Q value of coarse particles due to higher concentration, then subsequently the situation is quite the reverse. In 16 min the water weight for the salt particles with a radius of 1 μm is 60-80 tons; for 10 μm, about 200 tons. Towards 18-20 min, Q is 200-300 tons for 1-μm particles, 500-600 tons for 2.5 μm, and 900-1000 tons of water for the salt particles with a radius of 10 μm.

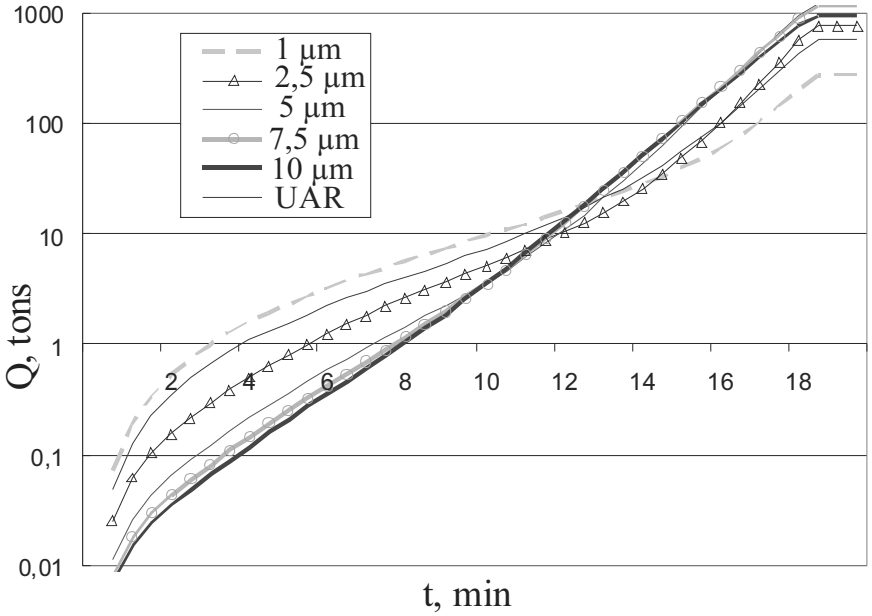


Fig. 3. Total water weight Q in tons accumulated on "salt" droplets in the volume of propagation of the hygroscopic reagent for different spectra of hygroscopic particles.

In the course of analysis of the dependence of droplet water content in the seeding volume on dispersion of hygroscopic particles, it was found that cloud droplets hardly respond to the seeding at all; they do not depend on size of the aerosol particles, which is explained by too low initial concentration of hydroscopic particles formed by an "As" rocket in the seeding volume. For an appreciable transformation of cloud droplets, it is necessary to increase the amount of introduced reagent by a factor of 10-100, which can be achieved only by seeding by equipped airplanes or by a great number of rockets.

Conclusions

1. When a hydroscopic reagent of the *NaCl* type with a particle size of 1-10 μm is introduced into cloud environment, "salt" droplets are formed, which grow in the course of 1-5 min due to the water vapor condensation and subsequently due to gravity coagulation with smaller cloud droplets.
2. At the liquid-water content on the order of 2 g/m^3, the time of the "salt"-particle growth to the size of raindrops (0.5 mm in diameter) is 17-20 min for the salt particles with a radius

of 1 μm; 12-14 min, for a radius of 5 μm; 10-12 min, for a radius of 10 μm. Higher liquid-water content of a cloud leads to more intense growth of "salt" droplets.

3. To have an effect on cumulonimbus clouds with a view to enhance precipitation, one can effectively use new-generation antihail rockets of the types "As", "Alan-3", and "Darg" filled with hygroscopic reagents. The use of particles with a radius of 7.5-10 μm is most advantageous. Thus, by the launching of one piece of "As" it is possible to obtain the fall-out of forced precipitation with a total weight of 500-1000 tons; taking into account multiplication of droplets in the course of splashing of large raindrops, amount of forced precipitation can be enhanced by several times more. Herein, the ratio of the forced-precipitation weight to the introduced-reagent weight is 5-6 orders of magnitude.

References

[1] M.T. Abshaev, A.M. Abshaev, M.K. Zhekamukhov, E.I. Potapov, I.A. Garaba, and E.A. Zasavitsky, Mold. J. Phys. Sci., 7, 2, 238, (2008).

[2] M.T. Abshaev, N.I. Mikheev et al., Tezisy Vseros. konf. po fizike oblakov i aktivnym vozdeistviyam na gidrometeorologicheskie processy (All-Russian Conf. on Physics of Clouds and Active Influence on Hydrometeorological Processes), Nalchik, 28-30, (2005).