Do the fundamental physical constants have the same values in different regions of space–time?

D. A. Varshalovich, A. V. Ivanchik, and A. Yu. Potekhin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 1–5 (September 1999)

An analysis of quasar spectra yields highly reliable constraints on the possible variation of the fine-structure constant α and the proton-to-electron mass ratio μ during cosmological evolution from the epoch corresponding to a cosmological red shift $z \approx 2.8$ (i.e., $\sim 10^{10}$ years ago) to the current epoch ($z=0$): $|\dot{\alpha}/\alpha| < 2\times 10^{-14}$ yr^{-1} and $|\dot{\mu}/\mu| < 2\times 10^{-14}$ yr^{-1}.
 [S1063-7842(99)00109-9]

1. INTRODUCTION

Constants characterizing various interactions are known to be related to one another and to depend on interaction energy.[1] However, this dependence is manifested only at very high energies (hundreds of MeV or more). Although the tabulated values of the constants pertain only to comparatively low energies, they cover the overwhelming majority of natural phenomena and experiments. It is these low-energy limits that will be examined in this report.

We note that it is actually only dimensionless combinations of physical constants, which do not depend on the choice of units of measure, that have fundamental significance. We shall consider two such combinations: the fine-structure constant $\alpha = e^2/\hbar c = 1/137.035\,999\,93\,(52)$ (the numerical value is after Ref. 2; the error in the last significant digits is indicated in parentheses) and the proton-to-electron mass ratio $\mu = m_p/m_e = 1836.152\,701\,(37)$ (i.e., the proton mass in the Hartree system of units; the value from Ref. 3 is given). Here $c = 299\,792\,458\,\mathrm{m\cdot s^{-1}}$ is the speed of light in free space (which is presently taken as a conversion factor for going from units of time to units of length[3]), $e = 4.803\,204\,251(10)\times 10^{-10}$ esu is the charge of an electron, $\hbar = 1.05457162(8)\times 10^{-27}$ erg·s is Planck's constant divided by 2π, $m_p = 1.672\,621\,62(15)\times 10^{-24}$ g is the proton mass, and $m_e = 9.109\,382\,1(8)\times 10^{-28}$ g is the electron mass (all the values were taken from Ref. 4).

The constant α is a basic parameter of quantum electrodynamics, while μ is related to the strong interaction coupling constant. The parameters α and μ are decisive for the spectra of atoms and molecules, respectively.

It can be seen from the numerical values given that most fundamental constants are measured within a relative error of $\sim 10^{-8}$. The reproducibility of these measurements includes the appreciable variation of the parameters during a short time interval, but does not include their variation during the existence of the Universe (~15 billion years). Moreover, the values of the constants might be different in spatially distant regions of the Universe.

The problem of the possible variation of the fundamental physical constants was first discussed by Milne[5] and Dirac,[6] who advanced their famous large numbers hypothesis. This problem was subsequently examined by many investigators (see, for example, Ref. 7, as well as the references cited in Ref. 8), but it has become especially acute in recent years in connection with the precipitous development of the grand unified theories (GUTs) for the strong and electroweak interactions, as well as some more general theoretical schemes, in which the gravitational interaction is also included.[9,10] It follows from the theories indicated that fundamental constants might have different values in other cosmological epochs, as well as in spatially distant regions of the Universe. That these values are identical in different, even causally unrelated regions of space–time, is not a trivial fact. However, different versions of the theory predict different spatiotemporal dependences of the fundamental constants. Experiments are needed to ascertain which of the theoretical models is correct.

This report presents the results of a study of the spectra of distant extragalactic objects, which permit the establishment of new, the most reliable to date, upper limits on the possible spatiotemporal variation of α and μ.

2. TESTING METHODS

Experimental tests for the possible inconstancy of fundamental constants can be classified according to the spatiotemporal regions of the Universe covered by them. The first group is comprised of laboratory measurements, which cover no more than a few years. The second group includes ''local tests,'' which are concerned with the Earth and the solar system and cover up to 4.5 billion (4.5×10^9) years. Finally, the third group is comprised of astrophysical testing methods, which are based on data from extragalactic astronomy and cover practically the entire period of existence of the Universe.

Laboratory tests are generally confined to comparing frequency standards based on various physical phenomena, which consequently depend differently on the values of different physical constants. If the constants would vary, standards which were initially matched to one another would become unmatched with the passage of time. One of the standards compared is generally the cesium frequency standard, which is presently taken as a basis for defining a unit of

time. In particular, different groups of researchers have compared the output frequency of stabilized lasers and masers with it. The accuracy of these experiments permitted the detection of relative changes in the fundamental constants at the level of $\sim 6\times 10^{-14}$ per year, but no statistically significant effect was discovered.[11]

''Local tests'' are based on an analysis of changes which would occur in the motion of the Earth and other bodies of the solar system, as well as in the physical conditions and processes on them, if the fundamental constants would vary. Such an analysis is capable of providing a higher accuracy than laboratory measurements, mainly because it permits tracing changes which occur over the course of a far greater time. For example, by analyzing the isotopic composition of meteorites and ancient terrestrial rocks, we can estimate the characteristic alpha- and beta-decay times of the long-lived radioactive elements in them and compare these times to the half-lives measured in the laboratory. The half-lives of such elements are very sensitive to the value of the fine-structure constant α. On the basis of such an analysis Dyson[7] showed that the relative change in α cannot exceed $\sim 10^{-14}$ per year. The limit $|\dot{\alpha}/\alpha|<2\times 10^{-15}$ yr^{-1} was subsequently obtained on the basis of refined data on the beta decay of ^{187}Re (Ref. 12).

A more rigorous estimate was obtained by Shlyakhter.[13] He used data on the isotopic composition of Sm in the ''spent fuel'' of a natural nuclear reactor, which operated 1.8 billion years ago at the site of the contemporary Oklo uranium mine in Gabon. A more detailed analysis was recently performed by Damour and Dyson.[14] They concluded that the rate of relative variation of the fine-structure constant $\dot{\alpha}/\alpha$ is no greater than 10^{-17} per year.

The weak spot in the local tests just described is their dependence on the model of the phenomenon studied. The model is usually fairly complicated and includes a number of physical parameters. In particular, Damour and Dyson assumed in their analysis that the electrostatic (Coulomb) energy of the excited ^{150}Sm nucleus into which a ^{149}Sm nucleus is converted after neutron capture exceeds the electrostatic energy of a ^{150}Sm nucleus in the absence of excitation. However, it is known from experiments that such an hypothesis is far from always correct: the root-mean-square radii of excited nuclei can be greater than or smaller than the radii of the corresponding nuclei in the ground state.[15,16] This constraint becomes even more doubtful if it is taken into account that different physical parameters are interdependent and could have varied synchronously during cosmological evolution. As was shown in Ref. 17, this could have weakened the limit by more than an order of magnitude. In addition, it was noted in Refs. 18 and 9 that if the fundamental constants depend nonlinearly on time, as is assumed in modern theories, constraints which are valid for one time interval do not apply to another. As we have noted above, there might also be a simultaneous spatial dependence. Thus constraints which are valid for the solar system cannot be arbitrarily extended to more distant regions of space and to earlier stages in the life of the universe. Only extragalactic astronomy affords the possibility of investigating these regions of space–time.

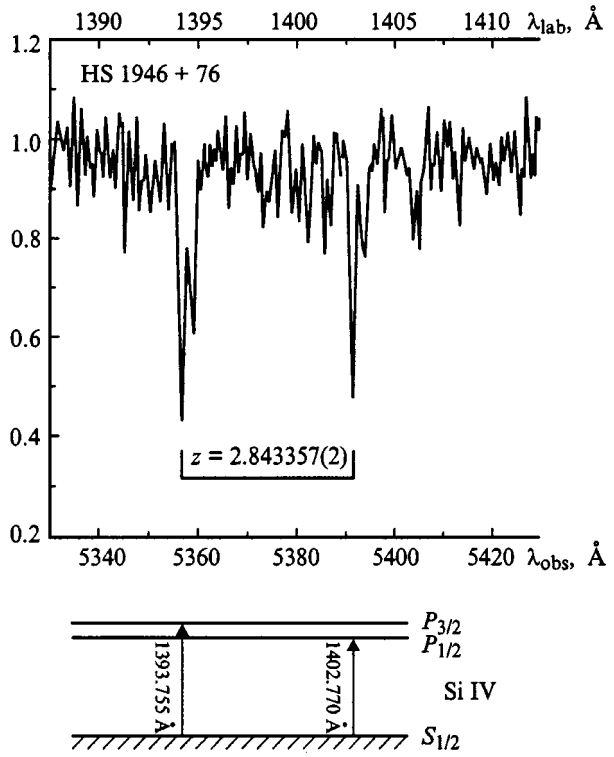

FIG. 1. Finely split doublet of the Si IV resonance line in the spectrum of the quasar HS 1946+76 recorded on the 6-meter telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences. Scheme of energy levels and transitions corresponding to the lines indicated.

For this purpose it would be useful to analyze the spectra of quasars, which are very powerful sources of radiation and can be seen for distances up to 10–15 billion light years. Along the way, the light coming from them has passed through clouds of interstellar gas in early galaxies and has been partially absorbed by them. Therefore, the spectral lines ''imprinted'' in quasar spectra contain information on the physical conditions and the state of matter in the early stages of the evolution of the Universe in various spatially distant regions of it.

The expansion of the Universe has resulted in increases in the wavelengths of the lines observed in quasar spectra (λ_{obs}) in comparison to their laboratory values (λ_0) according to the proportionality

$$\lambda_{obs}=\lambda_0(1+z),$$

where z is the cosmological red shift.

The value of z can be used to determine t, i.e., the ''age'' of the Universe in the epoch when the spectral line was formed. For example, Figs. 1 and 2 present fragments of the spectra of two quasars and exhibit absorption lines with $z\approx 2.8$. These absorption lines were formed when the Universe was 7–8 times younger than now. To obtain rigorous estimates of the variations of the fundamental constants, the lines studied must be sufficiently narrow and they must be recorded with a high spectral resolution and with a high

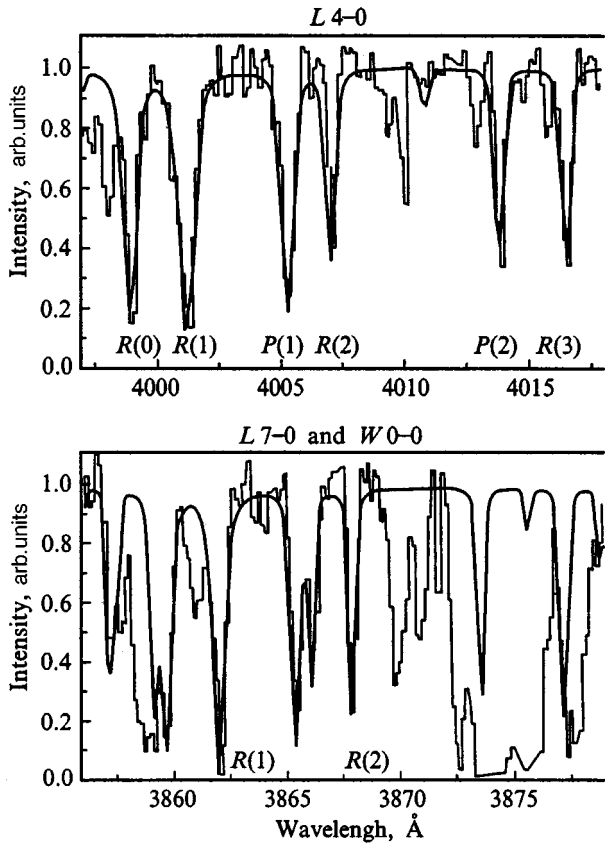

FIG. 2. Fragments of the spectrum of the quasar PKS 0528−250 containing absorption lines of H_2 in the Lyman and Werner series.

signal-to-noise ratio. The lines of the Si IV doublet shown in Fig. 1 and many of the lines of molecular hydrogen in Fig. 2 satisfy this condition.

3. DOES THE FINE-STRUCTURE CONSTANT VARY?

In order to ascertain whether the value of α has varied during cosmological time, it would be useful to investigate the fine splitting of the doublet lines of the Si IV, C IV, Mg II, and other ions, which are often observed in the spectra of distant quasars. An example of such spectral lines is shown in Fig. 1. The relative magnitude of the splitting $\delta\lambda/\lambda$ of these lines is proportional to α^2 (to within negligibly small corrections). Therefore, if the value of α varied with time, the relative splitting $\delta\lambda/\lambda$ should depend on the magnitude of the red shift.

In the first stage of our work we re-analyzed all the published data on finely split doublet lines observed in quasar spectra and complied a special catalog of wavelengths of these lines.[19] Altogether we examined about 1500 pairs of doublet lines with red shifts from 0.2 to 3.7. The analysis of these data demonstrated the absence of a statistically significant deviation of $\delta\lambda/\lambda$ from the current value. The rich observational material assembled in our catalog also allowed us to analyze the possibility of differences in the values of α in regions of the Universe that are causally unrelated to one another.[20] It was found that the dependence of the magnitude of the fine doublet splitting (and thus α) on the direction in the celestial sphere falls within the error $|\Delta\alpha/\alpha|<3\times10^{-3}$.

It should, however, be noted that in the overwhelming majority of cases the goal of the observers was not to perform exact measurements of the fine splitting; therefore, a large part of the data treated did not have a very high accuracy. The analysis performed allowed us to optimize the strategy for investigating the dependence of α on z. In the second stage of the work, our program of spectral observations of several quasars, which was aimed at achieving the highest possibility accuracy in measurements of fine splitting with large red shifts, was carried out on the 6-meter telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences.[21] As a result of the implementation of this program, as well as the use of observational data from other telescopes,[22–24] we determined the mean value of the possible deviation of the fine-structure constant for $z=2-4$ from its value for $z=0$ and the measurement errors (both the statistical and systematic errors)[25]

$$\Delta\alpha/\alpha=(-4.2\pm5.4[\text{stat}]\pm8.0[\text{syst}])\times10^{-5}. \quad (1)$$

A comparison of the data obtained on telescopes located in the northern and southern hemispheres allows us to tighten the constraint on the possible dependence of the fine-structure constant on the direction in the celestial sphere to $|\Delta\alpha/\alpha|<2\times10^{-4}$.

A result which has a formally higher accuracy was recently obtained on the basis of observations on the Keck I 10-meter telescope: $\Delta\alpha/\alpha=(-1.88\pm0.53[\text{stat}])\times10^{-5}$ for $z=0.6-1.6$. However, Webb *et al.*[26] did not take into account the systematic error, which results mainly from the uncertainty in the values of the reference (laboratory) wavelengths, and, as can be seen from (1), this error is dominant.

The upper limit corresponding to (1) (at the 2σ significance level) on the mean rate of variation of α during ~10 billion years is

$$|\dot{\alpha}/\alpha|<2\times10^{-14}\ \text{yr}^{-1}. \quad (2)$$

This constraint is five times tighter than the one which we previously obtained,[19,8] and its accuracy is 3.5 times better than that of the precision laboratory measurements in Ref. 11.

4. DOES THE PROTON-TO-ELECTRON MASS RATIO VARY?

The dimensionless constant $\mu=m_p/m_e$ is approximately equal to the ratio between the nuclear strong interaction coupling constant $(g^2/\hbar c)\sim14$ and the electromagnetic coupling constant $\alpha\approx1/137$. Here g is the effective coupling constant, which is calculated from the scattering amplitude of π mesons on nucleons.

The absorption lines of molecular hydrogen H_2 in quasar spectra can be used to ascertain whether the value of μ has varied during cosmological time. The variation of μ can be detected by comparing the wavelengths of various lines in a quasar spectrum and in the laboratory. The key to this technique is that the wavelengths of different lines depend differently on the parameter under study. This makes it possible to separate the cosmological red shift from the shift caused by the variation of μ.

The most suitable system of molecular lines with a large red shift for such an analysis is the system of absorption lines of the H_2 molecule with $z=2.8108$, which was observed by Levshakov and Varshalovich[27] in the spectrum of the quasar PKS 0528–250. We calculated the sensitivity coefficients of the wavelengths of H_2 toward the possible variation of μ (Ref. 28) and analyzed the observational data,[29] using the high-quality spectrum of PKS 0528–250 recorded by Lanzetta *et al.* on the CTIO 4-meter telescope. A fragment of this spectrum is shown in Fig. 2. The results of measurements of the wavelengths of 50 lines of molecular hydrogen with consideration of the sensitivity coefficients just mentioned led to the following estimate of the deviation ($\Delta\mu$) of μ in the epoch corresponding to $z=2.8108$ from its current value:

$$\Delta\mu/\mu=(-11.5\pm7.6)\times10^{-5}. \quad (3)$$

In addition, a multivariant statistical analysis of the measured spectrum, including a combined fit of a large number of H_2 lines with respect to their width and intensity, was performed. This analysis yielded

$$\Delta\mu/\mu=(8.3^{+6.6}_{-5.0})\times10^{-5}. \quad (4)$$

Both estimates are consistent within 2σ with the zeroth hypothesis that there is no variation of μ. Each of them corresponds to a constraint on the mean (during 10 billion years) rate of relative variation of the proton-to-electron mass ratio at the level

$$|\dot{\mu}/\mu|<2\times10^{-14}\ \ \mathrm{yr}^{-1}. \quad (5)$$

5. CONCLUSION

An analysis of quasar spectra has enabled us to establish tight upper constraints on the possible variation of α, i.e., the electromagnetic coupling constant, as well as the proton-to-electron mass ratio $\mu=m_p/m_e$. These quantities could have varied if the coupling constants of the strong and electroweak interactions had undergone changes. However, it was found that these constants did not vary within the statistical error during the 10^{10} years which have passed since the formation of the spectra of distant quasars. The upper limits found constrain the mean rate of possible variation of these parameters at the 0.02% level during a period of the order of 10 billion years, which covers 80–90% of the time of existence of the Universe. In addition, the values of the fundamental constant α were found to be identical (within a statistical error of 0.02%) in regions of the Universe which were not causally related in the period when the spectra were formed. This finding can be regarded as an argument which supports the so-called inflationary cosmological model, which presumes a general superfast inflation of the Universe according to an exponential law in a very early stage of cosmological evolution (see, for example, Ref. 30).

As a final note, we stress once again that the constraints which we obtained on the basis of an analysis of quasar spectra cover much more extensive regions of space and time intervals (corresponding to cosmological red shifts up to $z\approx3.7$) than do local tests (for example, the Oklo phenomenon, which is assigned only to the epoch with $z=0.09$ and to one point in space). In addition, they are far less dependent on model assumptions. They can therefore be recommended as the most reliable limits to date. They can serve as effective criteria for selecting permissible theoretical models of elementary interactions which predict changes in physical constants on the cosmological time scale.

This work was carried out within the State Scientific-Technical Programs ''Fundamental Metrology'' and ''Astronomy'' and was supported by the Russian Fund for Fundamental (Grant No. 96-02-16849a) and the Kosmion Scientific-Training Center.

[1] L. B. Okun, ''Fundamental constants of nature,'' Preprint ITEP-TH–56/96, Moscow (1996), 10 pp.; http://xxx.lanl.gov/abs/hep-ph/9612249.
[2] T. Kinoshita, IEEE Trans Instrum. Meas. **IM-46**, 108 (1997).
[3] B. N. Taylor and E. R. Cohen, J. Res. Natl. Inst. Stand. Technol. **95**, 497 (1995).
[4] E. R. Williams, R. L. Steiner, D. B. Newell, and P. T. Olsen, Phys. Rev. Lett. **81**, 2404 (1998).
[5] E. A. Milne, **139**, 409 (1937).
[6] P. A. M. Dirac, **139**, 323 (1937).
[7] F. J. Dyson, in *Aspects of Quantum Theory*, edited by A. Salam and E. P. Wigner, Cambridge University Press (1972), pp. 213–236.
[8] D. A. Varshalovich and A. Y. Potekhin, Space Sci. Rev. **74**, 259 (1995).
[9] T. Damour and A. M. Polyakov, Nucl. Phys. B **423**, 532 (1994).
[10] J. D. Barrow, *Theories of Everything. The Quest for Ultimate Explanation*, Clarendon Press, Oxford (1991).
[11] J. D. Prestage, R. L. Tjoelker, and L. Maleki, Phys. Rev. Lett. **74**, 3511 (1995).
[12] M. Lindner, D. A. Leich, R. J. Borg *et al.*, Nature **320**, 246 (1986).
[13] A. I. Shlyakhter, Nature **264**, 340 (1976).
[14] T. Damour and F. Dyson, Nucl. Phys. B **480**, 37 (1996).
[15] R. Engfer, H. Schneuwly, J. L. Vuilleumier *et al.*, At. Data Nucl. Data Tables **14**, 509 (1974).
[16] G. M. Kalvius and G. K. Shenoy, At. Data Nucl. Data Tables **14**, 639 (1974).
[17] P. Sisterna and H. Vucetich, Phys. Rev. D **41**, 1034 (1990).
[18] W. J. Marciano, Phys. Rev. Lett. **52**, 489 (1984).
[19] A. Y. Potekhin and D. A. Varshalovich, Astron. Astrophys., Suppl. Ser. **104**, 89 (1994).
[20] D. A. Varshalovich and A. Yu. Potekhin, Pis'ma Astron. Zh. **20**, 883 (1994) [Astron. Lett. **20**, 771 (1994)].
[21] D. A. Varshalovich, V. E. Panchuk, and A. V. Ivanchik, Pis'ma Astron. Zh. **22**, 8 (1996) [Astron. Lett. **22**, 6 (1996)].
[22] P. Petitjean, M. Rauch, and R. F. Carswell, Astron. Astrophys. **291**, 29 (1994).
[23] L. Cowie and A. Songaila, Astrophys. J. **453**, 596 (1996).
[24] P. J. Outram, B. J. Boyle, R. F. Carswell *et al.*, Mon. Not. R. Astron. Soc. (1998) (in press).
[25] A. V. Ivanchik, Candidate's Dissertation [in Russian], St. Petersburg (1998), 71 pp.
[26] J. K. Webb, V. V. Frambaum, C. W. Churchill *et al.*, Phys. Rev. Lett. **82**, 884 (1999).
[27] S. A. Levshakov and D. A. Varshalovich, Mon. Not. R. Astron. Soc. **212**, 517 (1985).
[28] D. A. Varshalovich and A. Yu. Potekhin, Pis'ma Astron. Zh. **22**, 3 (1996) [Astron. Lett. **22**, 1 (1996)].
[29] A. Y. Potekhin, A. V. Ivanchik, D. A. Varshalovich *et al.*, Astrophys. J. **505**, 523 (1998).
[30] A. Linde, Phys. Scr. T**36**, 30 (1991).

Translated by P. Shelnitz

Symmetry of the nonlinear collision operator matrix and new prospects in the moment method for solving the Boltzmann equation

A. Ya. Énder and I. A. Énder

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia; St. Petersburg State University, 199034 St. Petersburg, Russia

(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 6–9 (September 1999)

Traditionally, the moment method has been used in kinetic theory to calculate transport coefficients. Its application to the solution of more complicated problems runs into enormous difficulties associated with calculating the matrix elements of the collision operator. The corresponding formulas for large values of the indices are either lacking or are very cumbersome. In this paper relations between matrix elements are derived from very general principles, and these can be employed as simple recurrence relations for calculating all the nonlinear and linear anisotropic matrix elements from assigned linear isotropic matrix elements. Efficient programs which implement this algorithm are developed. The possibility of calculating the distribution function out to $8-10$ thermal velocities is demonstrated. The results obtained open up prospects for solving many topical problems in kinetic theory. © *1999 American Institute of Physics.* [S1063-7842(99)00209-3]

1. In the early classical studies expansion of the distribution function in orthogonal polynomials was employed to calculate the first few linear moments, which were needed for problems in transport theory.[1] Burnett[2] and Grad[3] formulated the basic principles of the nonlinear moment method. However, far-reaching progress in this area could not be made because of the great difficulty in calculating the matrix elements of the collision integral.

In the moment method for solving the Boltzmann equation, the distribution function is expanded into a series in Hermite polynomials:

$$f(\mathbf{v},\mathbf{r},t)=M(\mathbf{v})\sum_{i=0}^{\infty} C_i(\mathbf{r},t)H_i(\mathbf{v}),$$

$$H_{r,l,m}(\mathbf{v})=S^r_{l+1/2}(v^2/\tilde{T}_0)Y_{l,m}(\theta,\varphi). \tag{1}$$

Here $M(\mathbf{v})$ is a Maxwellian, $S^r_{l+1/2}(v^2/\tilde{T}_0)$ are Sonine (Laguerre) polynomials, and $Y_{l,m}(\theta,\varphi)$ are spherical harmonics. The system of moment equations has the form

$$DC_p/Dt=\sum_{i,j=0}^{\infty} K^p_{i,j}C_iC_j. \tag{2}$$

The operator D/Dt corresponds to the substantial derivative on the left-hand side of the Boltzmann equation and is well known. In the isotropic case with respect to the velocities, the expansion of the distribution function (1) is written only in Sonine polynomials. The matrix element $K^n_{k,m}$ is the projection of the collision integral of the kth and mth Sonine polynomials onto the Sonine polynomial with the index n.

In computer calculations the series (1) must always be truncated at a certain finite N. Turchetti and Paolilli[4] were the first who knew how to calculate all the matrix elements up to $N=13$ for ''pseudopower-law potentials'' in the isotropic case. The formula for the matrix element in Ref. 4 contains six nested sums, which restrict the possibility of increasing N further. In Ref. 5 we used the α representation[6,7] to derive formulas for calculating matrix elements, which are valid for any power-law potentials, and the number of summations was reduced to four. It was shown that calculations can be performed up to $N=25-30$ with guaranteed accuracy.

The transition to the α representation corresponds to expanding the distribution function in Maxwellian distributions with various temperatures ($\alpha=1/\tilde{T}$):

$$f(v,t)=\int_0^{\infty} M(v,\tilde{T})\varphi(\tilde{T},t)\,d\tilde{T}. \tag{3}$$

Here $\varphi(\tilde{T},t)$ is a distribution function in α space. We previously obtained a closed form of the Boltzmann equation in the α and $\alpha-u$ representations. In $\alpha-u$ space the expansion is written in Maxwellians with different temperatures and with different shear rates. The very simple representations of Sonine and Hermite polynomials in α and $\alpha-u$ spaces, respectively, were important for further work. A system of the Sonine polynomials $S^r_{1/2}$ which is orthogonal with a Maxwellian weighting function corresponds to a biorthogonal system of the functions s_R and s_L in α space

$$s^r_R(\tilde{T},\tilde{T}_*)=(\tilde{T}_*)^r\delta^{(r)}(\widetilde{T-\tilde{T}_*})/r!,$$

$$s^r_L=(1-\tilde{T}/\tilde{T}_*)^r. \tag{4}$$

In the isotropic case the presentation (1) corresponds to the distribution function $\varphi(\tilde{T})$ in the form

$$\varphi(\tilde{T},t)=\sum_{n=0}^{\infty} C_n(t)s^n_R(\tilde{T},\tilde{T}_0),$$

$$C_n(t)=\int_0^{\infty}\varphi(\tilde{T},t)s^n_L(\tilde{T},t)\,d\tilde{T}. \tag{5}$$

2. Let us consider a simple example. As the initial distribution function we take the Maxwellian distribution with the temperature T_0. It is clear that regardless of the choice of basis functions, the collision integral of such a distribution function is equal to zero. If, as is usually done, we expand this distribution function in Sonine polynomials near the same temperature T_0, a unique condition follows from the equality of the collision integral to zero: $K^0_{0,0}=0$. It is clear that a transition to another basis set, i.e., re-expansion of the same distribution function in Sonine polynomials near another temperature, cannot lead to a nonzero value for the collision integral. It was shown in Ref. 7 that such re-expansion leads to the following relations between the matrix elements:

$$\sum_{k=0}^{m} K^n_{k,m-k}=0, \quad n,m=0,\ldots,\infty. \tag{6}$$

Developing this idea, we can state that the collision integral is independent of the choice of the temperature of the Maxwellian weighting function not only for a Maxwellian distribution function, but also for an arbitrary distribution function. Let us examine what follows from this for the isotropic Boltzmann equation. We expand the distribution functions in two basis sets with the temperatures T_0 and T_1:

$$\varphi(T,t)=\sum_{k=0}^{\infty} C^0_k(t)s^k_R(T,T_0)=\sum_{r=0}^{\infty} C^1_r(t)s^r_R(T,T_1). \tag{7}$$

The infinite-dimensional vectors $\mathbf{C}^0$ and $\mathbf{C}^1$ are related by the simple expression

$$C^1_r=\sum_{k=0}^{\infty} d_{r,k}(T_1,T_0)C^0_k, \tag{8}$$

where the elements of the transition matrix $D(T_1,T_0)$ for going from one basis set to another are easily determined using (4) and (5) and have the very simple form

$$d_{r,k}(T_1,T_0)=\begin{cases}\binom{r}{k}(T_1-T_0)^{r-k}T_0^k/T_1^r, & r\geqslant k,\\ 0, & r<k.\end{cases} \tag{9}$$

Comparing the form of the collision integral in two basis sets, from the condition of invariance of the collision integral relative to the choice of the basis set we can obtain the relation between the matrix elements in the basis sets T_1 and T_0:

$$K^r_{r_1,r_2}(T_1)=\sum_{k=0}^{r} d_{r,k}(T_1,T_0)\sum_{k_1=r_1,k_2=r_2}^{\infty} \times K^k_{k_1,k_2}(T_0)d_{k_1,r_1}(T_0,T_1)d_{k_2,r_2}(T_0,T_1). \tag{10}$$

We differentiate (10) with respect to T_1 and set $T_1=T_0$. We take into account in the process that if $T_1=T_0$ the operator $\hat{D}$ transforms into the unit operator $\hat{E}$, since $d_{r,k}(T_0,T_0)=\delta_{r,k}$. When the differentiation with respect to T_1 is performed, only the derivatives of $\hat{D}$ must be calculated on the right-hand side of (10). It is clear from (9) that the derivative of $d_{r,k}(T_1,T_0)$ for $T_1=T_0$ is nonzero only if $r=k+1$ or $r=k$. As a result, after some transformations we obtain the basic expression relating the matrix elements to one another

$$\left(T\frac{d}{dT}-R\right)K^r_{r_1,r_2}(T)=rK^{r-1}_{r_1,r_2}(T)-(r_1+1)K^r_{r_1+1,r_2}(T) -(r_2+1)K^r_{r_1,r_2+1}(T), \tag{11}$$

$$(R=r_1+r_2-r). \tag{12}$$

For power-law potentials ($V\sim 1/r^{\kappa}$) the matrix elements $K^r_{r_1,r_2}$ depend on the temperature of the basis set in a fully specified manner, i.e., $K^r_{r_1,r_2}(T)\sim T^{\mu}$, where $\mu=(\kappa-4)/(2\kappa)$ (Ref. 5). Equation (11) now takes the form

$$(\mu-R)K^r_{r_1,r_2}=rK^{r-1}_{r_1,r_2}-(r_1+1)K^r_{r_1+1,r_2} -(r_2+1)K^r_{r_1,r_2+1}. \tag{13}$$

The equality (13) can serve as recurrence relations for calculating the matrix elements. It can be shown that to determine the nonlinear matrix elements it is sufficient to assign the linear elements of either the $K^r_{r_1,0}$ or the K^r_{0,r_2} type. One half of the linear elements are then specified in terms of the other half. To close the recurrence procedure, we also derived simple formulas for calculating the linear matrix elements.

Employment of the recurrence procedure enabled us to perform calculations of all the nonlinear matrix elements up to $N=128$ in about 30 min on a PC with a 486 microprocessor. Similar calculations using the six-sum[4] and four-sum[5] formulas would require 9×10^4 and 10 years, respectively.

The results of the calculations of distribution functions using the moment system (2) and the recurrence formulas (13) were compared with the exact BKW solution of the Boltzmann equation for Maxwellian molecules. For $N=128$ excellent agreement was observed up to 12 thermal velocities. For other cases with $N=128$ the maximum rate at which the segment of the series (1) accurately describes the distribution function reached 8–10 thermal velocities. As the number of moments increased, it was always possible to disclose the behavior of the tails of the distribution function and to move increasingly further toward higher velocities with increasing N. Figure 1 presents the relative distribution function f/f_{∞} (f_{∞} is the equilibrium Maxwellian distribution function) for the hard-sphere model (a) and a Coulomb interaction between the particles (b) with the initial sum of two BKW modes

$$f(v,0)=a_1M(T_1,v)(1+d_1S^1_{1/2}(mv^2/2kT_1)) +a_2M(T_2,v)(1+d_2S^1_{1/2}(mv^2/2kT_2)),$$

in which we set $a_1=a_2=0.5$, $T_1=1/3$, $T_2=8/9$, $d_1=1/9$, and $d_2=5/18$.

3. In contrast to the isotropic Boltzmann equation, in the anisotropic case there are no general formulas for calculating the nonlinear matrix elements (with the exception of Maxwellian molecules). In the axisymmetric case the distribution function is expanded in the Hermite polynomials $H_{r,l}$, where

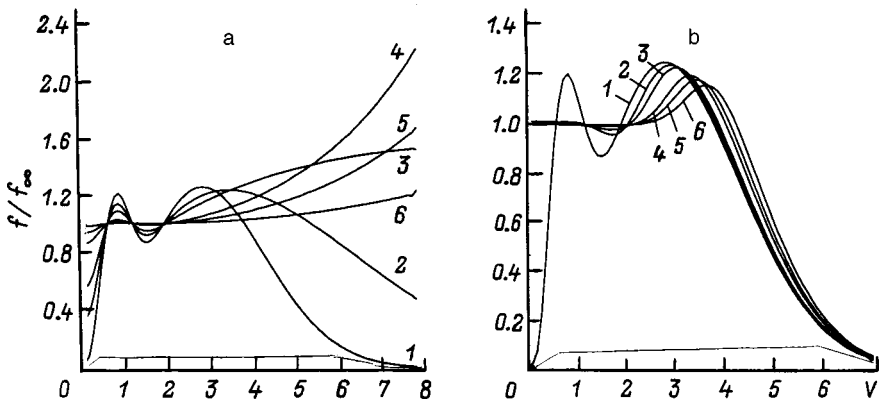

FIG. 1. Solution of the nonlinear Boltzmann equation with an initial sum of two BKW modes:a — for the hard-sphere model, b — for the Coulomb interaction model; *1–6* — $t/\tau^*=0$, 1, 2, 5, 7, and 10 (τ^* is the characteristic relaxation time).

$$H_{r,l}=c^l P_l(\cos\theta)S^r_{l+1/2}(c^2) \tag{14}$$

$$c^2=\frac{m}{2kT_0}(v_x^2+v_y^2+(v_z-u_0)^2), \tag{15}$$

which are orthogonal with the Maxwellian weighting function $M(c,T_0)=(m/(2kT_0\pi)^{3/2}\exp(-c^2))$.

The relations between the matrix elements $K^{r,l}_{r_1,l_1,r_2,l_2}$ are found from the invariance of the collision integral as the velocity u_0 and the temperature T_0 of the basis functions are varied:

$$\frac{l}{2l-1}K^{r,l-1}_{r_1,l_1,r_2,l_2}-\frac{(l+1)r}{2l+3}K^{r-1,l+1}_{r_1,l_1,r_2,l_2}$$
$$-\frac{(l_1+1)}{2l_1+1}K^{r,l}_{r_1,l_1+1,r_2,l_2}+\frac{l_1(r_1+1)}{2l_1+1}K^{r,l}_{r_1+1,l_1-1,r_2,l_2}$$
$$-\frac{(l_2+1)}{2l_2+1}K^{r,l}_{r_1,l_1,r_2,l_2+1}+\frac{l_2(r_2+1)}{2l_2+1}K^{r,l}_{r_1,l_1,r_2+1,l_2-1}=0, \tag{16}$$

$$T\frac{d}{dT}K^{r,l}_{r_1,l_1,r_2,l_2}(T)=(R+L/2)K^{r,l}_{r_1,l_1,r_2,l_2}(T)$$
$$+rK^{r-1,l}_{r_1,l_1,r_2,l_2}(T)-(r_1+1)$$
$$\times K^{r,l}_{r_1+1,l_1,r_2,l_2}(T)-(r_2+1)$$
$$\times K^{r,l}_{r_2+1,l_2,r_1,l_1}(T), \tag{17}$$

$$R=r_1+r_2-r,\quad L=l_1+l_2-l. \tag{18}$$

The results in Ref. 8 were used in the derivation of these relations. The relations obtained are valid in cases where there is a preferred direction in the system. Such a system can consist of, for example, particles immersed in an exceptionally strong magnetic field (of the order of 10^5 T). In this case the particles become asymmetric and are oriented along the magnetic field.

In standard kinetic theory (in the absence of a preferred direction) space is isotropic, and Hecke's theorem (see, for example, Ref. 9), which states that the matrix elements $K^{r_1,l}_{r_1,l_1,0,0}$ can be nonzero only if $l_1=l$, holds for the linear matrix elements. Using the recurrence relations (16)–(18), we can show that the generalized Hecke theorem holds for the nonlinear matrix elements: the matrix elements $K^{r,l}_{r_1,l_1,r_2,l_2}$ can be nonzero only if

$$|l_1-l_2|\leqslant l\leqslant l_1+l_2 \tag{19}$$

and the parity of l coincides with the parity of l_1+l_2. This theorem was previously proved in Ref. 10 by another method without employing recurrence relations.

The next important new fact in the case of unoriented particles, which follows from (16)–(18), is the relation between the linear anisotropic and linear isotropic matrix elements. It was shown that if the linear isotropic matrix elements are known, then all the remaining matrix elements can be uniquely specified in terms of them. In the few special cases for which explicit formulas existed[8,11] the results of our calculations coincided with the known results.

In the general three-dimensional case the spherical harmonics $Y_{l,m}(\theta,\varphi)$ appear in the Hermite polynomials. In this case there are possibilities for additional changes in the basis set. They include rotation of the reference system about the z axis, i.e., the transition $\varphi\rightarrow\varphi+\omega$, and deflection of the z axis from the original direction, i.e., $\theta\rightarrow\theta+\psi$. Relations obtained from the invariance of the collision integral relative to such changes in the basis set were derived. In the absence of a preferred direction in space the familiar Clebsch–Gordan coefficients satisfy such recurrence relations. Hence it follows that the matrix elements with indices $m\neq 0$ are related to the axisymmetric matrix elements through these coefficients. This fact was previously known (see, for example, Ref. 10). We note that fundamentally new results are obtained for asymmetric oriented particles.

Thus, it has been shown that a hierarchy of simple recurrence relations can be constructed in standard kinetic theory from a few principles, and all the matrix elements can be found, if the linear isotropic matrix elements of one type are known. Programs which implement these algorithms have been written. The results obtained open up possibilities for solving numerous unsolved problems in the kinetic theory of gases and for refining the already existing solutions. These problems include, for example, 1) the description of transport processes for Kn$\sim$1 both in gases and in a highly ionized plasma and the creation of a theory of transport for asymmetric oriented particles in a strong magnetic

field; 2) the kinetic description of the structure of weak shock waves and their nonstationary interactions; 3) the investigation of the distribution function at high velocities in order to describe physicochemical processes.

This work was carried out in part with support from the Russian Fund for Fundamental Research (Project No. 97-02-18080) and in part with support from the Federal Special-Purpose Program ''Integration'' (Program No. 326.53).

[1] S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases*, 2nd ed. [Cambridge University Press, Cambridge (1952); IL, Moscow (1960)].
[2] D. Burnett, Proc. London Math. Soc. **40**, 389 (1935).
[3] H. Grad, Commun. Pure Appl. Math. 331 (1940).
[4] G. Turchetti and M. Paolilli, Phys. Lett. A **90**, 123 (1982).
[5] A. Ya. Énder and I. A. Énder, Zh. Tekh. Fiz. **64**(10), 38 (1994) [Tech. Phys. **39**, 997 (1994)].
[6] I. A. Énder and A. Ya. Énder, Dokl. Akad. Nauk SSSR **193**, 61 (1970) [Sov. Phys. Dokl. **15**, 633 (1971)].
[7] I. N. Kolyshkin, A. Ya. Énder, and I. A. Énder, Zh. Vychisl. Mat. Mat. Fiz. **28**, 901 (1988).
[8] I. A. Énder and A. Ya. Énder, *Physical Kinetics* [in Russian], Izd. SPbGU, St. Petersburg (1983), pp. 197–215.
[9] L. Waldmann, ''Transporterscheinungen in Gasen von mittlerem Druck,'' in *Thermodynamik der Gase (Handbuch der Physik, Vol. 12)*, edited by S. Flügge [Springer-Verlag, Berlin (1958); pp. 295–514 Moscow (1970), pp. 169–414].
[10] V. V. Vedenyapin, IPM AN SSSR Preprint No. 59 [in Russian], Institute of Applied Mathematics, Academy of Sciences of the USSR, Moscow (1981), 15 pp.
[11] J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* [North-Holland, Amsterdam (1972); Mir, Moscow (1976)].

Translated by P. Shelnitz

Danger of the explosion of Callisto and the priority of space missions

É. M. Drobyshevskiĭ

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 10–14 (September 1999)

Ice is a protonic conductor, as has been demonstrated many times by electrolysis experiments. The dirty ices which comprise the thick ($\sim 10^3$ km) crusts of several distant moonlike bodies are subjected to bulk electrolysis by currents excited by the motion of such bodies in cosmic magnetic fields (for example, Voyager-1 measured a current amounting to $\sim 10^7$ A flowing through the Jovian satellite Io and its surroundings). The accumulation of electrolysis products in ice in amount equal to 10–15 wt. % renders such a solid solution capable of detonation. Global explosions of the crusts of moonlike bodies account for the origin and the known properties of many asteroids, short-period comets, planetary rings and small satellites, the formation of Titan's atmosphere, the differences between Jupiter's Galilean satellites, etc. Many predictions made on this basis have already been confirmed, and others are awaiting testing. According to all the signs, only the ices of the fourth Galilean satellite Callisto have not yet exploded. If they explode, the Earth will be subjected to concentrated bombardment by cometary nuclei, which will create a ''nuclear winter'' once every 60 years on the average. Therefore, a very high priority should be assigned to *in situ* investigations of Callisto for the purpose of determining the degree of saturation of its ices with electrolysis products. [S1063-7842(99)00309-8]

1. TWO RESEARCH STRATEGIES AND TWO COSMOGONIES OF THE SOLAR SYSTEM

Investigations of new phenomena can be divided into two stages: 1) an extensive (passive) stage, in which, without any guiding ideas, simple observations are performed and highly diverse facts are collected in an attempt to somehow link them to one another, and 2) an intensive (active) stage, in which there is already a working hypothesis and a search for predicted relationships and phenomena is conducted on its basis for the purpose of refining and confirming it. All the new facts obtained in the last 10 years during *in situ* investigations of solar system bodies are ''unexpected,'' ''puzzling,'' etc. This suggests that we are now in the first stage of investigation and that the 200-year-old ''obvious,'' but naive Kant–Laplace scenario of the origin of the solar system and its bodies from a vast disk of gas and dust surrounding the Sun, a scenario which has hitherto guided the planners of space missions, is fundamentally untrue. In fact, it was created *ad hoc* to explain the appearance of our planetary system only and at that time could not, in principle, be based on knowledge of the origin and evolution of stars and their systems, gas dynamics, etc. This is why none of its predictions have come true and it cannot explain a multitude of old and new facts without eliciting additional hypotheses.

The new planetary cosmogony is based on the modern understanding of the formation and evolution of multiple stars, a category to which almost all stars belong. It assumes that planetary systems appear in them as a by-product of their evolution or are a limiting case of close binaries formed as a result of the rotational–exchange breakup of a dense, rapidly rotating protostar, i.e., the last product of the multistage fragmentation of the original cloud of gas and dust as a consequence of the presence of the standard excess of angular momentum in it.[1,2] This approach, which gives a short ($\sim 10^4 - 10^5$ years) time scale for the formation of planets within a very dense, rapidly rotating proto-Jupiter, whose mass quickly drops from $\sim 1 M_\odot$ to $0.001 M_\odot$ due to outflow to the Sun and which therefore loses the planets being formed in it, consistently accounts for all the known facts. The recent discovery by S. Terebey *et al.* using the Hubble telescope (NASA Press Release 98-19) of a giant planet which was ejected from a newly forming pair of binary stars supports the new cosmogony and confirms one of its numerous predictions (see stage 5 in Fig. 1 in Ref. 2). Another verifiable prediction is the presence of a combined planet–comet cloud on the not too distant (50–3000 AU) periphery of the solar system.[2,3] This cloud is the main source of long-period (LP) comets (with periods >200 years) and contains $\sim 10-100$ as yet undiscovered moonlike bodies resembling Pluto, and larger bodies. Another important consequence is the creation of a totally new cosmogony for small solar system bodies. The latter, in turn, has consequences which may be very important for all of mankind. This approach at once shifts the investigation of solar system bodies into the second, or active, stage.

2. CLASSICAL COMETARY COSMOGONY. OLD AND NEW PROBLEMS

Within the Kant–Laplace hypothesis small bodies, viz., asteroids, comets, Trojans, small satellites, and planetary rings, are regarded as construction debris, which has remained since the time of formation of the planets. Conse-

quently, they cannot contain sources of energy, apart from insignificant sources (gas pockets, *ad hoc* amorphous ice capable of slowly crystallizing, etc.). It is believed that the initially created reservoirs of comets and other small bodies have scarcely been replenished and that they are only depleted with time. The traditional scenario of cometary activity presumes the sublimation of their dirty ices under the action of solar radiation followed by photolysis of the escaping vapors. The first such model, viz., the Vsekhsvyatskiĭ–Whipple ice model, was a great step in understanding comets. However, more attentive scrutiny revealed numerous inconsistencies in this picture. We shall list the most significant.

1) The great diversity in the manifestations of different comets. It cannot be ascribed only to their age. It cannot be understood in terms of a postulated condensation of cometary nuclei in a thoroughly stirred, highly agitated medium, which the gas in the circumsolar disk should be.

2) There is the old problem that the lifetime of the so-called ''parent'' molecules (in concept; no one knows what they actually are) in the field of solar radiation is much longer in laboratory simulations as compared with observations.

3) The causes of the burst activity of comets are not clear. It correlates with solar activity and is observed over distances up to 10 AU or more, where water ice does not sublime.

4) Bursts quite often precede the breakup of the nucleus into kilometer fragments that diverge with a velocity of 1–10 m/s, which requires a large amount of energy.

Numerous attempts to solve these problems have the distinct features of *ad hoc* hypotheses and frequently disregard not only other facts but even conservation laws. None of these attempts has had any clear-cut success.

The flight of an armada of space vehicles near the nucleus of Comet P/Halley in 1986 and the recent passage of the nucleus of Comet P/Shoemaker–Levy 9 near Jupiter, its breakup, and the descent of fragments onto Jupiter added many new facts, which do not have simple explanations within the traditional theories. We shall list the principal ones.

P/Halley: 1) jets of dusty vapors escaping from spots which comprise $<10\%$ of the area of the illuminated surface, despite the need for the solar energy impinging on 40% of the surface to sublime these vapors; the temperature of the dark surface of the nucleus, which is unexpectedly high ($\approx$400 K) for ice; 2) higher concentrations and temperatures of radicals, as well as positive and negative ions and heavy organic matter, closer to the nucleus (according to measurements as close as 700 km); 3) the presence of extremely active atomic carbon and ions of this element in the circumnuclear region; 4) a large excess of CO over CO_2 in the vicinity of the nucleus; 5) the discovery of numerous submicron CHON particles, which are unstable in the field of solar radiation; 6) cases of essentially instantaneous (within 9 min) vaporization of 4×10^8 kg of water ice; 7) a distant (at 14.3 AU) (double?) burst, which correlates with passage through the solar wind sector boundary.

P/Shoemaker–Levy 9: a) breakup of the nucleus into 21 fragments; 2) breakup events occurring weeks and even months after passage of the nucleus through the Roche limit of Jupiter, which cannot be attributed to the tidal forces of Jupiter alone; 3) acceleration of the large fragments formed upon breakup to relative velocities of 1 m/s; 4) the far greater relative velocity of the dust in comparison to the velocity of the large fragments; 5) an essential absence of traces of water in SL-9 at the time of its discovery 9 months after passing the perijove; 6) a local excess of Li in one part of the SL-9 nucleus, specifically in fragment L.

These examples can be supplemented by the classical recurrent bursts of Comet P/Schwassmann–Wachmann 1 beyond Jupiter, which are correlated with solar activity, and the inexplicable presence of CO^+ in it, as well as the discovery of the cometary activity of distant Chiron with the detection of CN on it.

3. GALILEAN SATELLITES, TITAN, AND THE NEW ERUPTIVE COSMOGONY OF SMALL BODIES

It follows from the new cosmogony of the solar system that all the Galilean satellites, which were among the last (and therefore unlost) planets formed in proto-Jupiter, should have a similar composition. This is not the case. It should be noted that the differences between the satellites cannot be explained within the old cosmogony, either.

A thorough study of the causes of these differences revealed that the detonation of dirty ices, which originally comprised the thick (with a thickness of $\sim700-800$ km as compared to a radius of the body equal to $\approx$2500 km) crusts of the satellites and became saturated with up to $\sim10-15$ wt. % bulk electrolysis products, i.e., $2H_2+O_2$, can be responsible for them. The electrolysis of solid ice, which is a protonic conductor, is a thoroughly understood and well-studied process.[4,5] It occurs in dirty ice on the boundaries of the numerous, fairly large mineral inclusions, which have electronic or hole conductivity. The electric current needed is generated by the motion of the moonlike body in the magnetic field of the planet or the solar wind. Owing to the high pressure, $2H_2+O_2$ is not released in the form of a gas, but accumulates in the form of a solid solution in the ice and is redistributed by solid-state convection throughout the crust.

The explosion of an icy Pluto-like planet, which had satellites, between Mars and Jupiter about 3.9 billion years ago probably produced the asteroids of the main asteroid belt[6,7] and, together with their subsequent collisional evolution, accounts for all of their known properties. It provides the source for the concentrated Imbrian bombardment of the Moon and permits evaluation of the intensity of the solar wind at that time ($\sim2\times10^{-11}\ M_\odot$/yr, in agreement with more recent estimates of other investigators). As a consequence of three global explosions of Io's crust, it lost all of its ices. Europa suffered two explosions, and Ganymede had one (probably 570 million years ago at the boundary between the Cryptozoic and Phanerozoic eons, which marks the date of a major geological catastrophe). Callisto's ices have not yet exploded.[8] All this explains the differences between the satellites, particularly the totally different topography of Ganymede and Callisto. The surface of the latter is the most

cratered in the solar system, but the size distribution of the craters differs strongly from that for the Moon and Mercury. On Ganymede there are two types of surfaces with different degrees of cratering, but despite expectations (with consideration of the focusing effect of Jupiter) the degree of cratering here is less than on Callisto (for more information on other details which can be explained by explosions, see Refs. 9 and 10).

The formation of Saturn's rings, Titan's atmosphere, and some comets is due to a recent explosion of the icy crust of that satellite. In fact, the distribution of LP comets of the Saturnian family with respect to Δ (the smallest distances between a comet's orbit and Saturn's orbit $\Delta \leqslant 2$ AU) has a sharp peak at $\Delta \leqslant 0.5$ AU (the radius of the Hill sphere of Saturn equals 0.57 AU). The LP comets of the Jovian family do not exhibit a similar peak.[11] If the significant eccentricity of Saturn's orbit ($e=0.056$) and the fact that its major axis completes one revolution after $P=47\,000$ yr are taken into account, it can be assumed that such a narrow distribution peak could still exist if its age is $\leqslant P/4 \approx 10^4$ yr.[12,13] Otherwise, some comets would have moved beyond the $\Delta=0.5$ AU limit. It is interesting that the concentration effect at $\Delta<0.5$ AU is even stronger for brighter (=younger) comets.[14]

Under the assumption of an explosion on Titan, the author, prior to the Voyager mission there, predicted[12] the existence of two populations (with centimeter and kilometer dimensions) in the rings, the presence of HCN, N_2, CO, and CO_2, i.e., products of the explosion of dirty ices, in Titan's atmosphere (the presence of methane was already known), and the effective radius of its solid crust. These predictions were all confirmed. The predictions of a fairly thin ice crust with light organic matter (1–10 km) on a deep (~700 km) liquid ocean, an excess thermal flux as a result of freezing of the ocean, and a similarity of the surface topograph of Titan, which is still being formed, to the topography of Ganymede, which has already refrozen since the explosion of its crust, are awaiting confirmation.

It therefore follows that fragments of the outermost layers of the exploding icy crusts of moonlike bodies containing ~10% organic matter and a chondritic mineral matrix (~10%) should also be saturated with $2H_2+O_2$ to a concentration close to the level necessary for combustion or detonation. It is natural to presume that these are the nuclei of short-period (SP) comets. Thus they have a powerful internal source of chemical energy, which can be activated by solar radiation or an electric current.

The former case is realized when sublimation products containing light organic matter and $2H_2+O_2$ are illuminated by solar radiation. It initiates ordinary combustion, which is capable of penetrating along jets of vapors to sites rich in $2H_2+O_2$ in the ice of a nucleus.

The latter case occurs when a nucleus travels in a strong circumplanetary magnetic field (the case of SL-9) or crosses current-carrying zones between solar wind sectors. Here electric currents flow along the gas jets, which are ionized by combustion, into the nucleus, where they are capable of igniting regions enriched with $2H_2+O_2$, including those which do not communicate with the surface.

The entire set of combustion processes along with their unexpected bursts and even detonation in some sites and smoldering in others, whose original cause is the presence of electrolysis products dissolved in the ice, readily explains all the diverse and puzzling manifestations of comets mentioned above. In fact, the presence of various (hot) ions and radicals near the nucleus, as well as atomic and molecular (C_2 and C_3) carbon (soot), a high CO/CO_2 ratio, CHON particles (smoke), etc. are explained at once by combustion taking place in sublimation products under conditions of a deficiency of the oxidizing agent.[15] The absence of water in the fragments of SL-9 may be due to the circumstance that the volatile substances had completely burned up at the time the comet was discovered, while the detection of Li in only one part of the nucleus can be attributed only to geochemical differentiation in the crust of a parent planet.[16]

The appearance of bursts, which frequently precede the breakup of a nucleus, is also obviously caused by the initiation of combustion or even detonation in such energized ices.

On the basis of the ideas developed above, it is fairly easy to give consistent explanations for other manifestations of comets.

The discovery of S_2 in the coma of several comets,[17] the fact that the concentration ratio of ortho-to-para water molecules emitted by nuclei is characteristic of a temperature of 25–50 K,[18] and the fairly prevalent opinion regarding the low density of nuclei ($0.2-0.3$ g/cm^3) are often interpreted as supporting the condensation cosmogony of comets. It is not difficult to see that the first two factors are simply explained by combustion with rapid expansion of the products in free space. The low density estimates are based on calculations of the reactive effect of escaping matter on the nucleus for a known rate of mass loss by it. The results of the estimates depend on the model adopted. In the published reports (for example, Ref. 19) no allowance was made for the jet character of a real escaping flow (it is known from gas dynamics that a jet has greater momentum than a flow escaping from a surface), or its consequent gas-dynamic rather than free-molecular character in this case. Of course, the higher initial heat content of the combustion products was not taken into account (a temperature of the subliming ice equal to only ≈200 K was taken). Consideration of these factors lowers the density of nuclei to $1-1.5$ g/cm^3.

The small fragments remaining from the explosion of a large ice fragment saturated with $2H_2+O_2$ are capable of acquiring velocities up to 5 km/s. Thus, the Martian satellites, the irregular Jovian satellites, and Trojans could be located in their orbits. Such an approach explains their known properties and provides predictions that can be tested (for example, the presence of ices in some Trojans and in Deimos and their absence in Phobos, the escape of some comets form the Trojan reservoir, etc.).[20,21]

4. CONSEQUENCES OF THE POSSIBLE EXPLOSION OF CALLISTO FOR THE EARTH

According to all the information available, Callisto's crust has not yet exploded. We shall list the main consequences of its possible explosion.[9,10]

1) A loss of $\sim 15\% \approx 1.6 \times 10^{25}$ g of mass, mainly in the form of water vapor, pyrolyzed organic matter, sand and stone, i.e., the chondritic mineral component originally contained in the dirty ice, as well as, more importantly, unexploded fragments of the outermost layers of the crust. They can measure up to several tens of kilometers. Their ice will also be saturated with $2H_2 + O_2$, which will render them capable of undergoing further explosions or combustion. They will be typical nuclei of SP comets.

2) After a year or two, Jupiter's sphere of influence will transform into a luminous body measuring $\sim 15°$, which will illuminate the Earth with scattered light ~ 60 stronger than that of the full Moon.

3) The collisions of ice fragments will cause new explosions, so that the secondary fragments with a total mass of $\sim 10^{24}$ g, which will acquire an additional velocity increment, will be capable of leaving Jupiter's sphere of influence and attaining heliocentric orbits of SP comets. If they have the size distribution characteristic of asteroids of the main asteroid belt, the number of such spheres with diameters >300 m will be $\sim 10^9$.

4) After 10–30 years the ice and mineral dust propagating along Jupiter's orbit will illuminate the earth $\sim 2-3$ times more intensely than the full Moon.

5) After several decades concentrated bombardment of the Earth by SP comets will begin. It will last $10^4 - 10^5$ years. It is known from the existing statistics that the mean probability of an SP comet of the Jovian family striking the Earth is $\sim 0.45 \times 10^{-9}$ per year for an impact velocity of ≈ 29 km/s (Refs. 9 and 22). Hence it follows that a fragment whose impact will be accompanied by the release of the energy of $\geqslant 1$ megatons of TNT will strike the Earth once a day, a fragment releasing the energy of $\geqslant 10^3$ megatons of TNT will strike once a year, and a fragment releasing the energy of $\geqslant 10^5$ megatons of TNT will strike once in 60 years. Impact of such force will throw so much dust into the atmosphere that it will cause a nuclear winter.[9,10]

5. MAIN CONCLUSIONS

1. Bodies of small mass are not remnants from the time of the formation of the solar system. They are products of the subsequent explosive and/or collisional evolution of (icy) moonlike (Pluto-like) planets. Only the new eruptive cosmogony of small bodies, which is based on the physically fairly reliable hypothesis of the possibility of the bulk electrolysis of ices, offers both a simple explanation for all the known high-energy manifestations of SP comets and the properties of other small bodies (asteroids, Trojans, and planetary satellites) and numerous predictions concerning the properties of these bodies and other bodies related to them.

Many of the predictions have already been confirmed, and others are awaiting testing. The new eruptive cosmogony has not run up against any substantiated objections in its 15 years of existence. Its wide acceptance has probably been held up only by psychological factors. The critical articles (for example, Ref. 23) appeal to emotions and are based either on unattentive reading of the publications on the new eruptive cosmogony or an unwillingness to study the elements of electrochemistry, gas dynamics, magnetohydrodynamics, the physics of combustion and explosion, etc. underlying it.[24]

2. The new eruptive cosmogony explains the properties of Titan and the known differences between the Galilean satellites. It singles out Callisto with its unexploded icy crust as a potentially dangerous object for the Earth. It follows from the foregoing that in the case of an explosion, the bombardment will be the strongest since the Imbrian event 3.9 billion years ago. The Earth's atmosphere will be poisoned, and its surface will be literally plowed to a depth of hundreds of meters. The known terrestrial resources are insufficient for preventing such bombardment. The most massive extinction of life forms during the entire history of the Earth will take place. The fate of *Homo sapiens* will not be an exception.

3. The principal far-reaching practical conclusion that can be drawn from the new eruptive cosmogony is the need to set up a purposeful second-generation space program for studying solar system bodies. Its main purpose should be to ascertain the degree of real danger from Callisto. The program should be divided into two stages: a) testing the validity of the new eruptive cosmogony by investigating SP nuclei *in situ* to detect molecular O_2 dissolved in their ices and consequent combustion in jets and sources of escaping matter; b) *in situ* investigations of Callisto's ices for the purpose of determining their degree of saturation with electrolysis products.

On the threshold of the 3rd millennium, Russia, with all of its resources and developed space technologies, could, remaining true to its centuries-old traditions, declare the protection and rescue of the Earth's civilization from possible danger as its national ideology and enter the international arena with appropriate initiatives. The program proposed above, which pursues a clearly defined large-scale goal that touches the interest of everyone, might unite the efforts of all the space-exploring powers both in evaluating the potential threat to mankind and in achieving an understanding of the actual ways in which solar system bodies form and evolve, which would facilitate exploiting them in the future. As a first simple step it would be useful to take into account the consequences of the new eruptive cosmogony and its recommendations in working out the details of the already planned missions to small bodies and satellites of planets.

[1] E. M. Drobyshevski, Nature **250**, 35 (1974).
[2] E. M. Drobyshevski, Astron. Astrophys. Trans. **10**, 211 (1996).
[3] E. M. Drobyshevski, Moon Planets **18**, 145 (1978).
[4] J. C. Decroli, H. Granicher, and C. Jaccard, Helv. Phys. Acta **30**, 465 (1957).
[5] E. M. Drobyshevski, V. A. Chesnakov, and V. V. Sinitsyn, Adv. Space Res. **16**, 73 (1995).
[6] E. M. Drobyshevski, Moon Planets **23**, 339 (1980).
[7] E. M. Drobyshevski, Astron. Astrophys. Trans. **12**, 327 (1997).
[8] E. M. Drobyshevski, Moon Planets **23**, 483 (1980).
[9] E. M. Drobyshevski, Earth, Moon, Planets **44**, 7 (1989).
[10] É. M. Drobyshevskiĭ, *The ''Callisto'' Project, Express Information No. 2-89* [in Russian], Izd. Akad. Nauk SSSR (1989), p. 24.
[11] V. P. Konopleva, *Cometary Exploration*, edited by T. I. Gambosi, Budapest (1982), Vol. 1, pp. 13–19.
[12] E. M. Drobyshevski, ''The history of Titan, of Saturn's rings and magnetic field, and the nature of short-period comets,'' Preprint PhTI-674, A. F. Ioffe Physicotechnical Institute, Leningrad (1980).

[13] E. M. Drobyshevski, Moon Planets **24**, 13 (1981).
[14] E. M. Drobyshevski, in *Evolution and Source Regions of Asteroids and Comets (IAU Colloquium 173)*, edited by E. M. Pittich, H. Rickman, and J. Svoreñ (1998).
[15] E. M. Drobyshevski, Earth, Moon, Planets **43**, 87 (1988).
[16] E. M. Drobyshevski, Astron. Astrophys. Trans. **13**, 215 (1997).
[17] M. F. A'Hearn, P. D. Feldman, and D. G. Schleicher, Astrophys. J. **274**, L99 (1983).
[18] M. J. Mumma, W. E. Blass, H. A. Weaver, and H. P. Larson, Bull. Am. Astron. Soc. **20**, 286 (1988).
[19] H. Rickman, Adv. Space Res. **9**(3), 59 (1989).
[20] I. I. Agafonova and E. M. Drobyshevski, Earth, Moon, Planets **33**, 1,111 (1985).
[21] E. M. Drobyshevski, Earth, Moon, Planets **40**, 1 (1988).
[22] D. Olsson-Steel, Mon. Not. R. Astron. Soc. **227**, 501 (1987).
[23] F. A. Tsitsin, V. M. Chepurova, and I. L. Genkin, Astron. Vestn. **27**(6), 55 (1993).
[24] É. M. Drobyshevskiĭ, Astron. Vestn. **29**, 572 (1995).

Translated by P. Shelnitz

Study of the equilibrium and motions of the medium in a rotating star with a magnetic field using tools from angular momentum theory

Yu. V. Vandakurov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 15–18 (September 1999)

The complicated processes occurring in a rotating magnetized medium are examined by representing all the vector quantities in the form of expansions in a complete system of orthogonal vector spherical harmonics. A separation of variables is ultimately achieved without a loss of accuracy despite the presence of nonlinear forces (the magnetic and Coriolis forces). The distribution of the rates of rotation and circulation motions in an adiabatically stratified, slowly rotating star or an atmospheric convection zone is studied as an example. The postulate of minimum entropy production from nonequilibrium thermodynamics is employed to find the most probable steady-state configuration. One solution satisfactorily describes the differential rotation observed on the Sun. The preliminary data support the notion that the superfast rotation of the type observed in Venus' atmosphere can also be explained within the theory discussed. © *1999 American Institute of Physics.* [S1063-7842(99)00409-2]

The current helioseismological data[1,2] provide evidence that the entire solar convection zone rotates differentially, with faster rotation of the medium in the equatorial region. It is usually assumed that such nonrigid-body rotation is shaped under the influence of a turbulent viscous force generated by convective motions of the medium. Rough estimates show that the turbulent viscosity should be approximately 14 orders of magnitude greater than the molecular viscosity of the medium (similar estimates made for conditions of convective heat transfer lead to appreciably smaller values of the turbulent viscosity[3]).

However, when there is such a strong increase in viscosity, entropy production (or, stated differently, energy dissipation) increases sharply, but, according to the fundamental postulate of nonequilibrium thermodynamics, entropy production should be at a minimum.[4] For this reason we assume that the distribution of the rotation rate in stellar or atmospheric convection zones should be established in accordance with that postulate. It is shown below that such an approach provides solutions which satisfactorily describe the observational data.

Since stellar and atmospheric convection zones are usually characterized by a small value of the superadiabaticity, we shall study an adiabatically stratified convection zone and find the motion structure for which the mean solenoidal viscous force, which is strictly related to the solenoidal Coriolis force, has its minimum value. In addition, the convective motions themselves will be assumed to be insignificant, although, of course, an instability which initiates a spontaneous transition from rigid-body to differential rotation can appear only in the presence of such motions. In addition, we shall assume that the rotation of the convection zone under consideration is slow.

In the case of a stationary, adiabatically stratified medium, the pressure is a function of density, and the equation of motion

$$\frac{1}{2}\nabla(\mathbf{v}\cdot\mathbf{v})+\mathbf{W}+\frac{1}{\rho}\nabla p+\nabla\Phi=\mathbf{F} \tag{1}$$

yields the relation

$$\text{curl } \mathbf{W}\approx\mathbf{D}, \tag{2}$$

where

$$\mathbf{W}=(\text{curl } \mathbf{v})\times\mathbf{v}, \tag{3}$$

$\mathbf{D}=\text{curl } \mathbf{F}$, $\mathbf{v}$ is the hydrodynamic velocity, $\mathbf{F}$ is the viscous force, p is the pressure, ρ is the density, and Φ is the gravitational potential.

Equation (2) characterizes the aforementioned strict relation between the solenoidal components of the forces. It is clear that as the mean value of the functional $|\text{curl}\mathbf{W}|$ over the convection zone decreases, the mean viscous force falls off and the entropy production consequently becomes lower. Our problem is thus to find a velocity distribution under which this mean force will have the smallest value (since we are dealing with very small velocity changes, the condition of an assigned mean angular momentum is satisfied automatically).

We seek a solution of the stated problem, assuming that all the motions of the medium are axially symmetric and symmetric relative to the equatorial plane, representing all the vector quantities in the form of series in an orthonormal system of the vector spherical harmonics $\mathbf{Y}_{J0}^{(\lambda)}$ with $\lambda=-1$, 0, or 1, which depend only on the angular variables, and using the same notation as in Ref. 5. For example,

$$\mathbf{v}=\sum_J v_{J0}^{(\lambda)}\mathbf{Y}_{J0}^{(\lambda)}=\sum_J\{\mathbf{i}_r v_{J0}^{(-1)}Y_{J0}+I^{-1}\times[\mathbf{i}_\vartheta v_{J0}^{(+1)}-\mathbf{i}_\varphi i v_{J0}^{(0)}](\partial Y_{J0}/\partial\vartheta)\}, \tag{4}$$

where $v_{J0}^{(\lambda)}=v_{J0}^{(\lambda)}(r)$; $I=[J(J+1)]^{1/2}$; $\mathbf{i}_r$, $\mathbf{i}_\vartheta$, and $\mathbf{i}_\varphi$ are the unit vectors of the spherical coordinate system; and Y_{J0} is a spherical harmonic.

It is important that the method under discussion permits a separation of variables within the exact equations, despite the presence of the nonlinear Coriolis force. The question of bringing the cumbersome expressions derived into a simpler form was discussed in our previous paper.[6] The series for the vector **W**, which is defined by formula (3), was also written out in Ref. 7.

In the case under consideration the problem reduces to finding the minimum of the functional $|\text{curl }\mathbf{W}|$, which is quadratic with respect to $v_{J0}^{(\lambda)}$, and the radial derivatives of these coefficients. Let us examine the simplest solution, which corresponds to an identical radial dependence of all the coefficients mentioned, i.e.,

$$v_{J0}^{(\lambda)}=\text{const}\cdot r^{\alpha}, \tag{5}$$

where α is a constant and real quantity.

This approximation does not impose other restrictions on the equilibrium configuration if the poloidal component of the velocity is insignificant. Otherwise, we must take into account the relation following from the continuity equation

$$\text{div }\rho\mathbf{v}=0. \tag{6}$$

In our approximation of slow rotation the density can be replaced approximately by its spherically symmetric component, and we then find that the density is approximated as

$$\rho=\text{const}\cdot r^{-\beta}, \tag{7}$$

where β is a constant.

As a result, we have

$$v_{J0}^{(+1)}=[(\alpha-\beta+2)/I]v_{J0}^{(-1)}. \tag{8}$$

With consideration of these relations, for the vector curl **W**, which we shall denote by **R**, we obtain

$$R_{J0}^{(-1)}=\frac{I}{2\alpha}R_{J0}^{(+1)}=\frac{i}{r^2}I\sum_{J_1J_2}G_{J_1J_2}^{J}\{(\alpha+1)I_2Z_{JJ_1}^{J_2}-(\alpha-\beta+2)I_1^2I_2^{-1}Z_{JJ_2}^{J_1}\}v_{J_10}^{(0)}v_{J_20}^{(-1)}, \tag{9}$$

$$R_{J0}^{(0)}=-\frac{i}{r^2}\sum_{J_1J_2}G_{J_1J_2}^{J}\{[(\alpha+1)I^2Z_{J_1J_2}^{J}+2\alpha I_1^2Z_{JJ_2}^{J_1}]v_{J_10}^{(0)}v_{J_20}^{(0)}+[I_1-(\alpha+1)\times(\alpha-\beta+2)I_1^{-1}][(\alpha-\beta+2)I^2I_2^{-1}Z_{J_1J_2}^{J}+2\alpha I_2Z_{JJ_1}^{J_2}]v_{J_10}^{(-1)}v_{J_20}^{(-1)}\}, \tag{10}$$

where

$$G_{J_1J_2}^{J}=\frac{(C_{J_10J_20}^{J0})^2}{4II_1I_2}\left[\frac{(2J_1+1)(2J_2+1)}{\pi(2J+1)}\right]^{1/2},$$

$Z_{ab}=a(a+1)+b(b+1)-c(c+1)$, $I_k=[J_k(J_k+1)]^{1/2}$, C_{ab}^{c} is a Clebsch–Gordan coefficient, and $k=1$ or 2.

In Eqs. (9) and (10) $J>0$, and the sum of indices $J+J_1+J_2$ is even (otherwise, the result would be zero). In addition, under the condition of symmetry relative to the equatorial plane the coefficients $v_{J0}^{(0)}$ have odd values of J, and in the other velocity coefficients J is even.

In the case of an identical radial dependence for all the modes under consideration, to find the mean value of the scalar product $\mathbf{R}\cdot\mathbf{R}^*$ it is sufficient to carry out averaging over the entire spherical surface. Here the asterisk denotes the complex conjugate. Making use of the orthogonality of the vector spherical harmonics, we obtain

$$[r/v_{10}^{(0)}]^4\sum_{\lambda J}R_{J0}^{(\lambda)}R_{J0}^{(\lambda)*}=\sum_{n}\varepsilon_{2n}\varepsilon_{2n}^{*}, \tag{11}$$

where ε_{2n} is a dimensionless variable and $n=1,2,\ldots$ is the number of the equation, i.e., $2n$ is equal to J if J is even and to $J+1$ if J is odd.

The explicit expression for ε_{2n} can be obtained after plugging in (9) and (10). Of course, Eq. (11) should be equated to the corresponding viscous force reduced to dimensionless form, but our problem is confined to finding the velocity distributions for which the root-mean-square value of all the ε_{2n}, i.e., $\overline{\varepsilon}$, has a minimum value. Obviously $N|\overline{\varepsilon}|^2$ is equal to the right-hand side of Eq. (11) if only N terms are kept in it.

Since $G_{11}^{2}=1/[4(30\pi)^{1/2}]$, it can be seen that in the case of rigid-body rotation, in which $\alpha=1$ and only a single coefficient $v_{10}^{(0)}$ out of the entire series of velocity coefficients is nonzero, all the ε_{2n} are equal to zero. In this case $v_{10}^{(0)}=-ir\Omega(8\pi/3)^{1/2}$, where Ω is the angular velocity. Such a state is probably unstable in the presence of even slow convective motions. Our calculations provide evidence supporting the existence of such instability, although this question is still in need of more detailed investigation.

We assume that $v_{J0}^{(0)}/v_{10}^{(0)}=u_J=u_{2k-1}$ and $v_{J0}^{(-1)}/|v_{10}^{(0)}|=f_J=f_{2k}$, where k is equal to $1,2,\ldots,N$. The coefficients u_J and f_J will then be real, and $u_1=1$. In Eq. (11) we clearly have $1\leqslant n\leqslant N$.

To solve the stated problem we compared various models, varying either one of the coefficients u_f, f_J or the parameter α and looking for the smallest value of $\overline{\varepsilon}$. Models with a number of rotational modes up to 30 were considered. In the first stage we studied models with motion close to rigid-body rotation in the approximation of negligibly small poloidal velocities, under which all $f_J\approx 0$ and in Eq. (11) $2n=J$. Solutions were found for values of α in the vicinity of 1.28 and 1.34. The first of these values describes rotation similar to that observed on the Sun, and we shall examine it in greater detail.

Table I presents the basic parameters of three modes calculated with different values of N, where N is equal to the number of rotational modes left (all u_J with indices $J>2N-1$ are assumed to be equal to zero). We present the parameter α, the two basic coefficients u_3 and u_5 (we recall that $u_1=1$), the last coefficient $u_{(2N-1)}$, the index $J=J_{\max\varepsilon}$ for which $|\varepsilon_J|=\varepsilon_2$ and the highest value $|\varepsilon_{J_{\max\varepsilon}}|$ is achieved, the corresponding maximum value, the minimum root-mean-square value found $\overline{\varepsilon}$, and the ratio of the equatorial angular

TABLE I. Parameters of the rotation models.

Model	Model No.	α	$u_3 \cdot 100$	$u_5 \cdot 100$	$u_{(2N-1)}$	$J_{\max \varepsilon}$	$\lvert\varepsilon_{J\max}\rvert$	$\overline{\varepsilon}$	r_Ω
1	11	1.28021	−4.9639	−1.1622	-6.8×10^{-5}	22	3.4×10^{-4}	1.9×10^{-4}	1.92
2	20	1.28014	−4.9630	−1.1621	-5.3×10^{-6}	32	6.1×10^{-5}	3.3×10^{-5}	2.24
3	29	1.28019	−4.9636	−1.1621	-3.2×10^{-7}	30	4.3×10^{-5}	2.1×10^{-5}	2.27

velocity to the angular velocity at the poles r_Ω. The variation of the ratio of neighboring coefficients in the series of u_J with increasing J can be undulating, as follows from Fig. 1. The latitude dependence of the angular velocity is shown in Fig. 2, in which the crosses denote experimental data obtained by measuring the Doppler shifts of lines.[8] The satisfactory fit between the theoretical and experimental data is evidence that the state with the minimum mean solenoidal viscous force under consideration is, in fact, realized on the Sun.

The heightened sensitivity of the value of the angular velocity in the immediate vicinity of the rotation axis toward the number N is noteworthy. Only if N is greater than or of the order of 20 does the dependence of the distribution of the angular velocity and the other parameters on N vanish. It should also be stressed that in the case of infinitely large N the requirement for a fairly rapid decrease in the coefficients u_J will probably not be satisfied. The latter is necessary because in the near-axial zone the derivative $(\partial Y_{J0}/\partial\vartheta)$ in Eq. (4) is proportional to $\vartheta J^{3/2}$, i.e., increases with increasing J. In other words, under real conditions the amplitude of very high-order modes will probably be smaller than the value determined by the minimization procedure. Even in that case the decrease in the mean solenoidal viscous force in comparison to the value corresponding to the unminimized model can be quite substantial, since $\overline{\varepsilon}$, as follows from Table I, falls off with increasing N.

In the case of convective planetary atmospheres the density usually drops sharply with altitude [in Eq. (7) $\beta \gg 1$], and models with intense meridional flows of the medium are of interest. A preliminary study of the model with $\beta=800$ provides evidence supporting the existence of a minimum for $\overline{\varepsilon}$ under the condition that α is close to or somewhat smaller than β. If it is also assumed that the thickness of the convection zone is 4–5 times greater than the ratio r/β, the angular velocity of the upper layers of the zone will be two orders of magnitude greater than that of the lower layers. Such a model could describe the superfast rotation of the upper layers of the Venusian atmosphere. In this case convection is

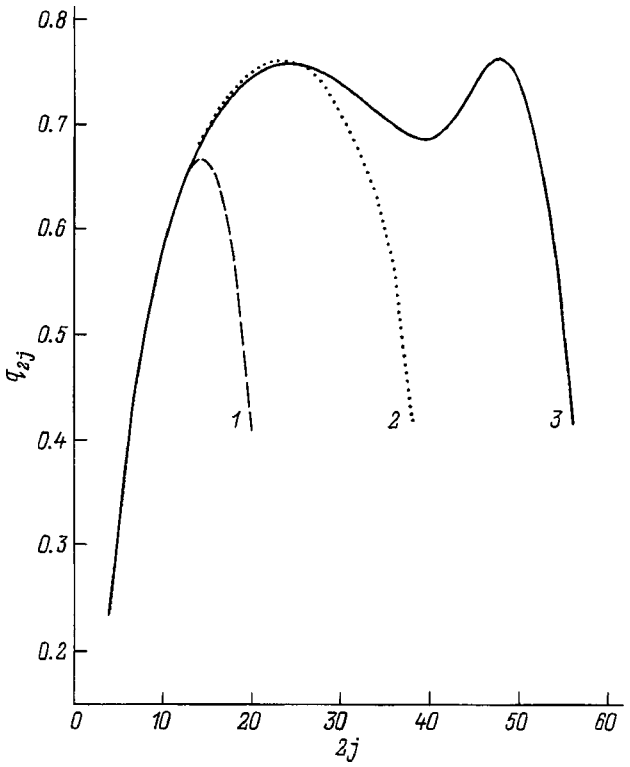

FIG. 1. Ratio between the amplitudes of neighboring coefficients $q_{2j}=u_{2j+1}/u_{2j-1}$ as a function of the mean value of the lower index of the two coefficients. Here $j=1,2,3,\ldots,(N-1)$. The first ratio q_2 is negative and is not shown in the figure. The numbers of the curves correspond to the numbers of the models in Table I.

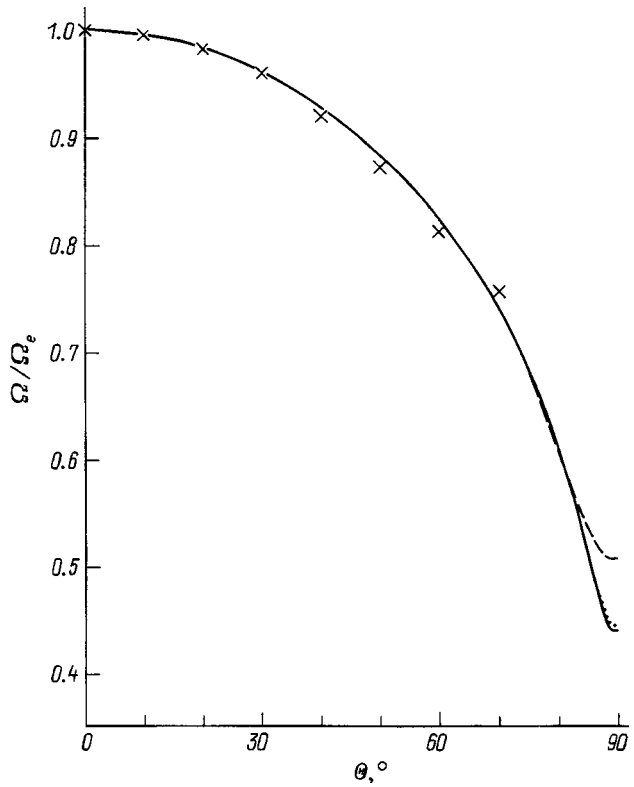

FIG. 2. Dependence of the angular velocity Ω normalized to the angular velocity at the equator Ω_e on the latitude Θ in the case of models 1–3 in Table 1 (the dashed, dotted, and solid curves, respectively). The crosses are experimental results reported by Howard.[8]

excited by the large horizontal temperature gradients.[9,10] This problem merits more detailed study.

If we are dealing with a magnetized nonrotating convection zone, then the magnetic force appears in Eq. (2) instead of **W**. It is not difficult to show that this equation will be solvable only in the case of a dipolar magnetic field.[6] The higher the multipolarity of the field, the greater will be the difference between the number of equations and the number of unknown variables. It has not been ruled out that the solution should be sought under the hypothesis that the characteristic diameter of the magnetic tube is small, comparable to the size of the chaotic fluctuations.[6] Possibly a solution to the enigma of the existence of small-scale magnetic tubes with large field fluxes will be found in this way.

Part of this work was carried out with financial support from the research fund of the St. Petersburg Scientific Training Center of the Ministry of Education of the Russian Federation and the Russian Academy of Sciences Program ''Electrophysics of High-Density Currents and High Magnetic Fields.''

[1] T. M. Brown, J. Christensen-Dalsgaard, W. A. Dziembowski *et al.*, Astrophys. J. **343**, 526 (1989).
[2] S. Tomczyk, J. Schou, and M. J. Thompson, Astrophys. J. **448**, L57 (1995).
[3] Yu. V. Vandakurov, *Convection on the Sun and the 11-Year Cycle* [in Russian], Nauka, Leningrad (1976).
[4] I. Gyarmati, *Non-Equilibrium Thermodynamics, Field Theory and Variational Principles* [Springer-Verlag, Berlin–Heidelberg–NewYork (1970); Mir, Moscow (1974)].
[5] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum: Irreducible Tensors, Spherical Harmonics, Vector Coupling Coefficients, 3nj Symbols* [World Scientific, Singapore–Teaneck, NJ (1988); Nauka, Leningrad (1975)].
[6] Yu. V. Vandakurov, Astron. Zh. **76**, 29 (1999) [Astron. Rep. **43**, 24 (1999)].
[7] Yu. V. Vandakurov, Pis'ma Astron. Zh. **24**, 450 (1998) [Astron. Lett. **24**, 384 (1998)].
[8] R. Howard, Annu. Rev. Astron. Astrophys. **22**, 131 (1984).
[9] A. D. Kuz'min and M. Ya. Marov, *Physics of the Planet Venus* [in Russian], Nauka, Moscow (1974).
[10] N. I. Burangulov, S. S. Zilitinkevich, and V. V. Kerzhanovich *et al.*, in *Dynamics of Venus' Atmosphere* [in Russian], Nauka, Leningrad (1974).

Translated by P. Shelnitz

Atomic effects accompanying the beta decay of tritium

Yu. A. Akulov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 19–21 (September 1999)

The results of the first experiment designed to measure the difference between the beta-decay constants of atomic and molecular tritium $\Delta\lambda=\lambda_a-\lambda_m$ are presented. The experimental scheme calls for the creation of two identical samples of a gas mixture containing helium-4 and molecular tritium followed by the treatment of one of them for the purpose of bringing the tritium into the atomic state. The value of $\Delta\lambda$ is determined by comparing the growth rates of the ratio of the concentration of radiogenic helium-3 to the concentration of helium-4 in the samples with molecular and atomic tritium. The value $\Delta\lambda=(4.6\pm0.8)\times10^{-12}$ s^{-1}, which corresponds to a relative change in the decay constant amounting to ~0.26%, is obtained.
 [S1063-7842(99)00509-7]

The subject of this report is the interaction of a beta electron formed by an atomic nucleus with electrons and electron vacancies in the atomic–molecular system surrounding the beta-active nucleus. According to the Fermi theory of beta decay, the probability of the emission of electrons by nuclei and their spectrum depend on the structure of the phase space accessible to a virtual beta electron in the atomic–molecular system. In this case the experimentally determined values of the fundamental parameters of beta conversion, viz., the decay constant λ and the cutoff energy $E_{\max}$ of the beta spectrum, are sensitive to the state of the electronic environment of the nucleus (as is manifested in the so-called chemical shifts of λ and $E_{\max}$). The examination of tritium-containing atomic–molecular systems is most promising from the standpoint of solving experimental and theoretical problems in the area of the investigation of the chemical shifts accompanying beta decay. Since the mass excess of the second neutron in the tritium nucleus is close to the Coulomb energy of the second proton in the helium nucleus, the energy released upon the beta decay of tritium is relatively small ($^3\mathrm{H}\rightarrow{}^3\mathrm{He}+e^-+\tilde{\nu}+18.6$ keV). In this case an appreciable portion of the virtual beta electrons are generated with wavelengths characteristic of atomic electronic systems. This leads to significant changes in the real beta spectrum and makes the chemical shifts of λ and $E_{\max}$ measurable. When the cutoff energy is measured, the shifts of $E_{\max}$ are masked by the unavoidable variation of the energy of the real beta electron as it moves toward the detector, which fundamentally complicates the identification of observable changes in $E_{\max}$. Such difficulties do not arise in measurements of the chemical shifts of the decay constant, since the decay rate is determined only by the density of vacancies for the virtual beta electron and clearly does not depend on the fate of the real electron leaving the circumnuclear space. This paper presents the results of the first experiment designed to measure the difference between the decay constants for atomic (λ_a) and molecular (λ_m) tritium $\Delta\lambda=\lambda_a-\lambda_m$.

Helium differential isotope mass spectrometry was used to determine $\Delta\lambda$. The method is based on a comparison of the growth rates of the ratio of the content of radiogenic helium-3 to the content of helium-4 in two samples, each of which contains helium-4 and some chemical compound of tritium. The important advantages of such an approach include the complete elimination of absolute measurements and significant simplification of the way in which the helium losses are taken into account, and the main sources of systematic error are thereby eliminated. In the case under consideration the experimental scheme called for the creation of two identical samples of a gas mixture containing helium-4 and molecular tritium at a time t_0, followed by the treatment of one of them for the purpose of bringing the tritium into the atomic state. The instant the tritium passed through a palladium membrane having an active area of 12 cm^2, a thickness of 0.01 cm, and a temperature of ~530 °C was taken as the starting point for measuring the time. The tritium which passed through the membrane was mixed with helium-4, and then the mixture was distributed among molybdenum glass ampuls, which had an inner diameter of 8 mm and a length of 350 mm. The partial pressure of tritium in the ampuls was 93 Pa, and the partial pressure of helium was 4–6 Pa. The $^3H_2$ molecules were dissociated by the energy from an rf discharge ignited in the ampul by a sinusoidal field with a frequency of 1 MHz and an intensity of ~100 V/cm, which was formed by external annular electrodes. Thermalized free tritium atoms were obtained by a resonant dissociation mechanism, under which the $^3H_2$ molecules receive the energy needed to break the interatomic bond as a result of superelastic collisions with mercury atoms excited in the discharge, which were preliminarily introduced into the ampul and vaporized by the discharge itself. The effective value of the current in the discharge was held at a level of ~40 mA, and the power dissipated in the discharge was 140–145 W.

Mass-spectrometric and optical methods were developed to determine the degree of disassociation of the molecules. In the mass-spectrometric measurements free hydrogen atoms passed from the discharge cell into the ionization chamber of

the mass spectrometer along a Teflon antirecombination channel in a molecular-flow regime. The value of the degree of dissociation (ε) was determined from the dependence of the current of atomic hydrogen ions I on the power P absorbed in the discharge cell. A horizontal plateau corresponding to a degree of dissociation of hydrogen close to 100% was achieved on the plot of $I(P)$ for the values of the current and field intensity in the discharge indicated above.[1] The optical method for determining the value of ε was based on a comparison of the intensity of the H_α, H_β, H_γ, and H_δ lines of the Balmer series, which were identified using a monochromator, and the bands with wavelengths of ~652, ~580, and ~410 nm from the spectrum of molecular hydrogen for two regimes of the rf discharge: a flow-through regime, under which gas is continuously pumped through the discharge tube, and a static regime. In the flow-through regime the time spent by a molecule in the space between the electrodes is only 0.003–0.01 s. Over this time practically no dissociation of the hydrogen occurs, and the spectrum of the discharge is dominated by molecular bands. In the static regime (i.e., without gas flow), at the same values of the pressure and the power dissipated in the discharge the intensity of the lines of the Balmer series increases by a factor of 10–15, while the intensity of the molecular bands decreases by a factor of 25 or more. A comparison of the coefficients of relative variation of the intensity reveals that the degree of dissociation of hydrogen in the static regime exceeds 93%. The statistical weights of the ionized and excited states of hydrogen in the discharge, according to calculations based on estimates of the current, electron density, and electron energy, does not exceed 10^{-3}%. Thus, the experimental conditions ensured the observation of the beta decay of tritium in the ground atomic and ground molecular states.

The equation of measurement for the helium differential method of determining $\Delta\lambda$ for the ^3H atom–$^3H_2$ molecule pair has the form[2]

$$\frac{(^3He/^4He)_a}{(^3He/^4He)_m}=\frac{\tau_a}{\tau_m}\left(k\varepsilon\frac{\Delta\lambda}{\lambda_m}+1\right).$$

Here $(^3He/^4He)_{m,a}$ are the mass-spectrometrically measured values of the helium isotope ratios in the samples with molecular and atomic tritium, and τ_m and τ_a are the effective exposure times of the samples, and the time that a portion of the tritium is in the atomic state (i.e., the existence time of the discharge) is expressed as $\tau_a\cdot k$, where $k<1$ is a dimensionless coefficient. The lengths of the time intervals τ_m and τ_a were calculated from the relations $\tau_{m,a}=(t_{m,a}-t_0)+\delta_{m,a}\cdot\eta_{m,a}$, where t_m and t_a are the time readings at the instants the tritium is removed from the samples, η_m and η_a are the time intervals between t_m and t_a, respectively, and the time of measurement of the helium isotope ratios in the samples, and δ_m and δ_a are the relative residual specific activities of tritium, which determine the rate of accumulation of helium-3 in the ''molecular'' and ''atomic'' samples during the time intervals η_m and η_a. The removal of tritium from the samples was accomplished by interaction of the gas mixture with a palladium membrane similar to the one used in filling the ampuls. The uncertainty in determining the starting point t_0 for the time measurements and in the values of t_m and t_a did not exceed 8 s. The relative residual activity due to the incomplete removal of tritium through the membrane was measured by an internal-flow proportional counter with a sensitivity equal to $\sim10^6$ tritium atoms[3] and amounted to $\delta_m=0.00115$ and $\delta_a=0.00095$ of the original specific activity of the mixture. The permeability of the palladium membranes toward helium, the absorption of helium by the ampul walls, the presence of an admixture of ^3He in the inert ^4He, and other factors which produce nonradiogenic changes in the $^3He/^4He$ ratio in the mixtures were monitored using a mass spectrometer having a sensitivity of $\sim3\times10^4$ atoms with respect to ^3He (Ref. 4). The amount of helium-3 formed in the samples during their exposure exceeded 2×10^{13} atoms. Thus, the possible systematic errors caused by helium losses were recorded at the $\sim10^{-6}$ % level.

The helium isotope ratios in the sample, which were close to 0.001 in absolute value, were measured using a modified MI1201 double-beam isotope mass spectrometer,[5] which provided a reproducibility in relative measurements of $^3He/^4He$ at the 0.02% level in a series of independent control experiments with a special helium–tritium calibration mixture. In the actual experiment for determining $\Delta\lambda$, five series of measurements, in each of which samples of the helium mixture from the ''molecular'' and ''atomic'' samples were alternately admitted into the chamber, were performed. As a result, a value of 0.90257 with a standard deviation of 0.00021 was obtained for $(^3He/^4He)_a/(^3He/^4He)_m$ from the left-hand side of the equation of measurement. The effective exposure time intervals had the values $\tau_m=271.80$ min and $\tau_a=244.90$ min, and the dimensionless coefficient was $k=0.70$, i.e., the discharge was sustained for 171 min. The value taken for the degree of dissociation ε was $0.95^{+0.05}_{-0.02}$.

The ratio $\Delta\lambda/\lambda$ determined from the equation of measurement for the values indicated above of the parameters appearing in the equation has a value of 0.00257 ± 0.00045. The error of the result corresponding to one standard deviation was determined by a quadratic form including the variances of all the parameters of the equation of measurement. The variance of the ratio τ_a/τ_m was calculated with allowance for the 95% positive correlation of the random quantities t_m and t_a due to the identical nature of the procedures for removing tritium from the two samples. When the absolute value of the difference between the half-lives of molecular and atomic tritium $\Delta T_{1/2}$ is calculated on the basis of the measured relative shift of the decay constant, it is sensible to use the value $T_{1/2m}=(12.296\pm0.017)$ years, i.e., the weighted-mean value of the two most recently published, closely agreeing estimates of the half-life of molecular tritium, one of which was obtained by helium isotope mass spectrometry,[6] while the other was obtained from the decay curve plotted by recording the bremsstrahlung.[7] Taking into account that $\Delta T_{1/2}/T_{1/2}=-\Delta\lambda/\lambda$ when $\Delta\lambda/\lambda\ll1$, we find $T_{1/2m}-T_{1/2a}=(0.0316\pm0.0055)$ years $=(11.5\pm2.0)$ days. For the absolute value of the difference between the decay constants λ_a and λ_m we obtain

$$\Delta\lambda = \lambda_a - \lambda_m = \frac{\ln 2}{T_{1/2m}}(0.00257 \pm 0.00045)$$
$$= (4.6 \pm 0.8) \times 10^{-12}\ \ \mathrm{s}^{-1}.$$

Techniques for the rapid tritiation of various hydrogen-containing compounds are being developed on the basis of the procedure devised for the production, stabilization, and diagnosis of atomic hydrogen, which will permit measurement of the chemical shirts of the decay constant for a long list of atomic–molecular systems. The determination of the lifetime of the tritium nucleus for several known states of its electronic environment can be interpreted as the calibration of a natural electronic probe, which has nuclear locality but generates electrons with an atomic wavelength. By subsequently measuring the chemical shifts $\Delta\lambda$ for an unknown structure of a phase volume accessible to a beta electron, we can draw conclusions regarding the presence of specific electronic states in the atomic phase space. Additional information on this can be provided by an analysis of the spectra of the relaxation radiation due to reorganization of the electronic levels in response to a change by one unit in the charge of one of the nuclei in the atomic–molecular system under investigation. One special feature and advantage of the method for structurally analyzing atomic–molecular systems using beta electrons from their own nuclei is the ideal correlation between the coordinates of the object being probed (the electron cloud) and the radiation source (the nucleus). This makes it possible to use quantum beta-electron diagnostics to investigate experimentally not only the energy structure of an electron ensemble but also its spatial configuration.

This work was performed with financial support from the Russian Fund for Fundamental Research (Project No. 97-03-33665) and the State Scientific-Technical Program ''Fundamental Metrology'' (Project No. 4.06).

[1] Yu. A. Akulov, B. A. Mamyrin, and P. M. Shikhaliev, Zh. Tekh. Fiz. **67**(5), 140 (1997) [Tech. Phys. **42**, 584 (1997)].
[2] Yu. A. Akulov, B. A. Mamyrin, and P. M. Shikhaliev, Pis'ma Zh. Tekh. Fiz. **19**(18), 72 (1993) [Tech. Phys. Lett. **19**(9), 559 (1993)].
[3] B. A. Mamyrin, Yu. A. Akulov, L. V. Khabarin, and V. S. Yudenich, Prib. Tekh. Éksp. No. **2**, 151 (1982).
[4] Yu. A. Akulov, B. A. Mamyrin, L. V. Khabarin, and V. S. Yudenich, Prib. Tekh. Éksp. No. **2**, 173 (1985).
[5] Yu. A. Akulov, B. A. Mamyrin, and P. M. Shikhaliev, Prib. Tekh. Éksp. No. **3**, 114 (1995).
[6] Yu. A. Akulov, B. A. Mamyrin, and L. V. Khabarin *et al.*, Pis'ma Zh. Tekh. Fiz. **14**, 940 (1988) [Sov. Tech. Phys. Lett. **14**, 416 (1988)].
[7] B. Budick, J. Chen, and H. Lin, Phys. Rev. Lett. **67**, 2630 (1991).

Translated by P. Shelnitz

Physics of atomic collisions in retrospect

J. Macek

University of Tennessee, Knoxville, Tennessee 379960-1501; Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 22–26 (September 1999)

The classic work by Mott and Massey, in which the scope of the physics of atomic collisions was defined, was published about 65 years ago. Since then, this field has undergone considerable development. In fact, all the theoretical methods named by Mott and Massey have been implemented to some extent. As for experiment, the measurements performed, which are differential with respect to several parameters, have provided for reliable testing of the mechanisms proposed. The physics of atomic collisions has been developed to the point that we can look back on the road taken and discover many achievements which have promoted its development. Progress in science is usually associated with highly concentrated efforts on the part of a critical number of investigators to solve a specific problem, which is widely regarded as being of great importance. Such a ''breakthrough'' is usually followed by rapid development of the field. In this respect, the physics of atomic collisions is no exception. It has known periods of highly concentrated efforts aimed at solving specific problems and breakthroughs followed by rapid development and subsequent periods of stagnation. The cycles have repeated: a new area for concentrated efforts is discovered, a breakthrough occurs, and a new concept is established. Some of these cycles are analyzed from the standpoint of their significance to atomic physics as a whole.
[S1063-7842(99)00609-1]

The status of the physics of atomic collisions as an independent branch of atomic physics was established in the late nineteen-fifties and the early nineteen-sixties, when the first international conferences on this subject were held. Two discoveries were made in that period, which, as is now clear, had a decisive influence on the development of the field. One of them is the prediction and discovery of resonances in the cross sections for the excitation of atoms and molecules by electron impact.[1] It was simultaneously discovered that such resonances, or autionizing states, are excited not only upon electron impact, but also upon photoabsorption and the impact of charged atomic particles. The investigation of these states became a widely known achievement, to which scientists at the A. F. Ioffe Physicotechnical Institute made a significant contribution.[2] Some of the phenomena discovered will be considered in greater detail. The second discovery, which was made in the Physicotechnical Institute in 1964 and which, at first glance, is not associated with resonances, was even more remarkable. In order to understand it, we should recall what an ''adiabatic'' state of a diatomic molecule is. In such a state the electrons are found in bound levels in a field of fixed Coulomb centers. It was expected that transitions between ''adiabatic'' states are unlikely during ''slow'' collisions (where the velocity of the projectile is much smaller than the mean orbital velocity of a bound electron). Thus, collisions in this energy range should be basically elastic. It was shown in a series of notable studies carried out at the Physicotechnical Institute[3] that in reality the opposite situation occurs: intense lines were discovered in the energy-loss spectrum at energies of the order of hundreds of electron volts. This discovery became possible owing to the use of the coincidence technique.[4] An analysis of the coincidence spectra showed that characteristic lines are excited when definite distances of closest approach of the colliding particles are achieved. In fact, when the electronic shells of atoms overlap, electrons are selectively ''promoted'' from inner to valence shells. The vacancies formed in the inner shells are usually filled by Auger transitions, which are accompanied by the escape of fast electrons with a fixed energy. To understand this process, it was necessary to abandon the existing theories regarding adiabatic states and introduce ''diabatic'' one-electron orbitals, such as those in the H_2^+ system.[5] This opened up the possibility of performing nonempirical calculations of inelastic processes. The existence of a quantitative theoretical foundation and experimental techniques that were continually being upgraded allowed the next 20 years to be devoted to testing the molecular-orbital model in ion–atom and atom–atom collisions.

The discovery of Auger electrons emitted when vacancies in the inner shells of atoms are filled[6] provided convincing proof of the validity of this model. Subsequently the characteristic spectra of electrons emitted from the projectile and the target particle were measured with fairly high resolution. For several years many characteristics of inelastic processes were studied, the cross sections for numerous transitions were reliably calculated, and it became possible to relate the energies of Auger transitions to an ionization state and even to identify very weak processes. The final triumph of the molecular model was possibly the discovery of x radiation and Auger electrons as a result of transitions between

short-lived quasimolecular states at the moment of impact.[7] Owing to this observation, all doubts that the molecular orbitals of H_2^+ are the key to understanding the mechanism of slow ion–atom collisions disappeared. Amazingly, the earlier discoveries of Afrosimov and co-workers, which were seemingly unrelated to the resonances observed upon photoabsorption and electron impact, ultimately turned out to be unequivocally related to Auger processes; therefore, investigations of ion–atom and atom–atom collisions are even more important. Slow collisions of many-electron atoms are accompanied by the formation of unusual states, which can be investigated by analyzing their decay products. The reliability of standard calculation methods, such as the Hartree–Fock method[8] and the *R*-matrix method,[9] has increased as the power of computers has grown. New methods for taking into account electron correlations in multiply excited states have been proposed.[10] Due to a lack of time, I am unable to list all the achievements resulting from the intensive study of these states which has been carried out by the community of scientists specializing in the physics of atomic collisions. I will mention only one discovery, which attests to the diversity of the new phenomena associated with the unusual, multiply excited resonant states formed during ion–atom collisions.

States of atoms with two inner-shell vacancies can form with a high probability during ion–atom collisions. Even in this case the atom usually relaxes by means of Auger transitions, which are essentially identical to the Auger transitions in atoms with one inner-shell vacancy. Such transitions are typically associated with the correlated motion of two electrons: one of them loses energy and fills the inner-shell vacancy, while the other escapes, carrying off a fixed amount of energy. In the case where two inner-shell vacancies form, they are usually filled by two independent two-electron transitions. In principle, there is a probability that the two electrons filling the vacancies will transfer their energy to one electron as a result of a three-electron transition, which is specified by the contribution of the correlations to the initial or final state. Such transitions were discovered by scientists working at the Physicotechnical Institute.[11] Their discovery is attributable to the development of highly sensitive instruments and to the fact that atoms with two inner-shell vacancies form in atomic collisions with cross sections close to the geometric value.

Along with the processes in the inner shells, transitions involving electrons in valence shells were investigated. Oscillations of the excitation cross sections[12] revealed effects associated with interference between different ''trajectories'' leading to the same final state. By plotting the excitation cross section as a function of the reciprocal velocity $1/v$, which is proportional to the collision time, we can determine the phase and the energy difference $E_1(t)-E_2(t)$ of the quasimolecular states which participated in the transition. The possibility of relating cross-section oscillations to levels of a quasimolecule gives us a valuable tool for analyzing atomic collisions within the molecular-orbital model adopted.

The further development of the conception of ''adiabatic'' states was associated with the discovery that slow electrons, which do not participate in Auger processes (such electrons are emitted as a result of direct ionization), have the following energy distribution:[13]

$$\sigma(E_e)\approx C\exp[-(E_e-\delta)/\alpha v].$$

This relation holds over a broad energy range, which corresponds to variation of the cross section $\sigma(E_e)$ by 5 orders of magnitude. The basis for this formula was given in Ref. 14. Following the remark in Ref. 15 that the adiabatic energy levels $E_n(R)$ should be regarded as separate branches of a single function $E(R)$ of the complex variable R, Ovchinnikov and Solov'ev[14] calculated this function for the H_2^+ system. The various branches of $E(R)$ are joined at branch points. The probabilities of nonadiabatic transitions between states whose levels are joined at branch points are calculated by estimating the phase integrals along paths which begin and end on the real R axis and circumvent the branch points. Solov'ev[16] discovered series of branch points, which have been termed ''superpromotions'' or S series, along which a $2p\sigma$ electron of the united atom can be promoted to a continuum state with an energy E_e. One of the remarkable features of this theory is that experimental data on cross sections can be described by the properties of the function of the adiabatic energy $E(R)$ at complex values of R.

The use of adiabatic states within the approximate hidden crossing theory led to definite success, but it ran into some other difficulties. The exact numerical solution of the time-dependent Schrödinger equation is a goal which has essentially been achieved owing to the continuous improvement of computer technology. Direct application of the standard method of coupled molecular eigenvalues, which is also known as the perturbed stationary state technique, is impossible, although, as follows from the experimental data, electrons move along molecular orbits in a first approximation. It can be presumed that the requirement of Galilean invariance rules out the possibility of using such states as a basis set for nonempirical calculations. It turned out that this difficulty can be overcome by many methods, the best of which, in my opinion, is the use of the scaling transformation proposed in Ref. 17. I will not discuss the details of this transformation and will only state that it leads to a satisfactory theory.

In fact, a satisfactory formulation of the problem in terms of adiabatic states reveals the main reason why they are unsuitable for performing complete nonempirical calculations of ionization, viz., the need to use an infinite number of states. This is illustrated by the Demkov–Osherov model.[18] In order to go from a bound state to the continuum, an infinite number of intermediate states must be crossed. If any of them is skipped, the result becomes erroneous in principle. For this reason, the use of adiabatic states is less suitable for solving the Schrödinger equation than other approximations, for example, integration of a partial differential equation or the use of multicenter atomic states. The existence of powerful modern computers makes the last two methods highly effective, but the general physical picture, in which the behavior of the orbitals of the H_2^+ system is a source of information on ion–atom collisions, is lost when they are used.

The solution of the dilemma described calls for reconstructing the diagrams of adiabatic energy levels. In this

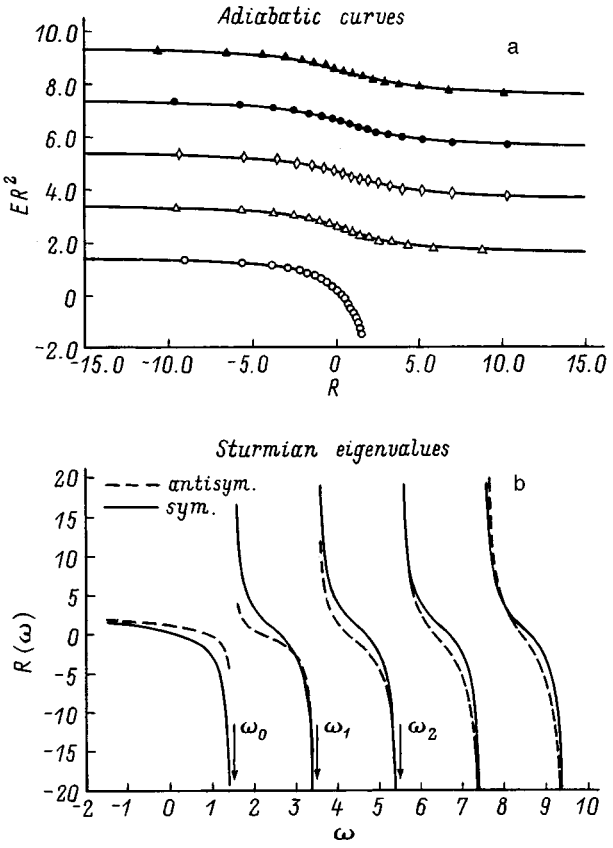

FIG. 1. Dependence of $\omega = E(R)R^2$ on R for a particle in a field of two zero-range potentials (a) and inverse dependence of $R(\omega)$ on ω (b). All the values are in atomic units.

case, the energy variable is replaced by the quantity $E(R)R^2$. Figure 1a shows plots of $\omega = E(R)R^2$ for a model system of an electron in a field of two zero-range potentials.[19,20] Let us turn this diagram through 90° and examine the function $R(\omega)$ (Fig. 1b). Only one eigenfunction $\varphi(R(\omega),q)$, i.e., the modified eigenfunction $S(\omega,q)$, corresponds to this function. The exact solution of the model problem is expressed in terms of the single-valued function presented above. It has been found that owing to the scaling properties of the Coulomb potential, such a device can be applicable to diatomic systems. In this case there is an infinite number of the functions $S_n(\omega,q)$, where $q = r/R$ is now the reduced coordinate of the electron, and the reduced energy was defined above. Unlike the case of an adiabatic basis set, in our case the amplitude of the transition between any initial and any final state, including the continuum, is specified by only one function. The development of this theory was a result of cooperation between scientists of the Physicotechnical Institute and the Oak Ridge National Laboratory.

The traditional course of the studies which we have carried out in cooperation with scientists of the Physicotechnical Institute has been a constructive interaction of theory and experiment. This tradition is presently being continued, although the financing of experimental work in the physics of atomic collisions is becoming increasingly problematic. In conclusion, I would like to present an example of how the theory which we developed predicted several new properties that were subsequently discovered in experiments.

The new states $S_n(\omega,q)$ are the eigenfunctions corresponding to the eigenvalues $R_n(\omega)$. Now these eigenvalues are the coefficients multiplying the potential $V(q)$ in the scaled Schrödinger equation, and at positive ω they acquire complex or even negative values. In the latter case the potential is effectively reversed, and the wave function is concentrated in the region midway between the two protons, where the local potential is described well by an harmonic oscillator. The electrons have a probability of advancing into the harmonic-oscillator region by means of a mechanism for movement at the top of the potential barrier; this leads to the appearance of a so-called T series of branch points. There are two possibilities for the electrons to reach the continuum. One of them leads to a σ-type distribution, and the other leads to a π-type distribution. Since the contribution of the corresponding trajectories is coherent, we should expect, as in the case of the oscillations[12] for electron excitation and capture processes, the appearance of similar oscillations of the electron distribution density. The distribution densities oscillate with a frequency which depends on the difference between the real parts of the complex eigenenergies of the harmonic oscillator, as is shown in Fig. 2. In fact, the oscillation frequency is a measure of the energy difference taken at complex values of the internuclear distance R. This energy difference decreases with the internuclear distance fairly slowly, as $R^{-3/2}$; therefore, the corresponding phase integral is fairly large:

$$\phi = v^{-1}\int_{R_0}^{\infty} \Delta E(R)\, dR \approx 2C_1/(vR_0^{1/2}),$$

where R_0 is of the order of unity and the real part of C_1 equals 4 in atomic units.[21]

Similar oscillations in the distribution of the very slow electrons formed in collisions of protons with helium atoms were discovered in Ref. 22.

The imaginary part of the phase integral also has a physical meaning, which is best traced in the context of electron-impact ionization. In this case another modification of adiabatic states, in which the hyperradius R plays the role of the adiabatic variable must be used. It turns out that the function $R(\omega)$ is suitable for describing electron-impact ionization, and a theory similar to the theory of ion–atom collisions can be developed. States at the top of the barrier also appear in it, and they can be described using the ''inverse adiabatic states'' $S(\omega,\hat{R})$, where $\hat{R}$ denotes all the coordinates of the particles, except the hyperradius R. One of the stages of cooperation between the Physicotechnical Institute and our team was devoted to applying the theory which we developed to electron–atom collisions.[23] In this case the imaginary part of C_1 is related to the famous Wannier threshold law. Our calculations exhibited very close agreement with the data from measurements of ionization cross sections in the threshold region. It can be concluded that adiabatic states are applicable, in principle, to essentially all low-energy processes in which atoms or molecules break up into three charged fragments.

In summation, it can be stated that the conception of adiabatic states has undergone considerable development in

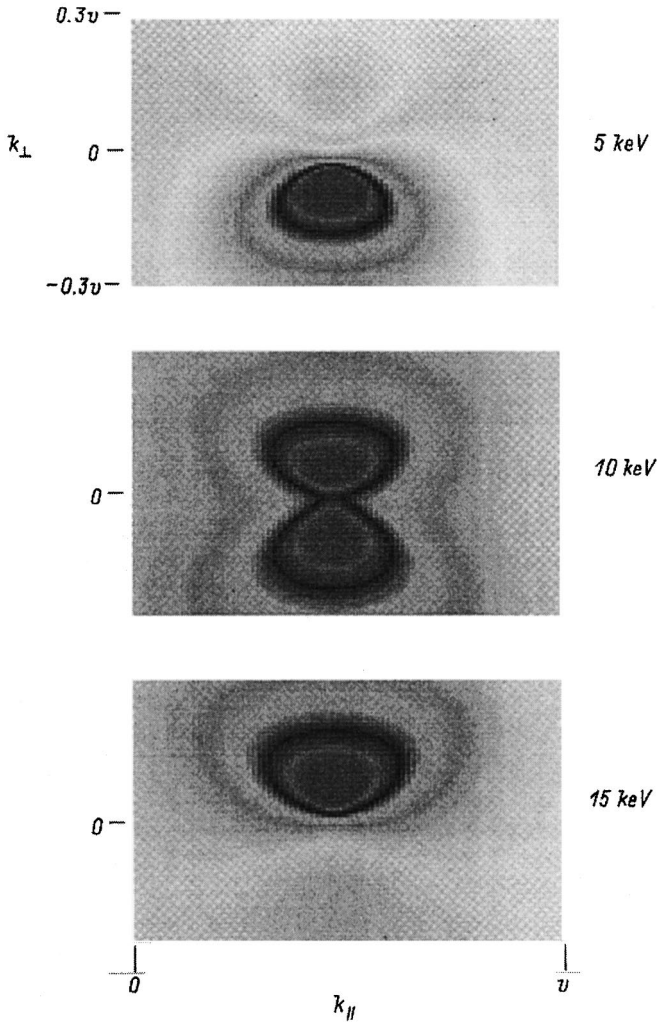

FIG. 2. Distribution densities of electrons emitted during collisions of 5-, 10-, and 15-keV protons with hydrogen atoms for a fixed impact parameter $b=1.2$ a.u. (from Ref. 21); $k_\parallel$ and $k_\perp$ — components of the electron wave vector parameter and perpendicular to the axis of the primary beam.

the last few years with an important contribution from the Physicotechnical Institute. This brief retrospective analysis has allowed us to recall past achievements and to predict the future creation of a complete theory of atomic systems consisting of a small number of particles on the basis of one of the versions of the theory of adiabatic states. Ion–atom collisions, however, are no longer at the forefront of atomic physics. Bose condensation and cold collisions are presently more promising and intriguing areas. However, in these cases, too, hyperspherical adiabatic states have been used to test the limits of applicability of the ordinary theory.[24] I cannot foresee the future, but I can speculate that adiabatic states will astonish us further with possibilities for application to the calculation of diverse dynamical processes.

[1] P. G. Burke and H. M. Shey, Phys. Rev. **126**, 147 (1962).
[2] G. N. Ogurtsov, Rev. Mod. Phys. **44**, 1 (1972).
[3] V. V. Afrosimov, Yu. S. Gordeev, M. N. Panov, and N. V. Fedorenko, Zh. Tekh. Fiz. **34**, 1613, 1624, 1637 (1964) [Sov. Phys. Tech. Phys. **9**, 1248, 1256, 1265 (1965)].
[4] N. V. Fedorenko, Zh. Tekh. Fiz. **24**, 784 (1954); V. V. Afrosimov and N. V. Fedorenko, Zh. Tekh. Fiz. **27**, 2557 (1957) [Sov. Phys. Tech. Phys. **2**, 2378 (1957)].
[5] W. Lichten, Phys. Rev. **131**, 229 (1963); **164**, 131 (1967); U. Fano and W. Lichten, Phys. Rev. Lett. **14**, 627 (1965).
[6] M. E. Rudd, T. Jorgensen Jr., and D. J. Volz, Phys. Rev. **151**, 28 (1966).
[7] F. W. Saris, W. F. van der Weg, H. Tawara, and R. Laubert, Phys. Rev. Lett. **28**, 717 (1972).
[8] C. F. Fischer, T. Braga, and P. Jönsson, *Computational Atomic Structure: An MCHF Approach*, Institute of Physics Publishing, Bristol (1997).
[9] G. W. F. Drake (ed.), *Atomic, Molecular & Optical Physics Handbook*, American Institute of Physics, Woodbury, N.Y. (1996), p. 536.
[10] J. H. Macek, J. Phys. B **1**, 831 (1968); C. D. Lin, Rep. Prog. Phys. **257**, 1 (1995).
[11] V. V. Afrosimov, Yu. S. Gordeev, and A. N. Zinoviev *et al.*, JETP Lett. **21**, 249 (1975).
[12] H. Rosenthal and H. M. Foley, Phys. Rev. Lett. **23**, 480; S. V. Bobashev, JETP Lett. **11**, 260 (1970).
[13] P. H. Woerlee, Yu. S. Gordeev, H. de Waard, and F. W. Saris, J. Phys. B **14**, 527 (1981).
[14] S. Yu. Ovchinnikov and E. A. Solov'ev, Comments At. Mol. Phys. **22**, 69 (1988).
[15] Yu. M. Demkov, in *Proceedings of the 5th International Conference on the Physics of Electronic and Atomic Collisions*, edited by I. P. Flaks and E. S. Solovyev, Leningrad (1967), p. 186.
[16] E. A. Solov'ev, Teor. Mat. Fiz. **28**, 240 (1976) [Theor. Math. Phys. **28**, 757 (1976)].
[17] E. A. Solov'ev and S. I. Vinitsky, J. Phys. B **18**, L557 (1985).
[18] Y. N. Demkov and V. I. Osherov, Zh. Eksp. Teor. Fiz. **53**, 1589 (1967) [Sov. Phys. JETP **26**, 916 (1968)].
[19] D. B. Khrebtukov and J. H. Macek, J. Phys. A **31**, 2853 (1997).
[20] S. Yu. Ovchinnikov, J. H. Macek, and D. B. Khrebtukov, Phys. Rev. A **56**, 2872 (1997).
[21] S. Yu. Ovchinnikov and J. Macek, Phys. Rev. Lett. **80**, 2298 (1997).
[22] R. Dörner, H. Khemliche, M. H. Prior *et al.*, Phys. Rev. Lett. **77**, 4520 (1996).
[23] J. H. Macek, S. Yu. Ovchinnikov, and S. V. Pasovets, Phys. Rev. Lett. **74**, 4631 (1995).
[24] B. D. Esry and C. H. Greene, *Abstracts of the 16th International Conference on Atomic Physics*, edited by W. E. Baylis and G. W. F. Drake, University of Windsor, Canada (1998), p. 101.

Translated by P. Shelnitz

Multiple-quantum radio-frequency spectroscopy of atoms: application to the metrology of geomagnetic fields

E. B. Aleksandrov, M. V. Balabas, A. K. Vershovskiĭ, and A. S. Pazgalev

S. I. Vavilov State Optical Institute All-Russia Science Center, 199034 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 27–30 (September 1999)

The evolution of the radio-frequency magnetic resonance spectrum of optically oriented potassium atoms in terrestrial magnetic fields as a function of the amplitude of the alternating magnetic field H_1 is investigated theoretically and experimentally. It is shown that among the set of observed n-quantum resonances which satisfy the selection rules $\Delta F=0$ and $\Delta m_F=\pm n$ (n is an integer), there is one resonance with the highest multiplicity $n=4$ ($m_F=2\Leftrightarrow m_F=-2$), which has exceptional properties. These qualities make this resonance a candidate for use in a weak-field quantum magnetometer with record-high characteristics. A preliminary investigation of an experimental model of a four-quantum potassium magnetometer is performed. © *1999 American Institute of Physics.* [S1063-7842(99)00709-6]

INTRODUCTION

The Zeeman effect in the ground-state radio-frequency spectra of several alkali-metal atoms has been utilized since the late 1950s for highly accurate measurements of the absolute value of the induction of weak magnetic fields mainly in the geomagnetic range. The demands for precision measurements of this kind are very diverse, and the accuracy requirements imposed have no upper bound, creating a situation that constantly stimulates the search for new solutions. The greatest absolute accuracy and variational sensitivity have been achieved using an optically pumped potassium-vapor magnetometer, which utilizes a narrow isolated resonance with $\Delta F=0$ and $m_F=2\Leftrightarrow m_F=1$ in the ground-state radio-frequency spectrum of one of the two stable isotopes of potassium.[1] Apart from the one-quantum transitions with $\Delta m_F=\pm 1$, multiple-quantum resonances with $\Delta m_F=\pm n$ appear when the amplitude of the resonance-stimulating alternating magnetic field H_1 is increased, and they are very effective in the quasiequidistant magnetic splitting system of the ground-state hyperfine sublevels. At the time of their discovery in the early 1950s,[2,3] these resonances seemed attractive because of their intrinsic width, which successively decreases with increasing n. However, it was soon found that their frequency depends on the amplitude of the alternating field and that the stronger it is, the higher is the order of the resonance. This dramatically reduced the interest in these resonances, and since then they were not employed in magnetic-field metrology.

In this report we turn our attention to the special properties of the resonance with the highest multiplicity $n=4$, which is essentially not displaced by the alternating field. In addition, the frequency of this resonance (in contrast to all the others) exhibits a strict linear dependence on the induction of the static magnetic field. Being very narrow and very intense, this resonance merits study as a candidate for use in a quantum magnetometer with record-high characteristics.

The easily predicted features of the four-quantum resonance might have provided an exceptional experimental technique back in the 1950s. However, in the experiments at that time experimentalists focused on three-quantum resonances for technical reasons: the four-quantum resonance would have required a radio-frequency field with an excessively high amplitude. Under our conditions we worked with such narrow resonances and in such weak static fields that only negligible amplitudes were needed to excite the four-quantum resonance.

ANALYSIS OF THE EVOLUTION OF THE MAGNETIC RESONANCE SPECTRUM OF POTASSIUM

In a weak magnetic field the ground state of the potassium atom forms two systems of almost equidistant magnetic sublevels separated by a hyperfine interval. An example of the energy structure of the ^{39}K isotope is shown in Fig. 1a, where the magnetic splitting has been greatly magnified for clarity. In a terrestrial field the spectrum of transitions with $\Delta m_F=\pm 1$ consists of six closely arranged lines in the vicinity of a frequency equal to $H\cdot 7$ Hz/nT, where H is the induction of the field in nanoteslas. The exact values of the frequencies of these transitions obey the familiar Breit–Rabi formula and were given in Ref. 4. Under the conditions of optical orientation the transitions within the $F=2$ state are most clearly expressed in the spectrum, being manifested by four almost exactly equidistant lines separated by an interval equal to $2bH^2$, where the value of b for the ^{39}K isotope is 106.327 GHz/T^2. In an average terrestrial field of 50 μT this interval amounts to about 500 Hz.

In a vacuum flask with a paraffin cover the intrinsic width of the resonances of potassium vapor is of the order of 1 Hz. The possibility of using one completely isolated, very strong line in the spectrum gives potassium a decisive advantage over cesium, which has been most popular in quantum magnetometry, but whose numerous overlapping resonances form a broad (of the order of 50–100 Hz) asymmetric line with an indefinite position for the maximum, creating a ma-

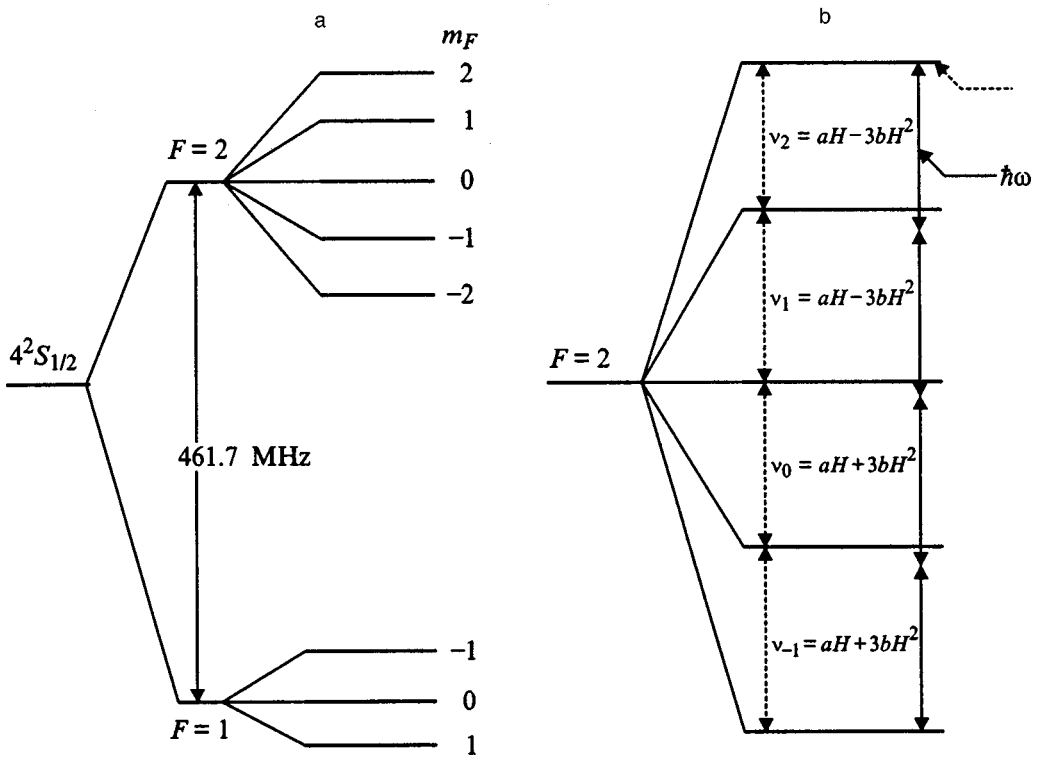

FIG. 1. Scheme of the ground-state energy sublevels of the ^{39}K atom (a) and detailed structure of the magnetic splitting of the upper ground-state hyperfine sublevel of the ^{39}K atom in a magnetic field with induction H with accuracy to the terms quadratic with respect to the field for $a = 7.004666$ GHz/T and $b = 106.327$ GHz/T^2 (b).

jor source of systematic errors in cesium magnetometers. Potassium magnetometers are completely free of errors associated with the influence of nearby resonances. The small width of the principal resonance together with its high amplitude (the signal-to-noise ratio is of the order of 10^4 in a 1-Hz band) ensures the record-high sensitivity of potassium magnetometers. Figure 2 presents a fragment of a record of the difference between the readings of two independent potassium magnetometers, whose sensors were separated by two meters. The differential record enabled us to suppress the natural fluctuations of the Earth's magnetic field by a factor of 100 and to approach the internal noise of the instruments. The noise track presented, which is confined to 1/4 pT, gives an upper estimate of the noise of the instruments. The record was made at the rate of 5 readings per second and provided a picture of the resolving power achieved in the instrument.

A spectrum of 6 isolated lines is observed for a fairly low amplitude of the radio-frequency field H_1, which is assigned by the relation $\gamma H_1 \leqslant \Gamma$, where γH_1 is the matrix element for a transition between adjacent sublevels and Γ is the intrinsic width of the transition. At such a low amplitude of H_1 each atom responds to the perturbation as a set of independent two-level systems. As H_1 is increased, the spectrum begins to change: the lines broaden and shift, and new lines corresponding (in terms of perturbation theory) to multiphoton resonances appear. Field-induced line broadening is well known in the theory of magnetic resonance. The field-induced displacement of a resonance in a two-level system is also well known as the Bloch–Siegert effect. However, a far stronger field-induced shift of the resonance is exhibited in multilevel quasiequidistant systems. In second-order perturbation theory the energy shift ΔE_j of a certain level j is given by the expression[5]

$$\Delta E_j = \Sigma_k |V_{jk}|^2 / (E_j - E_k + \hbar\omega), \tag{1}$$

where $|V_{jk}|$ is the matrix element associating the jth level with any possible kth level having an energy E_k, and $\hbar\omega$ is the quantum energy of the field H_1.

The formula presumes that the kth level lies above the jth level (otherwise, the sign in front of $\hbar\omega$ changes), and the denominator is fairly large compared to the width of the transitions, i.e., the displacement of levels under the action of a nonresonant field is considered (a resonant field does not

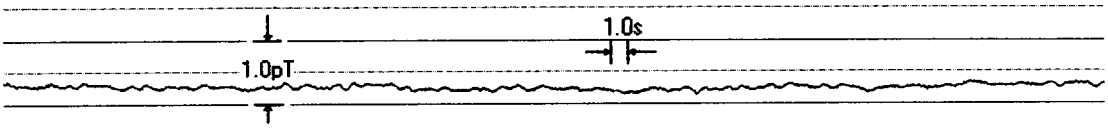

FIG. 2. Experimental record of the difference between the readings of two commercial GEM Systems potassium magnetometers in the Earth's natural field (Canada, April 1998).

shift the levels). The frequency shift of a transition is calculated as the difference between the energy shifts of the combining levels. As H_1 increases, the field-induced shift increases quadratically and becomes significant in the region where multiple-quantum resonances appear, increasing with increasing multiplicity of the resonance. The resonance of highest multiplicity is an exception. This can be seen by applying formula (1) to the calculation of the shifts of the $m_F=2$ and $m_F=-2$ levels under the action of the field H_1 with the frequency aH (Fig. 1b), which corresponds to 1/4 of the frequency of the unperturbed gap between the $m_F=2$ and $m_F=-2$ levels. This frequency exceeds the resonance value for the $m_F=2\Leftrightarrow m_F=1$ transition by $3bH^2$, while it falls short of the resonance value for the $m_F=-1\Leftrightarrow m_F=-2$ transition by the same amount. As a result, the upper and lower sublevels of the $m_F=2\Leftrightarrow m_F=-2$ transition are shifted identically in the same direction, so that the transition frequency itself remains unchanged (Fig. 1b).

The conclusion that there is exact compensation of the field-induced shifts of the levels is based on the use of perturbation theory and requires more definitive confirmation for two reasons. First, the excitation of the four-quantum resonance near saturation requires a fairly high value for H_1, which violates the criterion for a small perturbation. Second, in arriving at the conclusion that the field-induced shifts are compensated, the lines of the one-quantum resonances were assumed to be strictly equidistant, i.e., the terms higher than quadratic in the expansion of the transition frequencies in powers of H were discarded. In addition, the influence of the wings of neighboring resonances that are broadened at high values of H_1 on the position of the maximum of the four-quantum resonance should also be taken into account. For all these reasons it was decided to carry out an exact solution of the problem of magnetic resonance in an 8-level system under the conditions of optical pumping without restrictions on the amplitude of the field H_1. For this purpose we solved the Liouville equation for the density matrix, which was supplemented by a phenomenological relaxation matrix describing the optical pumping. During the search for a stationary solution in the rotating-field approximation, a system of 34 differential equations was reduced to a system of algebraic equations, which was solved using a personal computer for a grid of values of the frequency and amplitude of H_1.

The result is presented in Fig. 3 (Ref. 4). The lowest spectrum corresponds to the condition $\gamma H_1=2\pi$ Hz (the intrinsic width of the resonances was assumed to be equal to 1 Hz). The four resonances in the $F=2$ state can be seen here. The resonances in the $F=1$ state are too small for this scale. When γH_1 is increased to $2\pi\cdot 20$ Hz, the one-quantum resonances are strongly broadened and appreciably displaced, while the three two-quantum resonances are optimized. When $\gamma H_1=2\pi\cdot 80$ Hz, the two three-quantum resonances achieve maximum steepness, while the two-quantum resonances have already reached saturation and broadened. Finally, when $\gamma H_1=2\pi\cdot 190$ Hz, the four-quantum resonance dominates. An analysis showed that the frequency of this resonance responds to a 10% change in H_1 no more strongly than does the frequency of a one-quantum resonance in the case of the analogous change in H_1 in the vicinity of the optimum value $\gamma H_1=2\pi$ Hz, which is of the order of 50 fT and is thus essentially a negligible quantity. The steepness of the four-quantum resonance at the optimum is 7 times higher than the steepness of a one-quantum resonance.

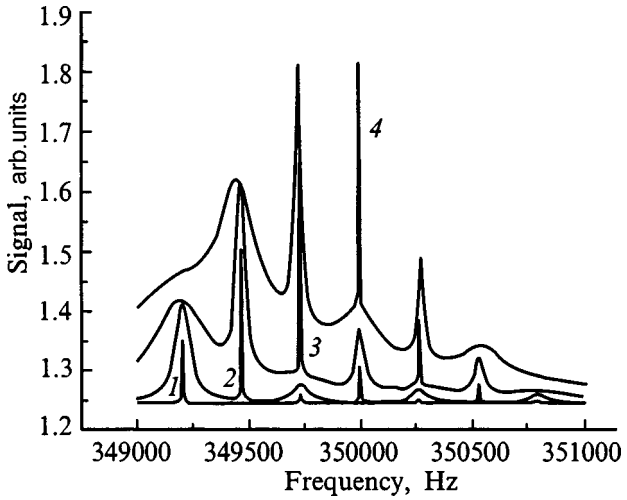

FIG. 3. Calculated ^{39}K magnetic resonance spectrum for four values of the field amplitude: *1)* 2π, *2)* $2\pi\cdot 20$, *3)* $2\pi\cdot 80$, *4)* $2\pi\cdot 190$ Hz. The calculation was performed for a static field H with an induction of about 50 μT.

Thus, the mathematical simulation confirmed the expected promising nature of the four-quantum resonance. This conclusion has also been confirmed by experimental testing.[6] In addition to the data presented in Ref. 6, we present a set of panoramic magnetic resonance spectra, which were recorded by the technique of low-frequency modulation of the frequency of H_1. The signal was recorded from the output of a synchronous detector, so that the resonance lines had the from of dispersion curves. In Fig. 4 the upper spectra correspond to higher intensities of H_1. The lowest spectrum displays the one-quantum resonances: three belong to the $F=2$ state and one of the $F=1$ resonances. In the next spectrum two two-quantum resonances dominate. The third spectrum displays broadened two-quantum resonances and

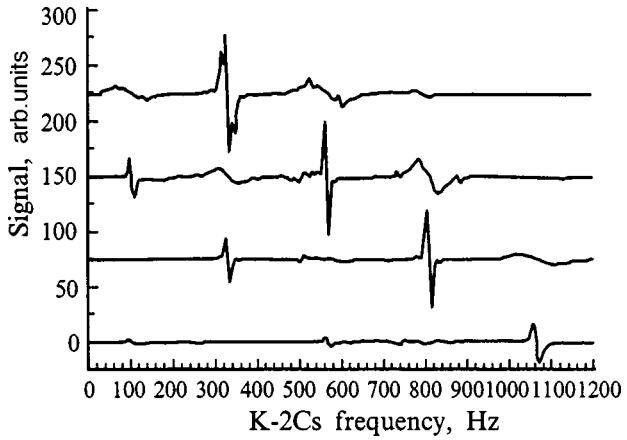

FIG. 4. Experimental potassium magnetic resonance spectra for various values of H_1. The spectra were recorded in a field with a magnetic flux density of about 50 μT. The frequency of the field H_1 measured relative to the doubled resonance frequency in cesium vapor, which was used to stabilize the magnetic field, is plotted along the horizontal axis.

clearly expressed three-quantum resonances. The uppermost spectrum exhibits the four-quantum resonance. The complicated shape of the resonances is associated with the excessively high modulation frequency superimposed on H_1, which generated nonstationary responses. We also note that the widths of the resonances are restricted by the instrumental width of the procedure for obtaining the spectra.

The use of the four-quantum resonance for the purposes of magnetometry is complicated by the narrowness of the resonance, which requires very slow scanning of the resonance and thus causes the instrument to have a low response rate. Therefore, a ''tandem'' scheme combining a self-generating cesium magnetometer, which provides for a fast response of the instrument, and a four-quantum potassium magnetometer, which guarantees a high absolute accuracy and linearity of the output frequency with respect to the measured field, was chosen. An instrument of this type is presently undergoing laboratory testing. Its resolving power is determined by the cesium magnetometer and is of the order of $1\ \mathrm{pT/Hz^{1/2}}$ (the rms noise). The long-term stability will apparently be limited by the photoinduced shifts of the potassium $m_F=2$ and $m_F=-2$ levels, which can be minimized by lowering the intensity of the optical pumping of potassium, if we bear in mind the high steepness of the four-quantum resonance.

CONCLUSION

The potassium atom in a fairly strong alternating field represents a new quantum system, which has, in particular, a new magnetic resonance line that is not displayed by the free atom. Until now, only one electronic paramagnet with a linear magnetic-field dependence of the resonance frequency, viz., ^{4}He in the metastable 2^3S_1 state, has been available to quantum magnetometry. The four-quantum potassium resonance has the same property, but is also a hundred times narrower. This circumstance allows us to recommend this resonance for standardizing the induction of fields in the geomagnetic range.

[1] E. B. Alexandrov, M. V. Balabas *et al.*, Laser Phys. **6**, 244 (1996).
[2] J. Brossel, B. Cagnac, and A. Kastler, C. R. Acad. Sci. URSS **237**, 984 (1954).
[3] P. Kusch, Phys. Rev. **93**, 1022 (1954).
[4] E. B. Aleksandrov and A. S. Pazgalev, Opt. Spektrosk. **80**, 534 (1996) [Opt. Spectrosc. **80**, 473 (1996)].
[5] A. M. Bonch-Bruevich and V. A. Khodovoĭ, Usp. Fiz. Nauk **85**, 3 (1965) [Sov. Phys. Usp. **8**, 1 (1965)].
[6] E. B. Aleksandrov, A. S. Pazgalev, and Zh. L. Rasson, Opt. Spektrosk. **82**, 14 (1997) [Opt. Spectrosc. **82**, 10 (1997)].

Translated by P. Shelnitz

Many-electron effects in atomic processes

M. Ya. Amus'ya

Racah Institute of Physics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

L. V. Chernysheva

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 31–35 (September 1999)

The major role of the collectivization of electrons in atoms and quasiatomic formations is demonstrated. The random-phase approximation with exchange, which permits allowance for these effects, is discussed in detail. The need to extend the scope of this approximation when some processes are considered, which is achieved by combining it with perturbation theory, is noted. The role of the collective effects is illustrated by the results of recently performed calculations of the photoionization cross sections of atomic iodine and its positive and negative ions, the single-electron ionization of Xe^+, resonance-enhanced photon emission in collisions of electrons with atoms and quasiatomic formations, the nondipole corrections to the angular distribution of photoelectrons, and the probabilities of two-electron transitions in which all the energy is released in the form of a single photon. [S1063-7842(99)00809-0]

INTRODUCTION

The purpose of this paper is to demonstrate the significant, and in some cases decisive, role of the so-called residual interaction between electrons in an atom. The residual interaction is the part of the ordinary Coulomb repulsion which cannot be taken into account by any choice of the mean atomic field in which the electrons are immersed.

The mean field in an atom is known to be fully comparable in strength to the nuclear field. Therefore, it is natural that it should be taken into account exactly. We shall show that the residual interaction also leads to significant many-electron (collective) effects, which must be taken into account exactly, rather than by perturbation theory. The parameter ξ, which gives an estimate of the role of collectivization, is the ratio between the potential energy W of the interaction of a given electron with all the simultaneously excited electrons and the excitation kinetic energy ΔE:[1)]

$$\xi = W/\Delta E \sim \left(\frac{1}{R}\cdot P_0^2 \cdot \frac{R^3}{R}\right) \Big/ \; P_0 \frac{1}{R} \sim R^2 \rho^{1/3}, \qquad (1)$$

where P_0 is the Fermi momentum, which is related to the density ρ by the expression $\rho = P_0^3/3\pi^2$, and R is the characteristic dimension of the system.

In heavy atoms $\xi \approx 1$, and in quasiatomic formations, such as metal clusters, $\xi \gg 1$, which indicates the significant or even very large degree of collectivization of the electrons in these objects.

Collectivization is most clearly manifested in photoabsorption, the inelastic scattering of fast electrons, and vacancy decay. These processes also take place in the one-electron approximation, but, as will be shown below, their course changes radically when the residual interaction is taken into account. There are also processes which can exist only when the residual interaction is taken into account, for example, the removal of two electrons by one photon.

The role of many-electron effects is most conveniently revealed by examining atoms, as well as positive and negative ions with subshells which are completely filled or half-filled with electrons. These effects are especially important in quasiatomic formations.

The investigations described in this report utilize the tools of many-body theory and Feynman diagram techniques.

THE HARTREE–FOCK APPROXIMATION AND THE RANDOM-PHASE APPROXIMATION WITH EXCHANGE

In order to reveal the role of many-electron effects, we must first clearly define what is selected as the one-electron approximation. Knowledge of what is taken into account by the latter is needed to avoid taking into account the same corrections twice. In this study the Hartree–Fock (HF) approximation is taken as the one-electron approximation. It can be regarded as the most exact of the systematic one-electron approximations. In it electrons are regarded as moving in a single self-consistent (i.e., determined by the states of the electrons themselves) nonlocal potential field.

The HF equations are solved numerically.[1] Having their solution, i.e., the wave functions and the electron energies, we can calculate the cross sections for the absorption and scattering of photons by atoms and the cross sections for the scattering (elastic and inelastic) of electrons on atoms. The results of such calculations differ significantly, sometimes even qualitatively, from the experimental data. This points to the importance of the collectivization of atomic electrons.

The simplest and theoretically most systematic method for taking it into account is the random-phase approximation with exchange (RPAE).[2] Within it, for example, the ampli-

tude $D(\omega)$ of the absorption of photons of energy ω is represented in the form of a sum of its HF value d and the additional term $\Delta D(\omega)$, which is defined by the expression $\Delta\hat{D}(\omega)=\hat{D}(\omega)\hat{\chi}(\omega)\hat{U}$. As a result, the entire RPAE equation in operator form is

$$\hat{D}(\omega)=\hat{d}+\hat{D}(\omega)\hat{\chi}(\omega)\hat{U}. \quad (2)$$

Here the second term describes the excitation (removal) of the electron under consideration according to the following mechanism: at first, an atomic electron passes as a consequence of the absorption of a photon of energy ω into a virtual excited state, where it resides for a time $\chi(\omega)$, and then, owing to the interelectron residual interaction $\hat{U}$, it donates the energy ω to the electron under consideration, which is excited or removed from the atom. Although (2) is an operator equation, which can be solved numerically on a computer,[1] its symbolic solution in the form

$$\hat{D}(\omega)=\hat{d}/(1-\hat{\chi}(\omega)\hat{U}) \quad (3)$$

is convenient for a qualitative RPAE analysis.

The following main results were obtained within the RPAE.

a) It was shown that the very high maxima in the cross sections for the absorption of photons by atoms and quasiatomic formations, i.e., giant resonances, are zeros of the denominator in (3):

$$1-\hat{\chi}(\Omega)\hat{U}=0. \quad (4)$$

We note that $\Omega=\Omega_0+i\Gamma$, where Γ is the reciprocal lifetime of the resonance. For the objects discussed in this report, Γ is only several times smaller than Ω_0, so that a giant resonance is a comparatively short-lived excitation of an atom. The existence of the solution of Eq. (4) is a consequence of the collectivization of electrons belonging to at least one subshell. The giant resonance is displayed most clearly above the threshold of the $4d^{10}$ subshell of Xe and its neighbors, as well as in metallic clusters and fullerenes.

b) Interference resonances and correlation minima were discovered. In an interference resonance (IR) the ionization amplitude of, for example, an outer subshell $\hat{D}_o(\omega_{\mathrm{IR}})$ is far greater than $\hat{d}_o$. This occurs because of the large amplitude $\hat{D}_i(\omega_{\mathrm{IR}})$ for the interaction of a photon with an inner subshell which is strongly coupled to that outer subshell by the residual interaction $\hat{U}_{io}$:

$$\hat{D}_o(\omega_{\mathrm{IR}})=\hat{d}_o+\hat{D}_i(\omega_{\mathrm{IR}})\hat{\chi}_i(\omega_{\mathrm{IR}})\hat{U}_{io},\ \ D_o(\omega_{\mathrm{IR}})\gg d_o. \quad (5)$$

At a correlation minimum (CM) destructive interference occurs, so that $\hat{D}_o(\omega_{\mathrm{CM}})=0$, while $\hat{d}_o\neq 0$. The interference resonances in Xe and its neighbors and the correlation minima in the subvalence shells of the noble gas atoms beginning from Ar were investigated most thoroughly.

There are some other important results in the RPAE, such as the discovery of giant autoionizing resonances (for example, in the Mn and Eu atoms), the prediction of ''autoionization'' accompanying the interaction of two continuous spectra (in contrast to ordinary autoionization, where one of them is a discrete excitation level), which is characteristic of negative ions with half-filled subshells (for example, Si^-), and the prediction of the suppression of the radiation-induced vacancy decay due to destructive interference between different channels of this process.[3]

BEYOND THE SCOPE OF THE RPAE

The RPAE has fundamental theoretical constraints, which call for extending its scope in some cases. For example, in the RPAE the ionization potential is equal to the HF energy of the electron removed with the opposite sign. Actually, however, there is a difference between these quantities, and it is very significant for outer and subvalence subshells. It is determined by the reorganization of the electronic subshells of the atom during ionization.

The fact that each atomic electron is affected not only by the HF energy but also by the polarization potential (a nonlocal quantity that depends on the energy of the electron), which is determined by virtual excitations of the atom, is also beyond the scope of the RPAE. For example, the polarization potential of alkaline-earth atoms is so great that it leads to a bound state, i.e., the formation of a negative ion.[4] Processes in which an external field or a projectile electron remove an atomic electron from an intermediate or inner subshell near its threshold also lie outside the RPAE: Auger decay of the vacancy formed sharply intensifies the attractive field acting on the slow electron removed from the atom. This leads to considerable quantitative changes in the form of both the Auger spectrum and the energy distribution of the slow electron, which, taken together, are known as the post-collision interaction.[5]

As has been found, if the electron removed from an atom has sufficient energy, it excites or removes another atomic electron with a high probability. Such ''knock-out'' has a significant effect on the energy, angular, and spin distributions of the electrons. This process is also beyond the scope of the RPAE.

The scope of the RPAE is usually extended by adding corrections to it, which can be taken into account using many-body perturbation theory (MBPT). For example, the polarization potential can be calculated to satisfactory accuracy in the second order of MBPT. However, the effect of the polarization potential on an electron should be taken into account exactly.

The theoretically rigorous extension of the scope of the RPAE is extremely complicated, since outside the RPAE even the simplest intermediate state includes two electrons interacting with one another, which are immersed in the atomic field. The description of such a state thus requires the solution of a three-body quantum-mechanical problem.

EXAMPLES OF STRONG MANY-ELECTRON EFFECTS

1. Photoionization cross section of iodine and its ions. The photoionization cross sections of Xe and Cs above the threshold of the $4d^{10}$ subshell are characterized by a high maximum and a giant resonance. It is natural to expect that the situation in iodine (I) is similar. Although experiment has demonstrated the presence of a maximum,[6,7] its height is three times smaller than in Xe. In order to clarify the situa-

tion, we performed calculations for I and I^{++} for the first time and repeated the calculations previously performed[8] for I^-, but in the framework of new computational possibilities, including up to ten interacting transitions, rather than two.[2)] As a result of the calculation, it was found that the photoionization cross sections of I and I^-, beginning at 5 Ry, differ only slightly and increase, reaching the same maximum of 29 Mb at 6.7 and 7.3 Ry, respectively. In the case of I^{++}, the region of discrete excitations is far broader, extending from 4 to 6.1 Ry. A photoionization cross section, which increases to 26 Mg at 7 Ry, is thus also found. At $\omega=10$ Ry the cross sections of I, I^-, and I^{++} essentially merge. The high reliability of the calculations performed allows us to state that the experimental data are erroneous. In any case, their absolute value requires a roughly threefold change in the normalization, which would increase the size of the giant resonance. The latter also affects the partial ionization cross sections of the $5p$ and $5s$ electrons. For example, it was shown using (5) that D_{5p} increases at its maximum under the influence of the $4d^{10}$ electrons by a factor of 2.5 in comparison to d_{5p}. This corresponds to more than a sixfold increase in the cross section in I, which reaches its greatest value at 6.5 Ry.

2. Single-electron ionization of Xe^+. It was shown comparatively recently[3] that the cross section for the single photoionization of Xe under the influence of a giant resonance has a distinct maximum above the threshold of the $4d^{10}$ subshell. There is considerable interest in the question of whether a similar maximum exists in the case of the single ionization of the Xe^+ ion, i.e., in the cross section for the formation of Xe^{++}. We[3)] performed calculations of the corresponding cross section within the RPAE. A very significant, more than an order-of-magnitude, increase in the ionization cross section of Xe^{++} under the effect of the $4d^{10}$ electrons was discovered. The calculations were performed by solving Eq. (5). The role of the second term was found to be considerably greater in Xe^+ than in Xe. The experimental data still have a considerable spread, and, in addition, there have not been any absolute measurements.[10] However, the presence of a broad maximum at 6.6 Ry is clear. After normalization of the experimental data to the calculation results, satisfactory agreement is observed over a very broad range of energies from 3 to 11 Ry. The high maximum corresponding to the discrete $4d\rightarrow 4f$ transition is reproduced, although it is shifted toward lower energies by 0.66 Ry in comparison to the calculated position.

3. Photon emission upon electron scattering. When electrons are scattered on atoms and quasiatomic formations, photons are emitted. Their source can be either the projectile electron or the target, which is virtually excited (polarized) during the collision. The total photon emission amplitude $A(\omega)$, like the photon absorption amplitude, can be represented in the form (2). In the case of fast electrons, the equation specifying $A(\omega)$ can be simplified, and its solution can be written in the form[11]

$$A(\omega,q)\sim\left[\frac{W(q)}{\omega}+\frac{4\pi\omega}{q^2}\alpha(\omega)\right]. \tag{6}$$

Here $\alpha(\omega)$ is the dynamic dipole polarizability of any target, i.e., atom or quasiatomic formation; q is the momentum transferred from the electron to the target; and $W(q)$ is the Fourier component of the static potential acting on the projectile electron. The first term in (6) describes ordinary bremsstrahlung (OB), and the second term describes the so-called ''atomic'' bremsstrahlung (AB). At small values of q, i.e., small scattering angles, the second term in (6) becomes considerably greater than the first. The strong degree of collectivization of the atomic electrons stipulates a large value for $\alpha(\omega)$ over a broad range of values of ω. The presence of a giant resonance in the spectrum enhances $\alpha(\omega)$. Therefore, AB is especially strong in polyatomic metal clusters and fullerenes, as well as in atoms with filled $4d^{10}$ and $4f^{14}$ electronic subshells. AB can also enhance the total flux of radiation with $\omega\approx\Omega$ [see (4)] by tens of fold for atoms and by many orders of magnitude for metal clusters and fullerenes. The role of OB is especially small in the latter.

4. Nondipole corrections to the angular distribution of photoelectrons. The ionization of atoms and quasiatomic formations by photons with energies amounting to tens, hundreds, and even thousands of electron volts is described in the dipole approximation, in which the photon momentum is neglected. The angular distribution of the photoelectrons is symmetric to a plane perpendicular to the direction of flow of the photons. Allowance for the quadrupolar corrections breaks this symmetry, and the photoelectron yield $I(\Theta,\omega)$ at the angle Θ to the direction of the photons is proportional to the expression[12]

$$I(\Theta,\omega)\sim[1+\beta(\omega)P_2(\cos\Theta)+\gamma(\omega)P_1(\cos\Theta)+\eta(\omega)P_3(\cos\Theta)]. \tag{7}$$

Here $P_i(\cos\Theta)$ $(i=1,2,3)$ are Legendre polynomials; $\beta(\omega)$, $\gamma(\omega)$, and $\eta(\omega)$ are fairly complicated expressions composed of the dipole and quadrupole matrix elements for the interaction of a photon with an atomic electron, as well as the scattering phases of the photoelectron wave in the field of the residual ion. The coefficients $\gamma(\omega)$ and $\eta(\omega)$ are proportional [unlike $\beta(\omega)$] to the photon momentum $\kappa=\omega/c$, where c is the speed of light.

The collectivization of atomic electrons has a significant effect on the dependence of β on ω. For example, for the Xe $5p$ electrons, β acquires an additional intense oscillation under the influence of the giant resonance.

Calculations of $\gamma(\omega)$ and $\eta(\omega)$ were recently performed for the Ar $3s$ electrons with consideration of the outer $3p^6$ subshell.[13] The influence of the latter was found to be very significant: the coefficients $\gamma_{3s}(\omega)=-\eta_{3s}(\omega)$ acquired an additional intense oscillation, so that, as a whole, these coefficients become totally different from their values in the HF approximation. We note that in the hydrogen approximation $\gamma_{ns}(\omega)$ is determined very simply and does not have any oscillations: $\gamma_{ns}(\omega)$, like $\eta_{ns}(\omega)$, is proportional to v/c, where v is the velocity of the photoelectron.

5. Two-electron transition with the emission of one photon. If two electrons simultaneously change their state and the energy released as a result is carried off by a single photon, the process occurs only by virtue of the presence of the residual interaction. Just this interaction permits the correlated, rather than independent, transition of two electrons.

The passage of a highly charged Ar^{18+} ion through a carbon film, as a result of which the ion acquires two electrons to become Ar^{16+} and emits one photon, was recently investigated experimentally.[14] The energy of the photon roughly equals

$$\omega_1 \cong 2(I_{1s} + v^2/2), \tag{8}$$

where I_{1s} is the ionization potential of Ar^{18+}, which equals $(18)^2 = 32$ Ry, and v is the velocity of the Ar^{18+} ion.

It can be seen from (8) that ω_1 is much greater than the energies of any atomic one-electron transitions. The probability of the process $Ar^{+18} + C \Rightarrow C^{++} + Ar^{+16} + \gamma$ was calculated using a model in which the field of the carbon atoms was neglected, the wave functions of the electrons were assumed to be purely Coulomb in the field of a nucleus with the charge $Z = 18$, and the electrons were assumed to have the energy $v^2/2$ in the initial state and to be bound in the $1s$ level in the final state. The interelectronic repulsion was taken into account in first-order perturbation theory. In effect, the process

$$e + e + A^{+n} \rightarrow A^{+(n-2)+} + \gamma, \tag{9}$$

was considered, and the cross section for it was calculated. The latter was compared with the cross section for single-electron recombination. Their ratio was equal to 3.6×10^{-6} in the calculation in Ref. 15 and 3.1×10^{-6} in the experiment. These figures are in satisfactory agreement with each other.

Reaction (9) should also be observed in electronic refrigerators of highly charged ions, in which the relative velocities of the electrons and ions can be considered equal to zero. The cross section of the process will depend not only on the electron density, but also on the structure of the electron beam, i.e., the presence or absence of correlations between the positions of the electrons.

At first glance, process (9) is the temporal reverse of two-electron photoionization; therefore, their cross sections should be related by the principle of detailed balance. However, the cross section of (9) actually corresponds to a very small part (especially for large v) of the cross section for double photoionization, specifically the part in which the ejected electrons fly off in the same direction with equal velocities. The main contribution to two-electron photoionization comes from the portion of the spectrum where one of the electrons carries off virtually all the energy of the photon and the other electron moves away slowly.

CONCLUSION

We have demonstrated the significant, and sometimes decisive, role of the residual interaction and the resultant collectivization of atomic electrons on the basis of general estimates and in some concrete examples. The RPAE developed to take it into account, which has been reconciled with the original single-particle HF approximation, has achieved an extremely high internal accuracy in atomic physics and led to success in describing many processes and characteristics. However, extending the scope of the RPAE is necessary to describe some of them. It is very difficult, although extremely desirable, to do this systematically in a theoretical treatment.

One of the closely related and important problems is the development of methods for solving two- and three-dimensional RPAE equations.[4)] In the case of success, it would be possible to quantitatively examine the electronic structure of complicated two-atom molecules, medium-weight atoms, and heavy atoms in strong external electric and magnetic fields, as well as aspherical metal clusters and fullerenes to the accuracy which has already been achieved in the description of isolated many-electron atoms.

There are numerous questions and corresponding objects of investigation, which can and should be investigated using the RPAE, such as, for example, the photoionization of excited states of atoms. Objects of investigation which seem exotic at first glance merit serious attention. One example is provided by highly excited spin-saturated states of atoms, i.e., states in which all the electron spins are oriented in the same direction. The lifetime of such states can easily reach several microseconds, which are sufficient for experimental investigations. Collectivization should be markedly greater in spin-saturated states than in ground atomic states.

An even more exotic object of investigation consists of electron bunches which are trapped and retained by a laser field. When the retention time is sufficiently long, a structure should appear in such bunches, and collective many-electron effects should be observed in processes involving them.

As a whole, the field of many-electron effects in atomic processes is virtually unlimited.

[1)] We have used the atomic system of units, in which $e = \hbar = m_e = 1$.
[2)] The calculations were performed together with S. T. Manson and N. A. Cherepkov.[9]
[3)] Together with S. T. Manson and N. A. Cherepkov.[9]
[4)] In the case of atoms, after separation of the angular and spin variables, these equations are one-dimensional.

[1] M. Ya. Amusia and L. V. Chernysheva, *Computation of Atomic Processes*, Institute of Physics Publishing, Bristol–Philadelphia (1997), 253 pp.
[2] M. Ya. Amusia, *Atomic Photoeffect*, Plenum Press, New York–London (1990), 272 pp.
[3] M. Ya. Amusia, in *VUV and Soft X-Ray Photoionization*, edited by U. Becker and D. A. Shirley, Plenum Press, New York (1996), pp. 1–46.
[4] G. F. Gribakin, B. V. Gul'tsev, V. K. Ivanov, and M. Yu. Kuchiev, J. Phys. B **23**, 4505 (1990).
[5] M. Yu. Kuchiev and S. A. Sheĭnerman, Usp. Fiz. Nauk **158**, 353 (1989) [Sov. Phys. Usp. **32**, 569 (1989)].
[6] L. Nahon, F. Svensson, and P. Morin, Phys. Rev. A **43**, 2328 (1991).
[7] G. O'Sullivan *et al.*, Phys. Rev. A **53**, 3211 (1996).
[8] M. Ya. Amusia, G. F. Gribakin, V. K. Ivanov, and L. V. Chernysheva, J. Phys. B **23**, 385 (1990).
[9] M. Ya. Amusia, N. A. Cherepkov, L. V. Chernysheva, and S. T. Manson, *DAMOP98, Abstracts of Papers*, Atlanta, Georgia (1998), p. 9.
[10] T. M. Kojima, M. Sano, N. Watanabe *et al.* in *Proceedings of the 20th ICPEAC*, Vienna (1990), p. MO 030.
[11] M. Ya. Amus'ya, *Bremsstrahlung* [in Russian], Énergoatomizdat, Moscow (1990), 204 pp.
[12] M. Ya. Amus'ya and V. K. Dolmatov, Zh. Éksp. Teor. Fiz. **79**, 1664 (1980) [Sov. Phys. JETP **52**, 840 (1980)].
[13] M. Ya. Amusia, A. S. Baltenkov, Z. Felfli, and A. F. Msezane, in *VUF'12*, San Francisco (1998), p. 12.
[14] A. Warczak *et al.*, Nucl. Instrum. Methods Phys. Res. B **98**, 303 (1995).
[15] V. L. Yakhontov and M. Ya. Amusia, Phys. Rev. A **55**, 1952 (1997).

Translated by P. Shelnitz

Spin exchange and chemi-ionization during collisions of polarized metastable helium atoms with ground-state cesium atoms

S. P. Dmitriev, N. A. Dovator, and V. A. Kartoshkin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 36–40 (September 1999)

The kinetics of the optical orientation of atoms in a helium–cesium gas-discharge plasma are considered, and kinetic equations describing the optical orientation of atoms in the case of two simultaneously occurring processes, viz., an elastic process (spin exchange) and an inelastic process (chemi-ionization), are derived. The rate constants of these processes are determined experimentally: $C_{se}=(2.8\pm0.8)\times10^{-9}$ cm^{3}s^{-1}, $C_{ci}=(1.0\pm0.3)\times10^{-9}$ cm^{3}s^{-1}.
 [S1063-7842(99)00909-5]

INTRODUCTION

Investigations of the interaction of excited polarized atoms of inert gases with ground-state alkali-metal atoms have been carried out for a fairly long time. The main interest in this problem is due to the possibility of indirectly polarizing one of the participants in a collision by the other, if the latter is preliminarily optically oriented by resonant optical radiation.[1–3] Orientation of a collision partner can occur in the situation under consideration as a result of two processes, viz., spin exchange and chemi-ionization. It was shown in Ref. 4 that these two processes have significant influences on one another. This must be taken into account both in performing calculations and in setting up the corresponding experiments.

The recent heightened interest in systems involving alkali-metal atoms is due to numerous reasons, both purely scientific and practical. The latter include, in particular, the use of indirect optical orientation in alkali-metal-atom–inert-gas-atom systems to obtain polarized atoms of nuclear paramagnets (primarily ^{3}He and ^{129}Xe atoms) for subsequent use in medicine, particularly in nuclear tomography,[5,6] as well as the creation and successful employment of alkali-metal–helium magnetometers with indirect optical orientation of the atoms.[3] Despite the interest in such systems, some important physical data, such as the cross sections of the respective processes, have been studied to an extremely small extent. Moreover, while experimental data on the ionization cross sections have already been reported in the literature (mainly from experiments on decaying plasmas), there are no data at all on the spin-exchange cross sections. This is because it is difficult to distinguish between the two simultaneously occurring spin-dependent processes in experiments with polarized particles, which have traditionally been used to measure spin-exchange rate constants.[4]

In the present work a successful attempt was made to solve this problem, and the spin-exchange cross section was measured for collisions of excited metastable helium atoms with ground-state cesium atoms. It was possible to accomplish this in the framework of experiments on the optical orientation and magnetic resonance of metastable helium atoms.

EQUATION OF THE EVOLUTION OF QUANTITIES OBSERVED IN THE He*(2^3S_1) – Cs($6^2S_{1/2}$) SYSTEM

When excited metastable helium atoms in the 2^3S_1 state collide with cesium atoms in the ground $6^2S_{1/2}$ state, ionization of the cesium atoms occurs because of the large internal excitation energy of the metastable helium atoms ($E=19.82$ eV). In addition, if one of the collision partners is preliminary polarized in some way, this polarization can be imparted to the other collision partner, if the total spin of the system is maintained during the collision, i.e., the following picture is observed:

$$\mathrm{He}(1^1S_0)+\mathrm{Cs}^+(2^1S_0)+e^-(m_3),$$

$$\mathrm{He}(2^3S_1,m_1)+\mathrm{Cs}(6^2S_{1/2},m_2)\rightarrow\mathrm{HeCs}^++e^-(m_3). \quad (1)$$

Process (1) takes place with the rate $1/\tau_2^i=C_{ci}N_i$, where C_{ci} is the chemi-ionization rate constant and N_i is the concentration of metastable helium atoms or ground-state cesium atoms. Reaction (1) is allowed when the total spin of the reactants m_1+m_2 is equal to the total spin of the reaction products m_3 (here m_i is the projection of the electron spin of the respective particle). Thus, if the cesium atoms are initially polarized (let the projection of their spin onto the preferred direction be $m_2=1/2$), the predominant polarization of the helium atoms as a result of a collision will be $m_1=1$.

The second process, which occurs simultaneously with chemi-ionization in the case of collisions of polarized particles, is spin exchange:

$$\mathrm{He}(2^3S_1,m_1)+\mathrm{Cs}(6^2S_{1/2},m_2)\rightarrow\mathrm{He}(2^3S_1,m_1')+\mathrm{Cs}(6^2S_{1/2},m_2'). \quad (2)$$

Here $m_1+m_2=m_1'+m_2'$ if the total spin of the system is conserved, and the rate of the process is $1/\tau_3^i=C_{se}N_i$, where C_{se} is the spin-exchange rate constant.

In this paper we consider the interaction of cesium-133 atoms with atoms of the ^{4}He isotope, which are diamagnetic in the ground state and exhibit electron paramagnetism in the

metastable state (the spin $\mathbf{S}_1=\mathbf{1}$). In the ground-state cesium atoms the electron spin $\mathbf{S}_2=\mathbf{1/2}$, and the nuclear spin $\mathbf{I}=\mathbf{7/2}$.

Since the collision time of the particles ($t_{\text{coll}}\sim 10^{-12}$ s) is much shorter than the hyperfine interaction time (for Cs atoms $\tau_{\text{hf}}\sim 10^{-10}$ s), we henceforth assume that evolution of the electronic orientation takes place during a collision, while the nuclear orientation varies in the time interval between collisions. The evolution of the longitudinal components of the quantities observed in the experiment, viz., the electronic orientation of the metastable helium atoms $\langle \mathbf{S}_{\text{He}}\rangle$, the electronic orientation of the Cs atoms $\langle \mathbf{S}_{\text{Cs}}\rangle$, and the electronic alignment of the metastable helium atoms $\langle \mathbf{Q}_{\text{He}}\rangle$ takes the following form [here we assume that $\langle \mathbf{S}\rangle=\text{Tr}(\rho\mathbf{S})$ and the components of the electronic alignment have the form $\langle \mathbf{Q}_{\text{He}}\rangle=\text{Tr}(\rho\mathbf{Q})$, where $Q_{\alpha\beta}=3/2(S_\alpha S_\beta+S_\beta S_\alpha)-\mathbf{S}^2\delta_{\alpha\beta}$, $\alpha,\beta=x,y$, or z, and ρ is the density matrix]:

$$\frac{d}{dt}\langle S_{\text{He}}\rangle^z=-\left(\frac{1}{3\tau_2}+\frac{1}{2\tau_3}\right)\langle S_{\text{He}}\rangle^z+\left(\frac{4}{9\tau_2}+\frac{4}{3\tau_3}\right)\langle S_{\text{Cs}}\rangle^z+\left(\frac{2}{9\tau_2}-\frac{1}{3\tau_2}\right)\langle S_{\text{He}}\rangle^z\langle Q_{\text{He}}\rangle^{zz}, \tag{3}$$

$$\frac{d}{dt}\langle S_{\text{Cs}}\rangle^z=-\left(\frac{1}{3\tau_2'}+\frac{1}{2\tau_3'}\right)\langle S_{\text{He}}\rangle^z+\left(\frac{1}{3\tau_2'}+\frac{4}{3\tau_3'}\right)\langle S_{\text{Cs}}\rangle^z+\frac{1}{3\tau_3'}\langle S_{\text{Cs}}\rangle^z\langle Q_{\text{He}}\rangle^{zz}, \tag{4}$$

$$\frac{d}{dt}\langle Q_{\text{He}}\rangle^{zz}=-\left(\frac{1}{3\tau_2}+\frac{3}{2\tau_3}\right)\langle Q_{\text{He}}\rangle^{zz}+\left(\frac{2}{3\tau_2}+\frac{1}{\tau_3}\right)\langle S_{\text{Cs}}\rangle^z\langle S_{\text{He}}\rangle^z. \tag{5}$$

The influence of the pump radiation can next be taken into account by introducing the respective terms into Eqs. (3)–(5): φ_p, φ_p', and φ_p'' ($\varphi_p=L_p/\tau_P$, τ_p is the pumping time for orientation or alignment, and L_P characterizes the radiation used). In addition, the contributions from diffusion and other relaxation processes which influence the evolution of the corresponding polarization moments can be taken into account in the equations. In matrix notation the system of equations takes the form

$$\frac{d}{dt}\begin{bmatrix}\langle S_{\text{He}}\rangle^z\\ \langle S_{\text{Cs}}\rangle^z\\ \langle Q_{\text{He}}\rangle^{zz}\end{bmatrix}=[A]\begin{bmatrix}\langle S_{\text{He}}\rangle^z\\ \langle S_{\text{Cs}}\rangle^z\\ \langle Q_{\text{He}}\rangle^{zz}\end{bmatrix}+[B]\begin{bmatrix}\langle S_{\text{He}}\rangle^z\langle Q_{\text{He}}\rangle^{zz}\\ \langle S_{\text{Cs}}\rangle^z\langle Q_{\text{He}}\rangle^{zz}\\ \langle S_{\text{He}}\rangle^z\langle S_{\text{Cs}}\rangle^z\end{bmatrix}+[\varphi_p]+[R], \tag{6}$$

where the matrices $[A]$, $[B]$, and $[\varphi_p]$ have the forms

$$[A]=\begin{bmatrix}-\left(\frac{1}{3\tau_2}+\frac{1}{2\tau_3}\right)-\frac{1}{\tau_d}-\frac{1}{\tau_R}; & \left(\frac{4}{9\tau_2}+\frac{4}{3\tau_3}\right); & 0\\ \left(\frac{1}{6\tau_2'}+\frac{1}{2\tau_3'}\right); & -\left(\frac{1}{3\tau_2'}+\frac{4}{3\tau_3'}\right)-\frac{1}{\tau_d'}-\frac{1}{\tau_R'}; & 0\\ 0; & 0; & -\left(\frac{1}{3\tau_2}+\frac{3}{2\tau_3}\right)-\frac{1}{\tau_d''}-\frac{1}{\tau_R''}\end{bmatrix}=\begin{bmatrix}-p_1 & r_1 & 0\\ r_2 & -p_2 & 0\\ 0 & 0 & -p_3\end{bmatrix}, \tag{7}$$

$$[B]=\begin{bmatrix}\frac{2}{9\tau_2}-\frac{1}{3\tau_3}\\ \frac{1}{3\tau_3'}\\ \frac{2}{3\tau_2}+\frac{1}{\tau_3}\end{bmatrix}, \tag{8}$$

$$[\varphi_p]=\begin{bmatrix}\varphi_p\\ \varphi_p'\\ \varphi_p''\end{bmatrix}, \tag{9}$$

and the matrix $[R]$ takes into account the various relaxation processes. It should be borne in mind that in (9) the respective matrix elements are manifested under definite pumping conditions. When metastable helium atoms are optically

pumped by linearly polarized light, the latter induces only alignment, but in the case of pumping by circularly polarized helium light, both orientation and alignment are induced simultaneously in helium. When cesium atoms are optically pumped, electronic orientation appears in the system of Zeeman sublevels. Orientation and alignment subsequently appear in the system of hyperfine sublevels of the cesium atom as a result of the hyperfine interaction.

We are interested in the case where optical orientation of the metastable helium atoms takes place. As a result of collisions of the latter with ground-state cesium atoms, the orientation is redistributed between the collision partners. It follows from (3) and (5) that the magnetic-resonance linewidths of the oriented (δf_{or}) and aligned (δf_{al}) metastable helium atoms due to collisions with ground-state cesium atoms have the form

$$\pi\delta f_{or}=1/(2\tau_3)+1/(3\tau_2), \tag{10}$$

$$\pi\delta f_{al}=3/(2\tau_3)+1/(3\tau_2). \tag{11}$$

Hence it is not difficult to see that the difference between the magnetic-resonance linewidths $\pi(\delta f_{al}-\delta f_{or})$ is equal to the rate of spin exchange between the helium and cesium atoms ($1/\tau_3=C_{se}N_{Cs}$). Thus, there is a real possibility for determining both the spin-exchange cross section and the chemi-ionization cross section from experiments on the optical orientation of atoms in the He–Cs mixture [bearing in mind that the cross section of the process equals $\langle\sigma(T)\rangle=C(T)/\langle v\rangle$, where $\langle v\rangle$ is the mean relative collision frequency of the particles].

MAGNETIC RESONANCE IN A SYSTEM OF OPTICALLY ORIENTED HELIUM AND CESIUM ATOMS

The influence of magnetic resonance on the system can be taken into account by supplementing the matrix (6) with diagonal elements of the following forms: a) for orientation

$$F_1(\omega_i)=\frac{\omega_i^2\cdot 1/p_i^2}{\omega_i^2+\Delta\omega^2+1/p_i^2},$$

where $i=1$ for helium and 2 for cesium; b) for alignment

$$F_2(\omega_1)=1/p_3\frac{3\omega_1^2(\omega_1^2+4\Delta\omega^2+1/p_3^2)}{(4\omega_1^2+4\Delta\omega^2+1/p_3^2)(\omega_1^2+\Delta\omega^2+1/p_3^2)}, \tag{12}$$

where $\omega_i=\gamma_i H_1$ is the amplitude of the resonant radio-frequency magnetic field, γ_i is the gyromagnetic ratio, $\Delta\omega=\omega-\omega_0$ is the detuning of the frequency, and $\omega_0=\gamma_i H_0$ is the resonant frequency.

Thus, the resonance in a system of oriented helium atoms, in a system of aligned helium atoms, and in a system of oriented cesium atoms can be described by introducing $F_1(\omega,\omega_1)$, $F_2(\omega,\omega_1)$, and $F_1(\omega,\omega_3)$, respectively.

When helium atoms are optically oriented by resonant radiation and magnetic resonance is excited at the frequencies of the Zeeman transitions of the helium atoms (the g factor equals 2), in the case of weak pumping we have the following results:

$$\langle S_{Cs}\rangle^z=\frac{r_2}{p_2}\langle S_{He}\rangle^z,$$

$$\langle S_{He}\rangle^z=\frac{\varphi_p p_2}{d}\left[1-(2\omega_1)^2\frac{p_1p_2/d}{p_1^2+\Delta\omega^2+(2\omega_1)^2p_1p_2/d}\right], \tag{13}$$

where

$$d=p_1p_2-r_1r_2,\quad \Delta\omega=\omega-2\omega_0,$$

and $\Delta\omega=\omega-1/2\cdot\omega_0$.

For magnetic resonance at the frequencies of the Zeeman transitions of the cesium atoms we obtain the following expressions, taking into account that the g factors of both hyperfine states are equal in magnitude and differ only in sign ($g=\pm 1/2$):

$$\langle S_{Cs}\rangle^z=\frac{\varphi_p r_2}{d}\left(1-\left(\frac{1}{2}\omega_1\right)^2\frac{p_1p_2/d}{p_2^2+\Delta\omega^2+\omega_1^2p_1p_2/d}\right),$$

$$\Delta\omega=\omega-1/2\cdot\omega_0,$$

$$\langle S_{He}\rangle^z=\frac{\varphi_p}{p_1}\left[1+\frac{r_1r_2}{d}\left(1-(2\omega_1)^2\times\frac{p_1p_2/d}{p_2^2+\Delta\omega_1^2+(2\omega_1)^2p_1p_2/d}\right)\right]. \tag{14}$$

When necessary, it is not difficult to obtain the analogous expressions for the case of the optical orientation of cesium atoms by resonant radiation.

Using the relations obtained above, we can write out an expression for the magnetic-resonance signal of the Cs atoms, which can be observed in the form of the change in the absorption of the pump radiation for the helium atoms for light parallel to the static magnetic field:

$$S_{\parallel}=-A_0\left|\frac{\Delta N_{He*}}{N_{He*}}\right|+A_1|\Delta\langle S_{He}\rangle^z|+A_2|\Delta\langle Q_{He}\rangle^{zz}|. \tag{15}$$

Here the coefficients A_0, A_1, and A_2 are determined by the relative probabilities of the transitions between the various Zeeman sublevels of the helium 2^3S_1 and $2^3P_{0,1,2}$ states, as well as by the intensities of the D_0, D_1, and D_2 lines of the helium pumping lamp and by the absorption contour of the cell. It should be noted that the last two terms in (15) are dominant, since the population changes at the moment of resonance are generally small. In addition, the possible variation of the contributions of orientation and alignment to the magnetic-resonance signal due to the collision processes should be taken into account. Such variation can result in a significant deviation of the shape of the magnetic-resonance line from the Lorentzian line shape.[7,8]

EXPERIMENTAL DETERMINATION OF THE SPIN-EXCHANGE AND CHEMI-IONIZATION CROSS SECTIONS FOR THE INTERACTION OF HELIUM AND SODIUM ATOMS

In the present work we have examined the interaction of ^{133}Cs atoms with optically polarized helium atoms in the

2^3S_1 state and experimentally determined the rate constant of spin exchange during collisions of polarized metastable helium atoms with ground-state cesium atoms for the first time. As was shown above, the decay rates of the orientation $\langle S_{\mathrm{He}}\rangle^z$ and the alignment $\langle Q_{\mathrm{He}}\rangle^{zz}$ of metastable helium atoms depend on collision processes of types (1) and (2) in accordance with (10) and (11). As can be seen from (10) and (11), the contributions to magnetic-resonance linewidth for aligned and oriented helium atoms should differ from one another. Just this difference permits determination of the rate constants of the two simultaneously occurring processes.

A standard system for investigating the optical orientation of atoms was used in the experiment. The spins of the metastable helium atoms were optically oriented by circularly polarized light[1)] and were aligned by unpolarized pump radiation ($\lambda=1.08$ μm) propagating along a magnetic field $H_z\cong 35$ mOe induced by Helmholtz coils within the magnetic shield. A helium capillary lamp, in which a high-frequency (hf) discharge was excited, served as the pumping source. The metastable state was populated by exciting the hf discharge within an absorption chamber containing gaseous ^{4}He ($P=1$ Torr at $T=300$ K) and metallic cesium. The gas-discharge chamber was placed in a thermostat. The working temperature was determined by a thermistor glued onto the coolest part of the absorption chamber. Variation of the temperature in the range 293−343 K permitted regulation of the concentration of cesium atoms from 2.7×10^{10} to 2.5×10^{12} cm^{-3}. Magnetic resonance was excited in the system of Zeeman sublevels of the 2^3S_1 helium atoms by applying an amplitude-modulated radio-frequency magnetic field $H_1=h\sin\Omega t\sin\omega t$, where $\Omega/2\pi=250$ Hz, and $\omega/2\pi$, i.e., the frequency of the radio-frequency field, was ~100 kHz, perpendicularly to the static magnetic field H_z. In the experiment the variation (at the modulation frequency Ω) of the intensity of the pump radiation passing through the absorption chamber was recorded with scanning of the frequency ω in the region of the magnetic-resonance frequency $\omega_0=\gamma H_z$ ($\gamma/2\pi=2.8$ MHz/Oe). At small values of the amplitude of the radio-frequency field $h\leqslant 0.1$ mOe (Ref. 7) the magnetic-resonance line shape (for both orientation and alignment of the helium atoms) was nearly Lorentzian, and the magnetic-resonance linewidth depended mainly on the collisional relaxation processes in the absorption chamber. The linewidth was determined from the full width of the magnetic-resonance signal at half-maximum.

To find the spin-exchange rate constant sought, the increment $\delta f_{\mathrm{or,\,al}}$ of the magnetic-resonance linewidth (for both oriented and aligned helium atoms) appearing in response to heating of the absorption chamber, i.e., $\delta f_{\mathrm{or,\,al}}=\Delta f_{\mathrm{or,\,al}}(T)-\Delta f_{\mathrm{or,\,al}}(T_0)$, where $\Delta f_{\mathrm{or,\,al}}(T_0)$ is the linewidth at $T=300$ K (~2−3 kHz), was determined as a function of the concentration of cesium atoms in the experiment. To construct the plots of $\delta f_{\mathrm{or,\,al}}$ versus the concentration of Cs atoms (N_{Cs}) the temperature was converted into the value of N_{Cs} in accordance with the data in Ref. 9.

Several series of measurements for determining $\delta f(N_{\mathrm{Cs}})$ both for orientation and for alignment were carried out in the experiment. The data obtained were averaged both within each series and for the entire list of series. An example of the plots of $\delta f_{\mathrm{or,\,al}}(N_{\mathrm{Cs}})$ obtained for one of the series of experimental measurements is presented in Fig. 1. As is seen, the contributions of the oriented (*2*) and aligned (*1*) atoms to the magnetic-resonance linewidth differ significantly.[2)] In fact, it follows from (3) and (4) that

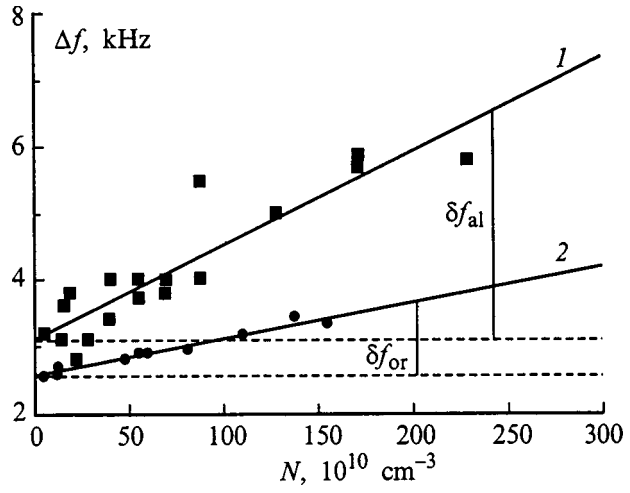

FIG. 1. Dependence of the magnetic-resonance linewidth of oriented (*2*) and aligned (*1*) metastable helium atoms on the concentration of ground-state cesium atoms (N_{Cs}). Here $\delta f_{\mathrm{or,\,al}}$ is the broadening of the magnetic-resonance line caused by the spin-exchange and chemi-ionization processes for oriented and aligned atoms.

$$\pi(\delta f_{\mathrm{al}}-\delta f_{\mathrm{or}})=N_{\mathrm{Cs}}C_{\mathrm{se}}. \tag{16}$$

The value of C_{se} determined from (16) was $(2.8\pm0.8)\times10^{-9}$ cm^{3}s^{-1}. The value of C_{se} obtained was taken into account in determining the chemi-ionization rate constant from (10) and (11): $C_{\mathrm{ci}}=(1.0\pm0.3)\times10^{-9}$cm^{3}s^{-1}. It should be noted that the dispersion of these measurements was caused mainly by the error associated with the determination of the concentration of alkali-metal atoms. Both the presence of temperature gradients on the surface of the chamber and the variability of the measurements of the working temperature contributed to this error.

In conclusion, it should be noted that while the spin-exchange rate constant was determined in the present work for the first time, the chemi-ionization rate constant has been determined repeatedly without consideration of the statistical multiplier $(1/3C_{\mathrm{ci}})$ in previous studies: $(0.44\pm0.16)\times10^{-9}$ cm^{3}s^{-1} (Ref. 9), $(0.45\pm0.09)\times10^{-9}$cm^{3}s^{-1} (Ref. 10). As can be seen from the results presented, satisfactory agreement is observed between our results and the results obtained in Refs. 9 and 10.

[1)]Pumping by circularly polarized radiation can bring about both orientation and alignment. The relative amount of alignment in the magnetic-resonance signal depends on the emission spectrum of the lamp and the length of the optical path in the absorption chamber. In our experiment this amount did not exceed 10%.

[2)]It should be noted that the width of the magnetic-resonance line of polarized helium atoms also depends on the rates of diffusion, spin exchange with electrons, etc. In the present experiment the variation (with increasing temperature) of the magnetic-resonance linewidth due to these processes did not exceed the measurement error.

[1] W. Happer, Rev. Mod. Phys. **44**, 169 (1972).
[2] S. P. Dmitriev and D. E. Denisov, Zh. Tekh. Fiz. **67**(6), 131 (1997) [Tech. Phys. **42**, 704 (1997)].
[3] E. V. Blinov, R. A. Zhitnikov, and P. P. Kuleshov, in *Optical Orientation of Atoms and Molecules. A Collection of Scientific Papers* [in Russian], VSOOAM, Leningrad (1987), pp. 12–16.
[4] G. V. Klement'ev, V. D. Mel'nikov, and V. A. Kartoshkin, Khim. Fiz. **4**(1), 37 (1985).
[5] H. Middleton, R. D. Black, B. Saam *et al.*, Magn. Reson. Med. **33**, 271 (1995).
[6] M. Gatzke, G. Cates, B. Driehuys *et al.*, Nature **370**, 199 (1994).
[7] V. A. Kartoshkin and G. V. Klement'ev, Opt. Spektrosk. **65**, 792 (1988) [Opt. Spectrosc. **65**, 468 (1988)].
[8] I. Langmuir and J. B. Taylor, Phys. Rev. **51**, 753 (1937).
[9] Yu. A. Tolmachev and D. Fogel', Opt. Spektrosk. **48**, 818 (1980) [Opt. Spectrosc. **48**, 451 (1980)].
[10] C. E. Johnson, C. A. Tipton, and P. G. Robinson, J. Phys. B **11**, 927 (1978).

Translated by P. Shelnitz

Stereodynamics of photochemical reactions

O. S. Vasyutinskiĭ

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 41–45 (September 1999)

The latest results obtained from investigating vector correlations in the photodissociation of di- and triatomic molecules are reviewed. Expressions for the polarization moments of atoms formed as a result of the photodissociation of an arbitrary molecule by polarized radiation, which depend on spatial and spin variables, are presented. The set of anisotropy parameters, which can be determined experimentally and then utilized to obtain a complete quantum-mechanical description of the primary photodissociation act, is discussed. Two methods for determining the orientation parameters, alignment parameters, and higher polarization moments of atomic photofragments are considered. They are based on a) ion imaging, which involves laser photoionization followed by mass-spectrometric detection of the photofragments and b) sub-Doppler laser polarization spectroscopy of the photofragments. The results of an experimental study of the photodissociation of Cl_2 molecules, which leads to the formation of $Cl(^2P_{3/2})$ atoms with aligned electronic angular momenta, are presented as an example. [S1063-7842(99)01009-0]

INTRODUCTION

The photodissociation of small molecules has attracted the fixed attention of researchers already for many decades,[1,2] since such processes play a major role in the physics and chemistry of the atmosphere of the Earth and planets, as well as in gas lasers and industrial systems. The nature of an elementary photochemical act is difficult to study, because there are several interdependent reaction channels under typical conditions, and the processes taking place are characterized by a picosecond time scale. Considerable progress has recently been observed in this area due to the development of new theoretical models and a consequent clearer understanding of the physical processes taking place and to the introduction of new experimental methods.

At the present time it is possible to theoretically describe and experimentally investigate the photochemical reactions which occur in small atmospheric molecules with consideration of all the possible spatial and spin degrees of freedom of the photofragments formed, as well as to closely approach the possibility of performing and interpreting a complete quantum-mechanical experiment.

The concept of vector correlations in the photodissociation of molecules has played a major role in developing the current theories on the nature of an elementary photochemical act. Attention was first focused on the correlation between the motion direction vector **k** of the photofragments formed as a result of the photodissociation of a diatomic molecule and the polarization vector **e** of the dissociating radiation more than 30 years ago in Ref. 3. Correlations between the directions of the electronic angular momentum vector **j** of a photofragment and the polarization vector **e** of the dissociating radiation were discovered in Refs. 4–6. Other types of vector correlations associated with the photodissociation of molecules have also been investigated (see, for example, Refs. 7 and 8). The current stage of the research in this area is based on a quantum-mechanical theory of photodissociation which takes into account all the possible types of vector correlations.[9] The predominantly employed experimental methods of study are ion imaging[10] and sub-Doppler laser polarization spectroscopy, and a great deal of attention has been attracted by studies of the orientation and alignment of the electronic angular momenta of atomic photofragments, which provide the most detailed information on the processes taking place.[11–20]

The purpose of this paper is to review and analyze the latest results obtained from investigating the vector correlations associated with the photodissociation of small molecules.

THEORETICAL TREATMENT OF CORRELATIONS BETWEEN THE SPATIAL RECOIL ANISOTROPY OF PHOTOFRAGMENTS AND THEIR ANGULAR MOMENTUM DISTRIBUTION

A general expression for the photodissociation cross section of an arbitrary molecule by polarized radiation, which includes a dependence on the vector **k** and on all the other spatial and spin variables of the photofragments formed in the ''instantaneous'' photodissociation approximation,[21] was obtained in Ref. 9. This expression includes a kinematic part, which is characterized by the vectors **e** and **k**, as well as a dynamic part, which is characterized by the dynamic functions $f_K(q,q')$, where K is the rank of the polarization moment (the irreducible statistical tensor[22]) of the photofragment, and q and q' label cyclic projections of the electric dipole moment of the molecule onto its axis. It was shown in Ref. 9 that the kinematic part for each polarization moment can be represented in a universal form, which is valid for fast photodissociation processes of small molecules. At the same

time, the values of the dynamic functions are different for different molecules and different experimental conditions and can be determined theoretically in a general form when the corresponding multichannel scattering problem is solved (see, for example, Ref. 21). The computational difficulties arising in this case are usually very great even for simple molecules, since it necessary to take into account diverse physical processes taking place on different portions of the potential surfaces, such as the simultaneous coherent optical populating of two or more excited states of a molecule from its ground state, nonadiabatic transitions between different potential surfaces, rotation of the molecular axis, the interaction between the photofragments formed at large distances between them, etc.

If the photodissociation radiation is linearly polarized along the Z axis, and the angles Θ and ϕ characterize the escape direction of the photofragments, the angular distribution for the number of photofragments (the rank zero statistical tensor[22]) can be represented in the following form:[3]

$$\rho_{00}=\frac{1}{4\pi\sqrt{2j+1}}[1+\beta_0 P_2\ (\cos\ \Theta)], \tag{1}$$

where j is the total angular momentum of the photofragment, and β_0 is the anisotropy parameter, which takes values in the interval $\beta_0=-1\ldots 2$ in accordance between the ratio between the probabilities of transverse and longitudinal optical transitions in the molecule.

The angular distribution for the dipole moment of the photofragments (the rank one statistical tensor ρ_{1Q} or the orientation[22]) can be represented in the form[9,15,23]

$$\rho_{10}=\frac{3\sqrt{3}}{4\pi\sqrt{2j+1}}\left[\alpha_1\ \cos^2\ \Theta+\frac{\gamma_1}{2}\ \sin^2\ \Theta\right], \tag{2a}$$

$$\rho_{11}=-\rho^*_{1-1}=-\frac{3\sqrt{3}}{4\pi\sqrt{2(2j+1)}}\times\exp(i\phi)\sin\ \Theta\ \cos\ \Theta\left[\alpha_1-\frac{\gamma_1}{2}-i\frac{\gamma_1'}{2}\right]. \tag{2b}$$

Here ρ_{10} is a diagonal element of the statistical tensor, and ρ_{11} and ρ_{1-1} are off-diagonal elements. It was assumed in deriving formulas (2) that the photodissociation radiation is right-circularly polarized and propagates along the Z axis. The three orientation parameters α_1, γ_1, and γ_1' in (2b) are normalized linear combinations of dynamic functions of rank $K=1$ with different values of q and q' (Ref. 15). Each of the terms on the right-hand side of (2) corresponds to a definite mechanism for the appearance of orientation of the angular momenta of the photofragments in the photodissociation reaction. For example, α_1 characterizes photodissociation resulting from perpendicular ($\Sigma\rightarrow\Pi$) optical transitions in the molecule, and γ_1 characterizes the photodissociation resulting from the coherent excitation of two states of different symmetry by a single photon ($\Sigma\rightarrow\Sigma,\Pi$). We note that parallel ($\Sigma\rightarrow\Sigma$) optical transitions in a molecule cannot lead to orientation of the angular momenta of the photofragments.[24]

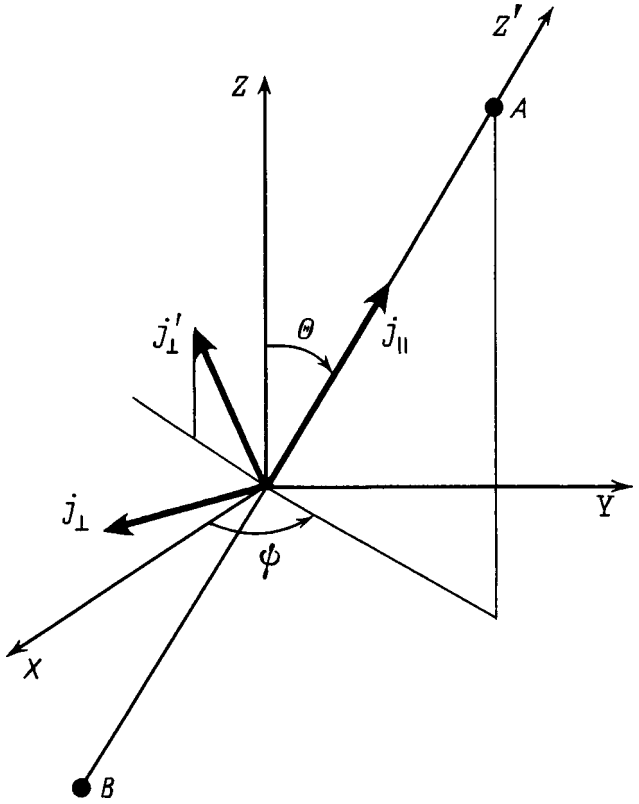

FIG. 1. Vector model of the orientation of the angular momenta of photofragments.[9] The components $j_\parallel$, $j_\perp$, and $j_\perp'$ of the total angular momentum of the photofragment are mutually perpendicular to one another.

The term proportional to the third orientation parameter γ_1' makes a contribution only to the off-diagonal elements of the statistical tensor (2b) and also describes dissociation involving the coherent excitation of Σ and Π states in the molecule.

The contributions of all three mechanisms described to the orientation of the angular momentum of a photofragment have a simple physical meaning and can be interpreted graphically using the vector model in the molecule coordinate frame shown in Fig. 1. It is seen that the angular momentum of a photofragment formed as a result of a perpendicular transition in a molecule is directed along the recoil axis, while the angular momenta formed as a result of the two possible coherent-excitation mechanisms are perpendicular to the recoil axis and to one another. All three orientation parameters can be determined directly from an experiment by investigating the angular dependences of the orientation of the electronic moments of the photofragments.

The angular distribution for the quadrupole moment of photofragments (the rank two statistical tensor ρ_{2Q} or the alignment[22]) was obtained in Refs. 9 and 16. For example, for the case where the photodissociation radiation is linearly polarized along the Z axis, the expression for the diagonal component of this statistical tensor can be represented in the form

$$\rho_{20}(\Theta,\phi)=\frac{\sqrt{5}V(j)}{4\pi\sqrt{2j+1}}\left\{P_2(\cos\ \Theta)\times[s_2-2\alpha_2 P_2(\cos\ \Theta)]-3\gamma_2\sin^2\times\Theta\cos^2\ \Theta-\frac{3}{4}\eta_2\ \sin^4\ \Theta\right\}, \tag{3}$$

where

$$V(j)=5\left[\frac{j(j+1)}{(2j+3)(2j-1)}\right]^{1/2}.$$

Here s_2, α_2, γ_2, and η_2 are four alignment parameters, which are normalized linear combinations of dynamic functions of rank $K=2$ with different values of q and q' (Ref. 15). Another (fifth) alignment parameter γ_2' appears only in the expression for the off-diagonal components of the statistical tensor. Each of the four terms on the right-hand side of (3) describes a definite mechanism for the appearance of alignment in a photodissociation reaction. For example, the terms with the alignment parameters s_2 and α_2 describe the contribution of the incoherent optical excitation of a molecule by means of perpendicular and longitudinal transitions, while the terms with the alignment parameters γ_2 and η_2 describe the contribution of the simultaneous coherent excitation of two states of different symmetry ($\Sigma\rightarrow\Sigma,\Pi$) and of two states of the same symmetry ($\Sigma\rightarrow\Pi,\Pi$), respectively. All five possible alignment parameters can be determined from an experiment by investigating the angular dependences of the alignment of the electronic moments of the photofragments.

The theory also predicts the existence of polarization moments of the photofragments of higher orders with ranks $K>2$ in photodissociation reactions. In principle, these quantities can also be determined in an experiment, but their contribution to the total signal is always considerably smaller than the contribution of the moments of lower rank. Consequently, the angular dependences of the high-rank polarization moments have not yet been observed experimentally. The set of anisotropy parameters with ranks from $K=0$ to $K=2j$ make up a complete set of physical observables characterizing an elementary photodissociation act.

EXPERIMENTAL METHOD AND RESULTS OBTAINED

Two models, which are based on ion imaging[10,16] and sub-Doppler laser polarization spectroscopy[11,14,15] were recently proposed for the purpose of experimentally investigating the spatial distributions of the orientation, alignment, and high-rank polarization moments in photodissociation reactions.

In a typical experimental setup which utilizes ion imaging two tunable lasers are employed for photodissociation of the molecules under investigation and detection of the photofragments formed, and their rays are directed toward one another, as shown in Fig. 2 (Ref. 16). The photofragments are detected by means of their multiphoton resonant ionization by the output of the probe laser followed by mass-spectrometric recording of the ions formed using a time-of-flight mass spectrometer. After traversing the drift chamber, the ion cloud is visualized using a photodetector consisting of a microchannel detector, a phosphorescent screen, and a video camera. The variation of the polarization of the output of the probe laser and the photodissociation laser makes it possible to determine the angular dependences of both the number of photofragments and their polarization moments.

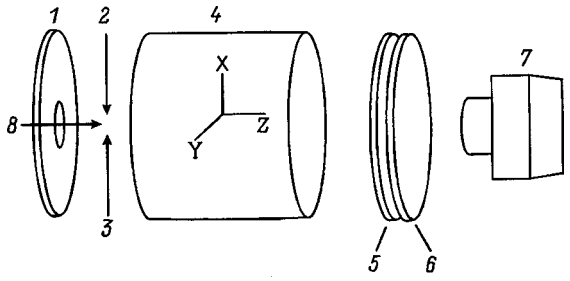

FIG. 2. Typical setup for an ion-imaging experiment: *1* — diaphragm, *2* — photodissociation laser beam, *3* — probe laser beam, *4* — ion drift chamber; *5* — microchannel detector, *6* — phosphorescent screen, *7* — video camera, *8* — beam of molecules investigated.

One of the experimental signals obtained by the method just described in Ref. 17 is shown in Fig. 3. In that work the photodissociation of the chlorine molecule into atoms in the ground $^2P^0_{3/2}$ state under the effect of linearly polarized radiation with a wavelength of 355 nm was investigated. The atoms were detected by the linearly polarized output of a probe laser with a wavelength of 234 nm by means of $2+1$ resonant ionization through the $^2D^0_{3/2}$ excited state. The Z axis in Fig. 3 is directed toward the mass spectrometer, and the four ion images correspond to four mutual combinations of the directions of the polarization vectors of the photodissociation radiation and the probe radiation.

As can be seen from Fig. 3, the recoil of chlorine atoms was observed predominantly in an axially symmetric pattern relative to the polarization vector of the photodissociation radiation in directions perpendicular to it. The corresponding angular distribution can be described by Eq. (1) with an an-

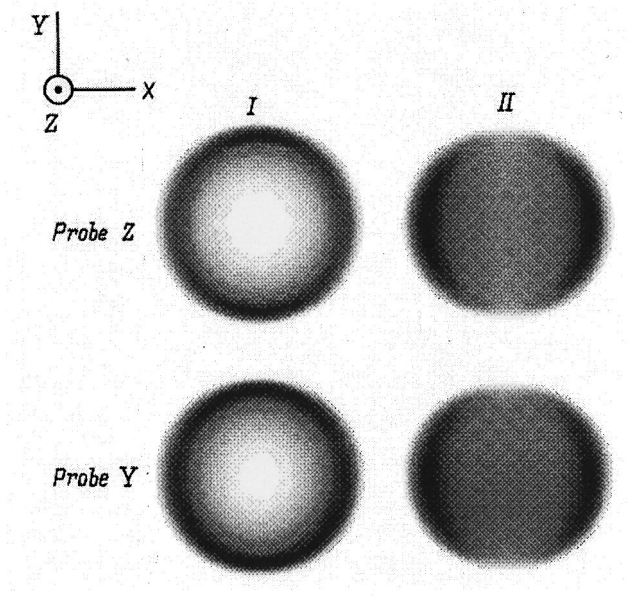

FIG. 3. Ion images obtained by ion imaging:[17] *I*, *II* — direction of the polarization vector of the photodissociation radiation along the Z axis and along the Y axis, respectively; the upper and lower rows correspond to the directions of the polarization vector of the probe radiation indicated in the figure.

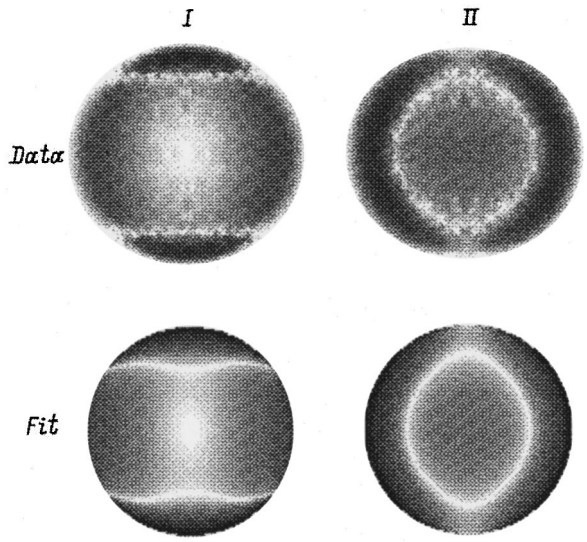

FIG. 4. Angular dependence of the alignment of chlorine atoms:[17] upper row — experiment, lower row — theory; *I*, *II* — same as in Fig. 3.

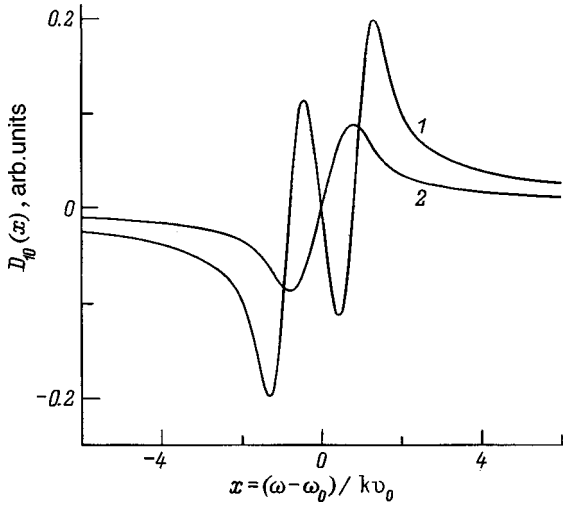

FIG. 5. Spectral dependences of the birefringence signal of the probe radiation after passing through vapors of oriented photofragments:[15] *1* — incoherent excitation of the $\Sigma\rightarrow\Pi$ transition in a dissociating molecule, *2* — coherent excitation of two states of Σ and Π symmetry; x is the relative detuning of the frequency of the probe radiation ω relative to the center of the absorption line of the photofragments ω_0, k is the wave vector, and v_0 is the absolute value of the recoil velocity of the photofragments.

isotropy parameter β_0 close to $\beta_0\approx -1$, which characterizes the perpendicular $X^1\Sigma_{0g}\rightarrow C^1\Pi_{1u}$ optical transition in the chlorine molecule.

In addition, a difference between the images corresponding to different directions of polarization of the probe laser output can be seen in Fig. 3. This difference is caused by the alignment of the electronic angular momenta of the chlorine atoms. Separating the contribution of the alignment to the total experimental signal by subtracting the images shown in the first and second rows in Fig. 3 from one another was proposed in Refs. 16 and 17. The corresponding difference images are shown in Fig. 4, in which the experimental signals are presented in the first row, and the theoretical results are presented in the second row. An analysis of these images permitted discovery of the existence of two different mechanisms for the photodissociation of Cl_2 molecules and determination of the anisotropy parameters α_2 and η_2.

In sub-Doppler laser polarization spectroscopy the photofragments are detected by a narrow-band tunable laser, whose frequency is scanned along the absorption line of the photofragments under investigation, and the total dichroism[11] or birefringence[14,15] signal of the probe radiation is recorded. The information obtained as a result on the angular dependences of the polarization moments of the photofragments is similar as a whole to the information obtained by ion imaging, but in simple cases sub-Doppler laser polarization spectroscopy is simpler to implement and can provide a better signal-to-noise ratio. As an example of the possible interpretation of the signals obtained, Fig. 5 presents the theoretical spectral dependences of the birefringence signal of linearly polarized probe radiation which has passed through a pair of spin-oriented photofragments for two different photodissociation mechanisms.[15] As is seen from this figure, the spectral dependence of the birefringence signal caused by incoherent transverse excitation of a dissociating molecule differs significantly from the spectral dependence of the signal caused by coherent excitation of the same molecule, making it possible to experimentally distinguish between these mechanisms.

Thus, as a result of the investigations performed, new methods for investigating photochemical reactions of practical importance have been proposed and approved, and new data on the primary photodissociation act of small molecules have been obtained.

This work was carried out with support from the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF) (Grant No. RP1-223) and the Russian Fund for Fundamental Research (RFBR) (Grant No. 93-02-18313).

[1] H. Okabe, *Photochemistry of Small Molecules* [Wiley, New York (1978); Mir, Moscow (1981), 500 pp.].
[2] A. M. Pravilov, *Photoprocesses in Molecular Gases* [in Russian], Moscow (1992), 350 pp.
[3] R. N. Zare and D. R. Herschbach, Proc. IEEE **51**, 173 (1963).
[4] R. J. Van Brunt and R. N. Zare, J. Chem. Phys. **48**, 4304 (1968).
[5] E. W. Rothe, U. Krause, and R. Duren, Chem. Phys. Lett. **72**, 100 (1980).
[6] O. S. Vasyutinskiĭ, JETP Lett. **31**, 428 (1980).
[7] J. P. Simons, J. Phys. Chem. **91**, 5378 (1987).
[8] P. L. Houston, J. Phys. Chem. **91**, 5388 (1987).
[9] L. D. A. Siebbeles, O. S. Vasyutinskii, M. Glass-Maujean *et al.*, J. Phys. Chem. **100**, 3610 (1994).
[10] A. J. R. Heck and D. W. Chandler, Annu. Rev. Phys. Chem. **46**, 335 (1995).
[11] Y. Wang, H.-P. Loock, J. Cao, and C. X. W. Qian, J. Chem. Phys. **102**, 808 (1995).
[12] Y. Mo, H. Katayanagi, M. C. Heaven, and T. Suzuki, Phys. Rev. Lett. **77**, 830 (1996).
[13] T. Suzuki, H. Katayanagi, Y. Mo, and K. Tonokura, Chem. Phys. Lett. **256**, 90 (1996).
[14] D. V. Kupriyanov, I. M. Sokolov, and S. V. Subbotin, Opt. Spektrosk. **80**, 728 (1996) [Opt. Spectrosc. **80**, 653 (1996)].
[15] B. V. Picheyev, A. G. Smolin, and O. S. Vasyutinskii, J. Phys. Chem. **101**, 7614 (1997).
[16] M. Ahmed, D. S. Peterk, O. S. Vasyutinskii, and A. G. Suits, J. Chem. Phys. **110**, 4115 (1999).

[17] A. S. Bracker, E. R. Wouters, A. G. Suits *et al.*, Phys. Rev. Lett. **80**, 1626 (1998).
[18] A. T. J. B. Eppink, D. H. Parker, M. H. M. Janssen *et al.*, J. Chem. Phys. **108**, 1305 (1998).
[19] T. P. Rakitzis, S. A. Kandel, and R. N. Zare, J. Chem. Phys. **108**, 8291 (1998).
[20] T. P. Rakitzis, S. A. Kandel, A. J. Alexander *et al.*, Science **281**, 1346 (1998).
[21] G. G. Balint-Kurti and M. Shapiro, in *Photodissociation and Photoionization*, edited by K. P. Lawley, Wiley, New York (1985), p. 403.
[22] E. B. Alexandrov, M. P. Chaika, and G. I. Khvostenko, *Interference of Atomic States* [Springer-Verlag, Berlin–New York (1993); Nauka, Moscow (1991), 255 pp.].
[23] O. S. Vasyutinskiĭ, Opt. Spektrosk. **54**, 885 (1983) [Opt. Spectrosc. **54**, 524 (1983)].
[24] O. S. Vasyutinskiĭ, Zh. Éksp. Teor. Fiz. **81**, 1608 (1981) [Sov. Phys. JETP **54**, 855 (1981)].

Translated by P. Shelnitz

Numerical simulation of gasdynamic phenomena

Yu. P. Golovachev, A. I. Zhmakin, and A. A. Shmidt

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 46–49 (September 1999)

The work in computational hydrodynamics related to investigations of the supersonic streamlining of bodies, the interaction of shock waves, multiphase flows, and hyposonic flows in devices for semiconductor technology is reviewed.
[S1063-7842(99)01109-5]

INTRODUCTION

Owing to the impressive progress in creating computer systems and advances in the development of methods for numerically solving nonlinear multidimensional problems, numerical simulation has become one of the principal methods for studying complex gasdynamic phenomena. The mathematical models of gasdynamic phenomena are highly complicated. Moreover, the flows considered are usually accompanied by diverse physicochemical processes with vastly different temporal and spatial scales. This imposes rigid requirements on numerical simulation methods. The solution of the problems under consideration with acceptable expenditures of computer resources requires, first, the use of the optimal mathematical and physical models and, second, efficient computational algorithms, which have a high resolving power. An important element of numerical simulation is the construction of computation meshes. In our studies we employ movable meshes, which adapt to the structure of the gasdynamic field. The adaptation of the mesh, which provides for the resolution of fine details of the flow under investigation, is performed during the solution of the problem either by varying the location of mesh points with a fixed total number of points or by adding and removing points in appropriate portions of the computation region.

SUPERSONIC STREAMLINING OF BODIES

Investigations of the supersonic streamlining of bodies are of interest for problems in aerodynamics and aerospace technology. The current problems of this kind are characterized by a need to take into account the effects of rarefaction, viscous-inviscid interactions, flow separation and attachment, and a long list of physicochemical processes initiated by heating of the gas in intense shock waves. In addition, molecular or turbulent transport processes have a significant influence in the entire disturbed region of a flow and thus call for the use of the equations of the dynamics of a viscous gas instead of the traditional models of inviscid flow and boundary layers.

The most general mathematical model of gas flows in a continuous-medium regime is the Navier–Stokes system of equations. Along with it, simpler models of ''truncated'' and parabolized Navier–Stokes equations are employed, depending on the specific conditions.[1] Turbulent flow regimes are simulated using Reynolds averaging and semiempirical models of turbulent transport. To allow for physicochemical processes, the gasdynamic equations are supplemented with balance equations for the corresponding components of the internal energy of the gas, continuity equations for the components of the gas mixture, and a radiation transfer equation.[1]

The mathematical models used lead to initial/boundary-value problems, and effective numerical methods of second-order accuracy based on the use of implicit constant direction schemes have been developed to solve them.[1–3] In these methods the values of the functions in each new step along the marching coordinate are found using iterations, in each of which one-dimensional linear systems of difference equations with three-diagonal block matrices of coefficients are solved. When nonequilibrium physicochemical processes are taken into account, the computation algorithm includes a Newton iteration process.

The numerical simulation procedure just described has been applied to the solution of a long list of current problems in the supersonic streamlining of bodies, including hypersonic streamlining under the conditions of the motion of space vehicles in planetary atmospheres,[4,5] the highly nonstationary flows appearing in response to a rapid change in flight conditions,[6–8] the streamlining of bodies in inhomogeneities of various origin,[9–11] and separated flows. As an example, Fig. 1 presents the results of the solution of the classical problem of the formation of a flow field around a circular cylinder when a plane shock wave acts on it. The upper part of Fig. 1 shows the density isolines obtained from a computation for a single moment in time, and the lower part shows an experimental shadow photograph of such a flow, which was kindly supplied by M. P. Syshchikova. It can be seen that the calculation satisfactorily reproduces all the details of the experimentally observed shock-wave structure of the flow.

SIMULATION OF MULTIPHASE FLOWS

Heterogeneous mixtures are generally highly nonequilibrium systems, which are characterized by a large set of physicochemical processes, which take place both within each of the phases and on the phase boundaries. The investigation of multiphase flows is of interest both from the

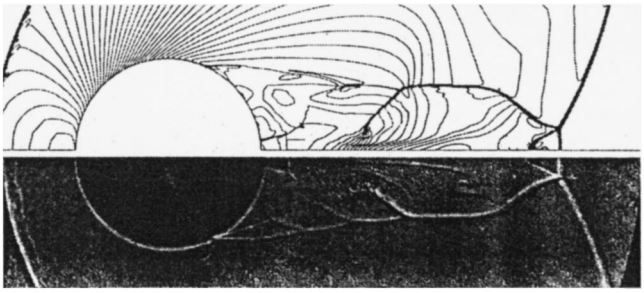

FIG. 1. Diffraction of a plane shock wave on a circular cylinder. $M_\infty = 3.0$, $\gamma = 1.29$, $\mathrm{Re} = 2.07 \times 10^5$, $t = 5.8$.

standpoint of the development of theoretical models of the mechanics of continuous media and because of the broad spectrum of applications (problems in power engineering, chemical engineering, transportation, the creation of new materials, etc.).

A novel Lagrange–Euler algorithm has been developed for numerical investigations of flows of multiphase media. It permits the analysis of each of the phases in the most natural way: the carrier (dispersion) phase is described using continual models (the Euler, Navier–Stokes, and other such models), and the dispersed phase is described by a discrete test-particle model. Such an approach has great flexibility, which permits including different models of the components of the heterogeneous mixture, physicochemical processes taking place in them, and mechanisms of interfacial transport in the treatment. An algorithm for simulating the response of a condensed medium to a pulsed high-energy disturbance, which is a combination of the finite-element method and TVD schemes, has also been developed.

The algorithm and programs developed have been used to simulate the supersonic streamlining of bodies by gas flows containing solid particles or water drops;[14] the erosion failure of bodies streamlined by supersonic sputtered gas flows;[14] the propagation of shock waves in gas suspensions, the ignition of gas suspensions, and the transition from combustion to detonation;[15] the high-speed interaction of a drop with a solid barrier or a liquid layer;[16] and the propagation of shock and detonation waves in condensed media undergoing a phase transition on the wave front.[17]

As an example, Fig. 2 presents a comparison of the results of an experimental investigation (on the left)[18] and numerical simulation (on the right) of the interaction of a cylindrical drop with a solid barrier. Density isolines are shown. The velocity of the steel projectile interacting with the lateral surface of a motionless cylindrical drop is $V = 110$ m/s, and the time between frames is 1 μs. As can be seen from the comparison presented, the algorithm developed provides a faithful picture of the propagation of shock waves and rarefaction waves, the formation of radial near-wall cumulative jets, and the appearance and development of cavitation bubbles.

SIMULATION OF HYPOSONIC FLOWS

An extensive list of viscous gas flows of practical importance are characterized by slow (compared with sonic) velocities and significant temperature variations. Such flows are realized, for example, in heat-exchange processes, in flow-type pulsed-periodic lasers, and in epitaxial reactors for growing semiconductor structures from the gas phase.

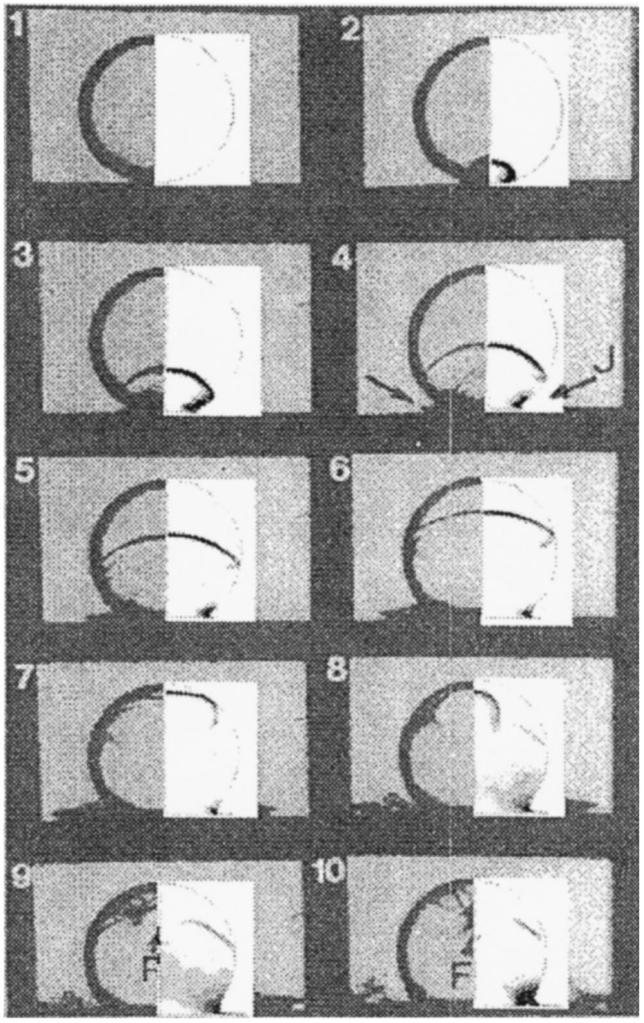

FIG. 2. Interaction of a cylindrical drop with a solid barrier. $V = 110$ m/s.

The class of problems described cannot be solved within the familiar Boussinesq approximation. At the same time, the use of the complete Navier–Stokes equations leads to formidable problems associated with the tediousness of the calculations, losses of accuracy, and the complexity of formulating the boundary conditions on the boundaries being penetrated. The most faithful model is provided by the so-called hyposonic Navier–Stokes equations, which are obtained in the limit of small values of the Mach number and the hydrostatic compressibility parameter.

Several numerical methods based on an implicit modification of the Chorin–Hart method for incompressible flows[19] and block iteration[20] have been developed for investigating hyposonic flows on structured walls. A conjugate-gradient method has been used to speed up convergence of the iterations in finding the pressure correction, and fast Fourier transformation has been used in the calculations of flows in simple regions.[21]

The growth of layers of GaAs and the solid solution

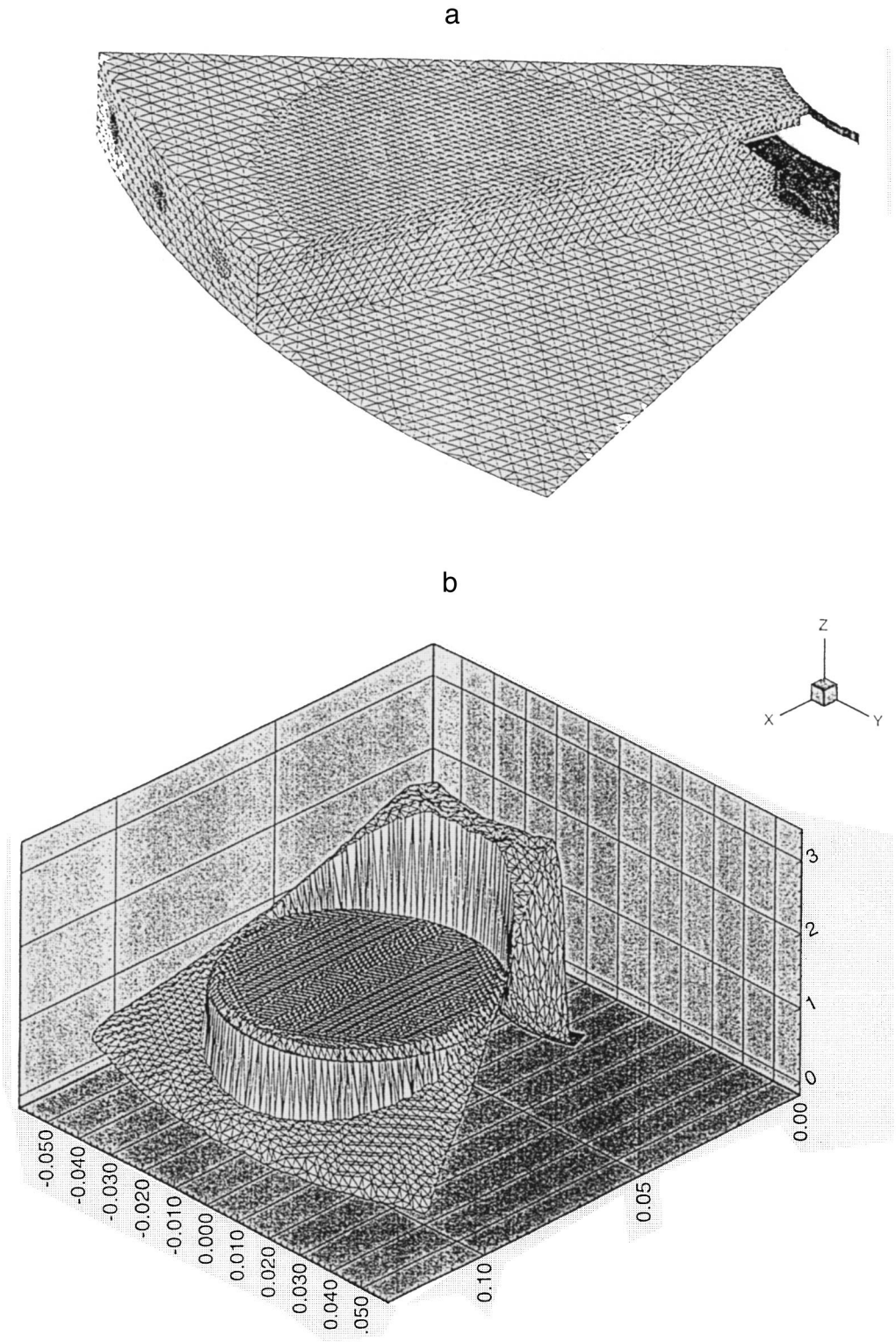

FIG. 3. Computation mesh and distribution of the growth rate of GaAs for a planetary reactor: *1–10* — successive moments in time, $\Delta\tau = 1$ μs.

AlGaAs in two- and three-dimensional formulations, nonstationary processes involved in the synthesis of heterostructures, and the possibility of controlling the composition of epitaxial layers have been investigated.[22–24] A combination of a macroscopic gasdynamic approach and a kinetic description within the Monte Carlo method has permitted simulation of growth on a profiled surface, and a modified numerical method has permitted an investigation of liquid-phase epitaxy on a moving substrate.[25]

Methods for calculating hyposonic flows on two-dimensional and three-dimensional unstructured meshes have recently been developed.[26,27] To obtain a quantitative description of processes in real industrial systems, the flow of undiluted gas mixtures is considered, and the thermal conduc-

tivity in solid blocks and radiative heat exchange are taken into account.

As an example, Fig. 3 presents the computation mesh (a) and the epitaxial growth rate of GaAs (b) for a planetary reactor with rotating substrates.

[1] Yu. P. Golovachev, *Numerical Simulation of Viscous Gas Flows in a Shock Layer* [in Russian], Fizmatlit, Moscow (1996).
[2] Yu. P. Golovachev and F. D. Popov, Zh. Vychisl. Mat. Mat. Fiz. **12**, 1292 (1972).
[3] V. L. Belousov, Yu. P. Golovachev, and V. V. Zemlyakov, Mat. Model. **6**(10), 66 (1994).
[4] Yu. P. Golovachov, Int. J. Heat Mass Transf. **4**, 649 (1981).
[5] Yu. P. Golovachev and N. V. Leont'eva, Uch. Zap. Tsentr. Aérogidrodyn. Inst. **13**(5), 40 (1982).
[6] Yu. P. Golovachev and N. V. Leont'eva, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 2, 186 (1989).
[7] V. L. Belousov, Yu. P. Golovachev, and N. V. Leont'eva, Model. Mekh. **2(19)**, No. 2, 31 (1988).
[8] D. Drikakis, D. Ofengeim, E. Timofeev, and P. Voinovich, J. Fluid Struct. **11**, 665 (1997).
[9] Yu. P. Golovachev and N. V. Leont'eva, Izv. Akad. Nauk SSSR, No. 3, 143 (1985).
[10] Yu. P. Golovachev and N. V. Leont'eva, Inzh.-fiz. Zh. **51**, 955 (1986).
[11] V. L. Belousov and Yu. P. Golovachev, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 183 (1990).
[12] Yu. P. Golovachev and N. V. Leont'eva, Zh. Tekh. Fiz. **68**(10), 20 (1998) [Tech. Phys. **43**, 1154 (1998)].
[13] Yu. P. Golovachov and N. V. Leontieva, Comput. Fluid Dyn. Rev. **5**, 439 (1997).
[14] Yu. P. Golovachov, M. S. Ramm, V. L. Belousov *et al.*, in *Gas Dynamics*, edited by Yu. P. Koptev, Nova Science, New York (1992), pp. 231–247.
[15] Yu. A. Kurakin and A. A. Shmidt, Pis'ma Zh. Tekh. Fiz. **23**(6), 26 (1997) [Tech. Phys. Lett. **23**(3), 221 (1997)].
[16] A. V. Chizhov and A. A. Shmidt, Pis'ma Zh. Tekh. Fiz. **22**(3), 57 (1996) [Tech. Phys. Lett. **22**(2), 117 (1996)].
[17] A. V. Chizhov, V. O. Chistyakov, and A. A. Shmidt, Pis'ma Zh. Tekh. Fiz. **23**(9), 33 (1997) [Tech. Phys. Lett. **23**(5), 347 (1997)].
[18] M. Lesser and J. Field, in *Proceedings of the 18th International Symposium on Shock Waves*, Sendai (1991), pp. 61–72.
[19] A. I. Zhmakin and Yu. N. Makarov, Mat. Model. **2**(7), 62 (1990).
[20] Yu. N. Makarov, E. A. Subashieva, A. N. Zabolotskikh *et al.*, in *Proceedings of the 5th International Conference*, Swansea (1993), pp. 1298–1309.
[21] L. A. Kadinskii, Yu. N. Makarov, and A. I. Zhmakin, in *Gas Dynamics*, edited by Yu. P. Koptev, Nova Science, New York (1992), pp. 261–278, 1298–1309.
[22] Yu. N. Makarov and A. I. Zhmakin, J. Cryst. Growth **94**, 537 (1989).
[23] Yu. N. Makarov and A. I. Zhmakin, Cryst. Res. Technol. **25**, 31 (1990).
[24] Yu. N. Makarov, M. S. Ramm, E. A. Subashieva, and A. I. Zhmakin, J. Cryst. Growth **145**, 271 (1994).
[25] D. Z. Garbuzov, E. V. Zhuravkevich, A. I. Zhmakin *et al.*, J. Cryst. Growth **110**, 955 (1991).
[26] Yu. N. Egorov and A. I. Zhmakin, Comput. Mater. Sci. **11**, 204 (1998).
[27] A. I. Zhmakin, Commun. Numer. Methods Eng. **13**, 219 (1997).

Translated by P. Shelnitz

Physics of spherical tokamaks

A. Sykes

EURATOM/UKAEA Fusion Association, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 50–57 (September 1999)

Spherical tokamaks are a limiting case of conventional tokamaks, combining simple design with attractive physical characteristics. Being of potential importance for the controlled nuclear fusion program in their own right, spherical tokamaks also contribute much to our understanding of the physics of conventional tokamaks. For instance, they are contributing to the modeling of confinement scalings for ITER and to research on general plasma properties such as energy confinement and MHD processes. Results obtained for the first generation of spherical prototype tokamaks are reviewed and the main trends for further research are indicated. © *1999 American Institute of Physics.* [S1063-7842(99)01209-X]

INTRODUCTION

In 1986 Peng and Strickler[1] outlined numerous potential theoretical advantages of the concept of ''spherical tori'' in which the aspect ratio $A=R/a$ is substantially smaller than that of conventional tokamaks, being close to unity. Robinson[2] suggested that, as well as exhibiting typical ''tokamak'' confinement and MHD stability because of the similar values of the safety factor q, the plasma in these ''spherical tokamaks'' will also possess various favorable characteristics inherent in spheromaks and reverse-pinch devices. Predicted advantages included the natural elongation of the plasma in the vertical direction ($k\sim 2$), high toroidal beta, and so-called ''isodynamic'' regions providing improved confinement. These predictions have now been confirmed experimentally, particularly in the START (Small Tight Aspect Ratio Tokamak) spherical tokamak at Culham Laboratory[3] and also in various other small spherical tokamaks (such as CDX-U, HIT, TS-3, MEDUSA, and ROTAMAK-ST).

These promising results have stimulated the development and construction of the next generation of spherical tokamaks, including GLOBUS-M (Russian Federation), MAST (Great Britain), PEGASUS and NSTX (US), ETE (Brazil), and TST (Japan).

The new geometry posed new problems and also served as an important test of magnetic confinement theory. Recent theoretical studies stimulated by the experimental results indicate that spherical tokamaks may have additional advantages in terms of stability and plasma confinement. The next generation of large spherical tokamaks is now being built with important specific aims: they should improve our understanding of the physics of toroidal plasma confinement systems, check out new theoretical predictions, and indicate the way to achieving high-beta compact fusion facilities having various applications such as materials testing and power production.

1. RESULTS OF EXPERIMENTS ON SPHERICAL TOKAMAKS

The first spherical tokamaks were produced by inserting a central rod into existing experimental fusion facilities such as Rotamak at Lucas Heights,[4] the Heidelberg Spheromak,[5] and the SPHEX ''Rodamak.''[6] Each of these small experiments yielded important and reliable results on plasma equilibrium and stability. For instance, it was observed in the Heidelberg Spheromak that a small current in the rod stabilizes the tilting instability, and the creation of a toroidal field in SPHEX improved the efficiency of energy transfer from the gun to the plasma, but all these were cold-plasma experiments. The first experimental verification of the concept of spherical tokamaks using a hot plasma (>100 eV) was carried out at Culham Laboratory using the small tight aspect ratio tokamak START.[3] Although START was constructed only to prove the concept of spherical tokamaks, important results have been achieved with the following typical parameters: major radius $R\sim 0.3$ m, minor radius $a\sim 0.23$ m, aspect ratio ~ 1.3, plasma elongation 1.8, plasma volume $0.5\,\mathrm{m}^3$, $\langle n_e\rangle \sim 5\times 10^{19}\,\mathrm{m}^{-3}$, vacuum toroidal field on axis (at R_0) ~ 0.3 T, and plasma current up to 310 kA.

1) Plasma equilibrium characteristics.

One of the characteristic features of spherical tokamaks is the improved efficiency of utilizing the toroidal field as the aspect ratio A decreases, because of an appreciable increase in the edge safety factor q_a and toroidal effects. This was confirmed by accurate probe measurements of the magnetic field in the low-temperature Rotamak-ST at Flinders.[7] Whereas at the center of the plasma the magnetic surfaces are similar to those in a conventional tokamak, at the plasma edge (with an extremely low aspect ratio ~ 1.1) the field lines are mainly concentrated in the inner regions of high magnetic field, which leads to a substantial increase in q. The safety factor at the plasma edge in a typical Rotamak-ST discharge exceeds 20 with a cylindrical equivalent $q_{\mathrm{cyl}} = 5a^2kB_t/RI_p = 2.1$.

Such an increase in the safety factor q has a number of consequences. For instance, the toroidal field at the center of existing spherical tokamaks and the next generation under construction is usually ten times lower than that in conventional tokamaks, and the normalized current $I_N = I_p/(aB_0)$ may reach high levels before q falls below 2. Since the

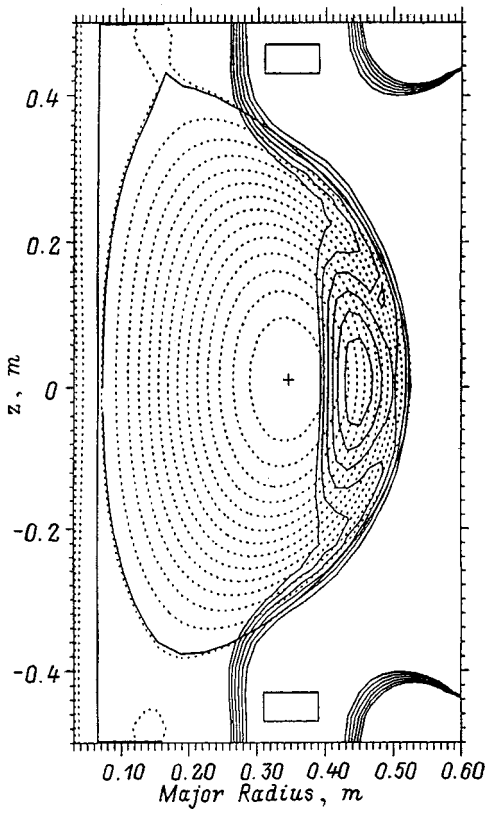

FIG. 1. Observation of a magnetic well (local $|B|$ minimum) in shot #35533 with high beta ($\beta_T \sim 40\%$). The dashed lines show the magnetic surfaces, the solid lines give the $|B|$ contours.

Troyon scaling indicates that β is proportional to I_N, this implies that high values of β may be attained (see the next section). Furthermore, at the high β values obtained in START, the associated diamagnetic effect compensates for the paramagnetic effect produced by the high normalized current, so that a magnetic well may appear on the $|B|$ profile (Fig. 1) (the results plotted in Figs. 1–6 pertain to START). This improves plasma stability[8] and also, by modifying the trajectories, reduces the fraction of trapped particles and consequently enhances the current generation efficiency. The efficiency of utilization of the toroidal field may be characterized by a simple parameter: the ratio of the plasma current to the current in the central rod I_p/I_{rod}. In recent spherical tokamak experiments, the ratio I_p/I_{rod} has exceeded one. In START the ratio $I_p/I_{\text{rod}}=1.2$ was obtained in a plasma with aspect ratio $A=1.3$ (Ref. 3), and in Rotamak-ST the ratio $I_p/I_{\text{rod}}=1.1$ was recently obtained for $A=1.1$ without any indications of the MHD limit being reached.[7] In TS-3 in Tokyo, a transition regime has been studied between a spheromak and a spherical tokamak and the results of numerical calculations[9] indicate that the ratio I_p/I_{rod} may be increased substantially, and for very low values $A=1.1$ it may exceed 3.

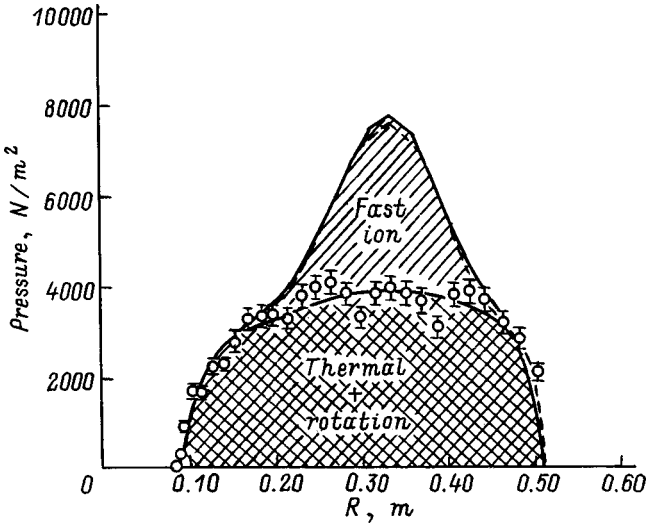

FIG. 3. Measurement of β values at the center. Solid curve — sum of thermal energy (determined by Thomson scattering and neutral charge-exchange spectroscopy) and fast ion pressure (with artificially imposed isotropy), dashed curve — results of calculations using the EFIT code based on magnetic measurements.

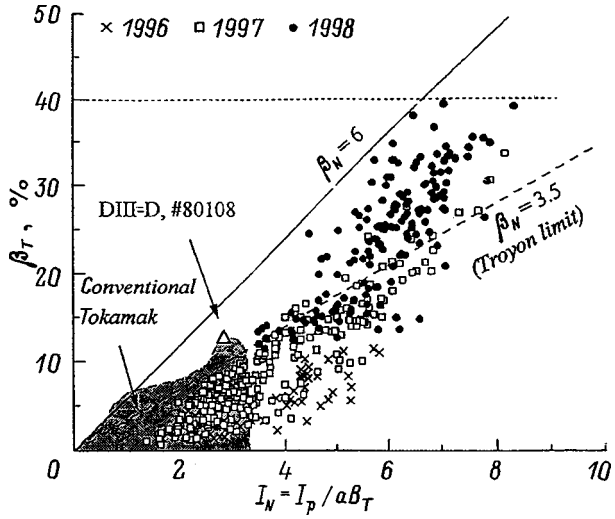

FIG. 2. Measurements of average β_T values.

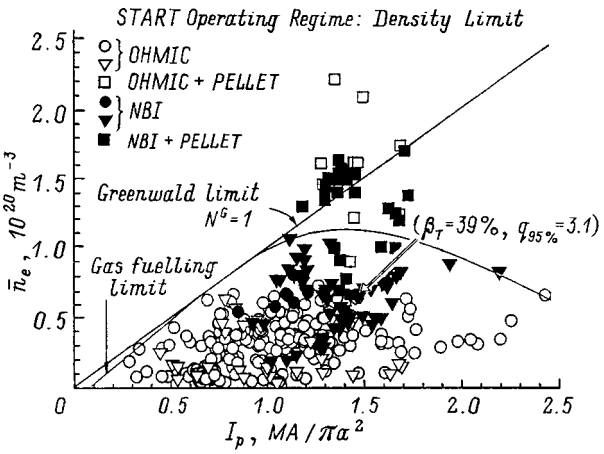

FIG. 4. Region of operating parameters demonstrating the increased accessibility for pellet injection using the Frascati/Riso injector. The limit for gas fueling discharge regimes is shown.

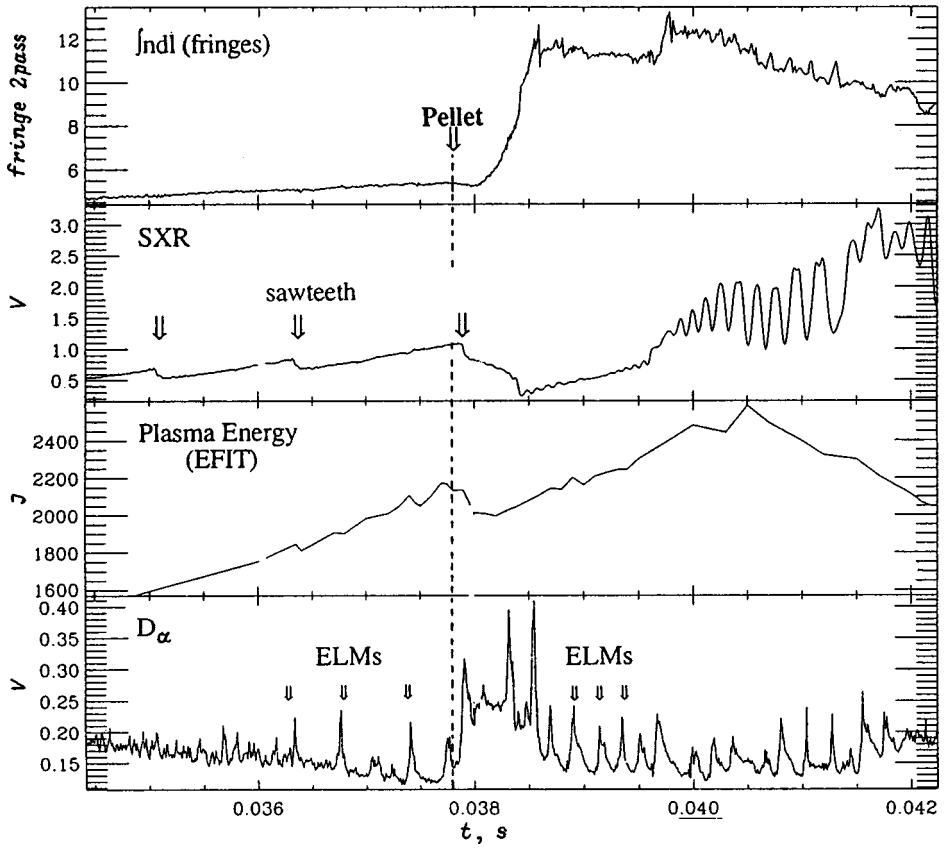

FIG. 5. Pellet injection into a discharge with ELM instabilities in a neutral beam injection regime.

2) High beta.

The ratio β of the plasma thermal energy to the magnetic energy is a characteristic of the efficiency of a fusion device. When a 1 MW neutral particle injector loaned by Oak Ridge National Laboratory (US) was used on START, an average $\beta\sim 40\%$ was obtained (Fig. 2), which is more that three times the highest value obtained in a conventional tokamak [here β is defined as $\beta_T=2\mu_0\int pdV/(VB_0^2)$, where V is the plasma volume and B_0 is the vacuum toroidal field in the geometric center of the plasma]. The factor β_N which is the ratio of β to the normalized current I_N, was increased to $\beta_N\sim 5.9$ and it was also demonstrated that these high beta values ($\beta_T>30\%$ and $\beta_N>4$) can be sustained over times several times longer than the energy confinement time.[10]

3) Central beta in the START tokamak.

The central beta, defined here as $\beta_0=2\mu_0 p_a/B_a^2$, where p_a and B_a are the plasma pressure and the vacuum toroidal field measured on the magnetic axis, may reach 100% in START. Whereas the average beta is adequately determined by processing the magnetic measurement data using the EFIT code,[11] the central beta is not determined as reliably and requires additional analysis using data on the electron temperature and density profiles obtained using Thomson scattering diagnostics for 30 points. The profile of the thermal ion temperature was measured by CELESTE diagnostics of neutral particle spectra, while the behavior of the fast ions was calculated using the Monte Carlo method. The profiles of the various components are shown in Fig. 3 for time $t=36$ ms and discharge #36544 with $\beta_T=31\%$ and $\beta_0=99\%$. Good agreement is observed when the total energy is compared with the results of the analysis using the EFIT equilibrium code (dashed curve).

4) Range of operating parameters.

a) Density limit.

High densities $n_e\geqslant 10^{20}\,\mathrm{m}^{-3}$ with the Greenwald number $G=\pi a^2 n_e/I_p[10^{20}\,\mathrm{m}^{-1}/\mathrm{MA}]\sim 1$ were obtained in START in high-pressure regimes. The range of operating parameters was extended still further by injecting frozen deuterium pellets upward into the plasma from the inside. The pellet vaporizes and supplies working gas to the central regions of the plasma. The results obtained for START indicate that the plasma density is approximately trebled, thereby extending the range of operating parameters, as shown in Fig. 4, approaching the Greenwald number $G\sim 1.5$ and the Murakami parameter $M=n_eR/B[10^{20}\mathrm{m}^{-2}/T]\sim 2.4$ (Refs. 12–14). Figure 5 shows oscilloscope traces of signals obtained in neutral-injection discharge #34881 with peripheral ELM MHD activity under conditions of pellet injection. A twofold increase in density averaged over the line of observation was recorded: sawtooth oscillations cease but the ELM activity continues, a picture consistent with the successful injection of deuterium into the central part of the discharge.

b) q limit.

A remarkable feature of high-beta discharges in a double-null divertor configuration is that they can operate with extremely low values of $q_{95}=2.3$. The initial calcula-

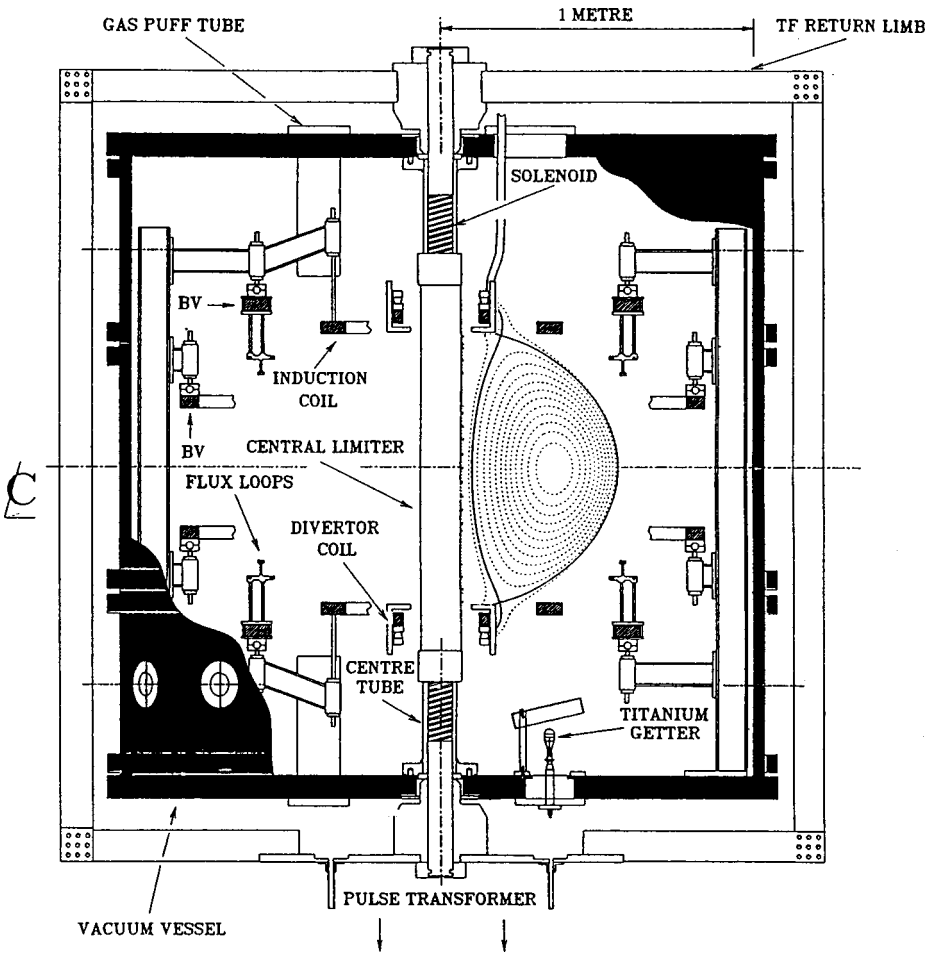

FIG. 6. Schematic of experiment showing proximity of the divertor coils to the plasma in a typical regime with high parameters.

tions used a plasma model with a limiter in which an edge safety factor $q\sim 4$ was assumed for discharges with low aspect ratio and large coefficients of vertical elongation and triangularity of the outer magnetic surface. Recent calculations using a plasma model with a separatrix showed that values of $q_{95}\sim 1.1$ are quite attainable (assuming some stabilization by the walls).

c) Beta limit.

START demonstrated that low-aspect-ratio configurations exhibit enhanced stability at very high beta. In a plasma with $\beta_N > 4$ and low shear, the second region of stability for ballooning modes was approached in the center. The values of β_N obtained in START are considerably higher than the conventional Troyon limit $\beta_N = 3.5$ shown in Fig. 2, which is in striking contrast to the situation obtaining in large collisionless tokamaks with normal aspect ratios, where neoclassical tearing modes limit β_N to approximately 2. Recent equilibrium calculations[15] have shown that the stabilizing Glasser term is significantly greater under typical conditions in spherical tokamaks. Research using the next generation of large spherical tokamaks with hotter, less collisional plasmas will determine whether all these favorable predictions are sufficient to sustain the high β_N regimes obtained in START.

5) Internal reconnection of magnetic field lines.

The range of operating parameters in START in high-density low-q regimes was usually limited by an effect known as internal reconnection of magnetic field lines, or an internal reconnection event (IRE). This effect has features similar to the ''soft disruptions'' observed in conventional tokamaks (including a precursor and thermal energy dumping) and usually takes place once or twice during each START discharge and is also observed in CDX-U (Ref. 16), Medusa (Ref. 17), and HIST (Ref. 18). This effect appears to be caused by low-m MHD modes; for instance, a small IRE sometimes occurs at the early stages of a START discharge when the value of q at the center initially falls below 2 and the $m=2$ instability grows rapidly. Strong gas puffing or a reduction in the toroidal field (and/or an increase in the plasma current) can also lead to a series of IREs .

Hayashi *et al.*[19] proposed an interesting nonlinear model of the IRE process. They showed that at low aspect ratios the $m=2/n=2$ and $m=1/n=1$ modes can interact with each other (as observed in CDX-U), and reconnection of internal and external magnetic fields may lead to current surges, a drop in thermal energy, and the helical perturbations observed in the experiments.

Equilibrium calculations based on magnetic probe data and profile measurements show that a consequence of IREs in START is a flattening of the current profile. This is con-

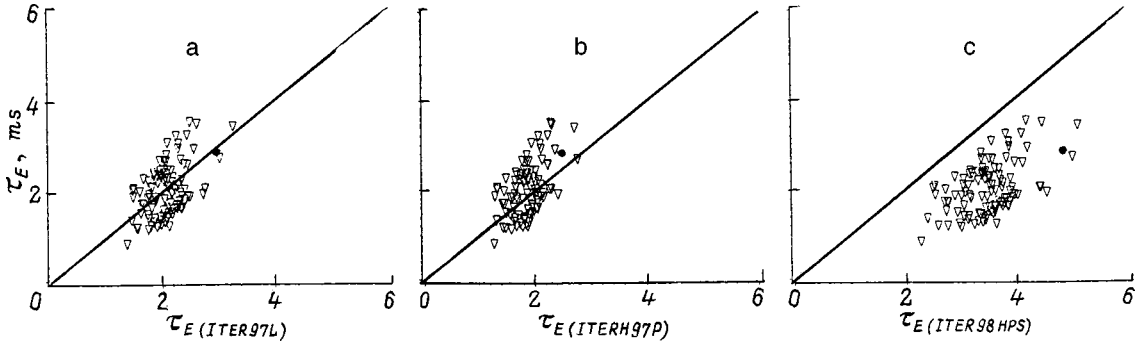

FIG. 7. Comparison between START data (Ohmic regime and neutral injection regime) with recent ITER scalings for the *L* (a) and *H* mode (b). Recent START discharges with improved confinement are not shown: c — the same data compared with the recent Kardaun ITER-98 scaling.

sistent with relaxation theory, which predicts that the plasma will be transferred to a state of minimum magnetic energy which, in a tokamak, is associated with a flattening of the current profile. At low aspect ratios, the plasma has a larger natural vertical elongation, which increases still further as the internal inductance decreases, as was observed experimentally. This leads to an increase in q at the plasma edge and has a stabilizing effect (Sec. 1.7).

6) Instabilities in spherical tokamaks caused by high-energy particles.

Neutral particle beam heating of the plasma in START was accompanied by several different classes of MHD instabilities.[11] The excitation of Alfvén instabilities by high-energy ions in START is obviously a harmless effect, which may be an inherent feature of spherical tokamaks characterized by low toroidal fields. However, the extent to which the instabilities observed in START are attributable to characteristics unique to this tokamak has yet to be determined. Calculations of particle trajectories show that a considerable number of the beam ions intersect the last closed magnetic surface and become neutralized. This effect which leads to an unstable spatial velocity distribution will not be so important in MAST, and in spherical tokamak reactors it will be altogether negligible. Hence, one source of instability in START will most likely be absent in large spherical tokamaks.

7) Stability against major disruptions.

Prior to October of 1995, internal reconnection events (IREs) in the START plasma with low aspect ratio ($A<1.8$) did not cause much harm, their only consequences being a reduction in the stored thermal energy, a current surge, and increased plasma elongation. No major current disruptions had been observed (except in regimes with $A>1.8$). An explanation for this apparent ''immunity'' to major disruptions was put forward in Ref. 20. As a result of the flattening of the current profile caused by IRE, a low-aspect-ratio plasma significantly increases its vertical elongation, which leads to an increase in q_{95} and an increased safety factor. Moreover, the reduction in the plasma inductance leads to an appreciable increase in the plasma current, which counteracts any decrease in the major radius of the plasma ring and thereby prevents intensified interaction between the plasma and the inner cylinder of the chamber.

The close proximity of the divertor coils to the plasma prevents any increase in elongation, as can be seen from Fig. 6. In fact, after the divertor coils were installed in START in 1995, discharge disruptions began to occur at low q. It is hoped that the large gap envisaged in the next machine, the Mega-Amp Spherical Tokamak (MAST), will restore the high stability against major disruptions. This stability would be a major advantage in future large facilities with high plasma currents and is another important property that needs to be studied in the next generation of spherical tokamaks.

8) Energy confinement in the START tokamak.

Exact calculations of the energy confinement time in START are difficult because of the short discharge duration and the indeterminacy in the measurements of the absorbed beam energy. In order to determine the stored plasma energy in the thermal component, the total stored plasma energy is first determined using the EFIT code from carefully calibrated magnetic measurement data, and then the calculated fraction of fast ions is subtracted. The deposited power P is taken as the sum of the Ohmic power and the absorbed neutral beam power. For this purpose the fraction of the power passed straight through and the beam particle losses on the first orbit (in total these may reach 50% in START because of the relatively low current and the small plasma dimensions) are subtracted from the total beam power.

Figure 7 compares results obtained for START (for the *L* mode with additional heating without any indications of the *H* mode) with the generally accepted ITER-97 scaling for confinement times in conventional-aspect-ratio tokamaks in the *L* mode (a) and in the *H* mode (b). It should be noted that the predictions of this empirical scaling for confinement times in the *H*- and *L* modes for aspect ratios typical of START are very similar, although there is experimental evidence (see below) to suggest that improved confinement regimes very similar to the *H* mode may exist in START. This anomaly may be eliminated by modifying the dependence on the aspect ratio in the ITER scaling for the *H* mode proposed recently by Kardaun,[21] which does not alter the predictions for ITER but gives improved predictions for the START data. In Fig. 7c the new scaling (known as ITER-98HPS) is

compared with the data base for the L mode regime in START. The point denoted by the filled circle corresponds to discharge #35533 with 40% beta. The START discharges in which we observed the peripheral instabilities described in the next Section are currently being analyzed.

It is predicted that the more accurate data on confinement times obtained in the next generation of spherical tokamaks will have a major influence on tokamak scalings.

9) Improved confinement characteristics in the START tokamak.

Recent START discharges which have demonstrated exceptional results have revealed various characteristics typical of H-mode improved plasma confinement regimes.[22] These include long periods accompanied by ELM instabilities, changes in the poloidal rotation, a drop in the D_a line intensity (which can be identified from the fact that the plasma edge becomes significantly more clearly defined), a density pedestal at the plasma edge, and a spontaneous density rise. The rapid increase in energy confinement observed in the periods between ELM instabilities indicates significantly improved confinement.

The discharges with ELM instabilities observed in START are the first indication that an improved confinement regime with a transport barrier may be achieved in a spherical tokamak. Although in typical discharges the plasma parameters significantly exceeded the usual threshold for a transition to the H mode,[23] this was only achieved in certain discharges with improved vacuum conditions and for specific combinations of parameters. This may be attributed to the relatively poor vacuum inside the system and/or the high neutral particle concentration at the periphery exacerbated by the large ratio of vacuum volume to plasma volume (of the order of 20:1). Thus, details of the transition to the H mode are another feature of spherical tokamaks requiring further study.

Calculations made by Rewoldt *et al.*[24] showed that the growth rate of instability caused by collisionless electrostatic toroidal drift modes decreases very rapidly as the aspect ratio decreases for $A<1.5$. Later calculations of electromagnetic gyrokinetic microinstability[25] showed that in spherical tokamaks with $A\sim1.4$ and high $\beta\sim45\%$, broad MHD-stable pressure profiles have a velocity shear exceeding the linear growth rate of all the drift modes and ITG modes caused by the ion temperature gradient. These calculations predict that in the next generation of machines transport caused by plasma turbulence may be reduced considerably (this may well occur to some extent in the START experiments).

10) Physics of the peripheral region of the START tokamak.

The thermal load on the surface facing the plasma is an important problem at low aspect ratios, since the power release from the plasma may be distributed over a smaller area. A detailed study of the released power distribution over the four main points on the divertor plates in a standard double-null divertor configuration revealed strong asymmetry between the outer and inner sides of the machine, and also some asymmetry between the top and bottom on the inner circumference.[26]

The regions of energy release on the divertor plates are usually narrow (of the order of 0.5–1.5 cm), which indicates that large gradients exist at right angles to the region outside the separatrix because of the appreciable expansion of the poloidal flux in comparison with the equatorial plane (approximately 8–12 times).[27] The observed asymmetry between the top and bottom appears to be caused by particle drift induced by the electric field in the region beyond the separatrix, which may predominate over drift in the toroidal magnetic field. However, the asymmetry between the outer and inner sides of the divertor plates may well be unrelated to the particle drift effect.

The results of observations indicate that the power flux from the plasma to the outer layer beyond the separatrix usually exceeds the value predicted from a comparison of relative surface area. The highest power flux densities (around 4–8 MW/m^2 in neutral-beam START discharges) were observed (randomly) on the outer side of the divertor plates.

11) Halo currents in spherical tokamaks.

In conventional-aspect-ratio tokamaks, a plasma having a vertically elongated poloidal cross section is initially unstable with respect to vertical displacements. If the feedback system does not operate during termination of the discharge, the vertical displacement of the plasma and current dumping are frequently accompanied by the appearance of halo currents flowing through the composite circuit closed by the plasma and its surrounding conductors. The toroidal asymmetry of the halo currents generates lateral and tilting forces which act on the vessel. Results obtained with conventional tokamaks such as COMPASS-D (Ref. 28) indicate that the total halo current may reach 40% of the plasma current with appreciable toroidal asymmetry (the toroidal peaking factor reaches ~3.5).

Although the START plasma can conserve vertical stability under appreciable elongation (which depends on the current profile), vertical displacements may take place at greater elongations and/or during a disruption. Measurements of the halo currents confirmed the CDX-U findings[16] that in a spherical tokamak the symmetric halo currents in the central column are very small ($<3\%$ of the total plasma current), have a small asymmetry ($<5\%$), and thus should not present problems for the next generation of tokamaks. Results obtained for COMPASS, START, and CDX-U are compared in Fig. 8.

These results are consistent with recent models which describe toroidal asymmetry in terms of inductively coupled elements of the plasma–wall circuit.[29,30] The mechanism depends on the composite circuit closing the torus in the poloidal and toroidal directions. At low aspect ratios the relatively short length of the circumference around the inner cylinder of the vessel allows induced currents to flow in the toroidal direction, creating a stabilizing effect which impedes the evolution of halo currents.

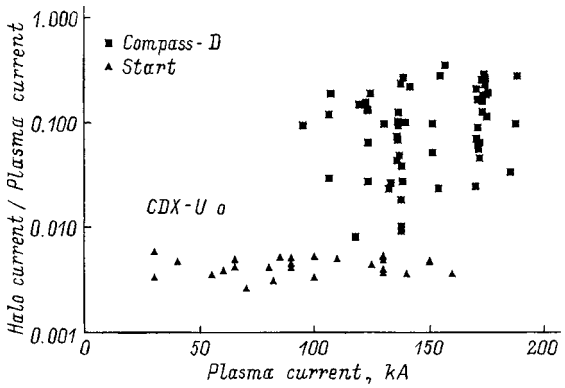

FIG. 8. Comparison of halo currents measured on COMPASS-D (~2.6 aspect ratio) and START (aspect ratio ~1.3–1.5).

CONCLUSIONS

As has been shown in the previous Sections, START and other prototypes of spherical tokamaks exhibit many very promising features ($\beta_T \sim 40\%$, $\beta_0 \sim 100\%$, magnetic well, improved confinement regime, stability against disruptions, and so on). The results of these experiments allow one to view the physics of toroidal-confinement systems from a different angle and to reduce the uncertainty in the use of scalings. An important task for the next generation of machines with hot collisionless plasmas and mega-ampere currents, such as GLOBUS-M, PEGASUS, ETE, MAST and NSTX, is to check out these results in collisionless regimes and to expand the tokamak data base by developing scalings for aspect ratio and linear dimensions.

Potentially, spherical tokamaks may provide a cheap route to achieving controlled thermonuclear fusion through the creation of bulk neutron sources and nuclear power plants.[31] However, it is first important to demonstrate that cw operation (predicted theoretically) is possible by effectively generating a plasma current with a high fraction of bootstrap current, and that the system can operate with the large vertical elongation of the plasma required to achieve high efficiency in a nuclear power plant.

The author would like to thank the START and Neutral Injector groups at Culham Laboratory, and many other colleagues at Culham and throughout the world who have supplied their data. This work is jointly supported by Euratom and the UK Department of Trade and Industry. The neutral particle injector for START was loaned by the US Department of Energy, the pellet injection experiments were carried out in collaboration with Riso and ENEA (Frascati), and the EFIT computer code was supplied by General Atomics.

[1] Y-K. M. Peng and D. J. Strickler, Nucl. Fusion **26**, 769 (1986).
[2] D. C. Robinson, *Fusion Energy and Plasma Physics* (World Scientific, Singapore 1987), p. 601.
[3] A. Sykes *et al.*, Plasma Phys. Controlled Fusion **39**, 247 (1997).
[4] G. A. Collins, G. Durance, G. R. Hogg, J. Tendys, and P. A. Watterson, Nucl. Fusion **28**, 255 (1988).
[5] H. Bruhns, R. Brendel, G. Raupp, and J. Steiger, Nucl. Fusion **27**, 2178 (1987).
[6] P. K. Browning *et al.*, Phys. Rev. Lett. **68**, 1722 (1992).
[7] I. R. Jones, C. Deng, I. M. El-Fayoumi, and P. Euripides, Phys. Rev. Lett. **81**, 2075 (1998).
[8] D. Palumbo, Nuovo Cimento **53**, 507 (1968).
[9] M. Yamada, A. Morita, N. Pomphrey *et al.*, Nucl. Fusion **36**, 1210 (1996).
[10] M. Gryaznevich *et al.*, Phys. Rev. Lett. **50**, 3972 (1998).
[11] L. L. Lao *et al.*, Nucl. Fusion **25**, 1611 (1985).
[12] C. Ribeiro, in *Proceedings of the 25th EPS Conference on Controlled Fusion and Plasma Physics*, Prague, 1998, p. 741.
[13] S. Medvedev, *ibid.*, p. 658.
[14] M. Gryaznevich, K. McClements *et al.*, *ibid.*, p. 724.
[15] S. E. Kruger, C. C. Hegna, and J. D. Callen, Phys. Plasmas **5**, 455 (1998).
[16] M. Ono *et al.*, Proc. IAEA Fusion Energy, Montreal, Vol. 271 (1996).
[17] R. Fonck, in *Proceedings of the International Workshop on Spherical Tokamaks*, Culham, England, 1996, Vol. 1, p. 49.
[18] M. Nagata *et al.*, *ibid.*, Vol. 2. p. 717.
[19] T. Hayashi, to be presented at the 17th IAEA Fusion Energy Conference, Yokohama, 1998.
[20] A. Sykes, Phys. Plasmas **4**, 1665 (1997).
[21] O. J. W. Kardaun, in *Proceedings of the ITER Confinement Workshop*, PPPL, Princeton, 1998.
[22] P. G. Carolan *et al.*, Plasma Phys. Controlled Fusion **40**, 615 (1998).
[23] J. A. Snipes, in *Proceedings of 24th EPS Conference on Controlled Fusion and Plasma Physics*, Berchtesgaden, 1997, Vol. 3. p. 961.
[24] G. Rewoldt *et al.*, Phys. Plasmas **3**, 1667 (1996).
[25] W. Dorland and M. Kotschenreuther (private communication).
[26] K. Morel, G. F. Counsell, and P. Helander, presented at the 13th PSI Conference, San Diego, 1998, J. Nucl. Mater (to be published).
[27] G. F. Counsell *et al.*, Paper presented at the 13th PSI Conference San Diego, 1998, J. Nucl. Mater (to be published).
[28] A. W. Castle and G. G. Morris, in *Proceedings of the EPS Conference*, Kiev, 1996.
[29] A. Caloutsis and G. G. Gimblett, Preprint UKAEA FUS. 393 (1998), Nucl. Fusion (to be published).
[30] N. Pomphrey, J. M. Bialek, and W. Park, Nucl. Fusion **38**, 449 (1998).
[31] D. C. Robinson, in *Proceedings of the International Symposium on Plasma Dynamics in Complex Electromagnetic Fields*, Kyoto, 1997, p. 157.

Translated by R. M. Durham

Globus-M spherical tokamak

V. K. Gusev, V. E. Golant, E. Z. Gusakov, V. V. D'yachenko, M. A. Irzak, V. B. Minaev, E. E. Mukhin, A. N. Novokhatskiĭ, K. A. Podushnikova, G. T. Razdobarin, N. V. Sakharov, E. N. Tregubova, V. S. Uzlov, and O. N. Shcherbinin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

V. A. Belyakov, A. A. Kavin, Yu. A. Kostsov, E. G. Kuz'min, and V. F. Soĭkin

D. V. Efremov Research Institute of Electrophysical Apparatus, 189631 St. Petersburg, Russia

E. A. Kuznetsov and V. A. Yagnov

Troitsk Institute for Innovative Fusion Research, 142092 Troitsk, Moscow Province, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 58–62 (September 1999)

The main parameters and design of the Globus-M spherical tokamak are described, including the modifications introduced after construction was completed. The experimental program and possibilities for optimizing the operating regimes of the facility are discussed. The research program envisages beginning with Ohmic heating experiments and continuing with work on rf heating and generation of drag currents. Recent results of modeling the propagation of rf waves in the 10–50 MHz and 2.45 GHz ranges are presented and the possibility of using these to control the profiles of the plasma parameters is discussed. A brief description is given of the primary diagnostics and the further development of the diagnostic complex is discussed.
 [S1063-7842(99)01309-4]

The Globus-M spherical tokamak has been developed by member organizations under a project of the International Scientific-Technical Center and has been manufactured at the ''Leningrad Northern Plant'' State Enterprise (St. Petersburg). The tokamak has been installed in a new experimental facility at the A. F. Ioffe Physicotechnical Institute of the Russian Academy of Sciences. The area of the two-storey building is 600 m^2. In addition to the device, the facility will also accommodate technological equipment and systems for diagnostics and auxiliary plasma heating.

DESCRIPTION OF TOKAMAK

The final parameters of the Globus-M spherical tokamak were improved over the initial design:[1] the nominal plasma current was increased to 0.5 MA and the toroidal magnetic field was increased to 0.65 T. The major radius R is 0.36 m and the minor radius $a=0.24$ m, which corresponds to an aspect ratio of 1.5, where the vertical elongation of the plasma column may reach 2.2 and the triangularity 0.6.

The 1.1 m^3 vacuum vessel is an all-welded austenitic stainless-steel structure having 38 diagnostic ports with a total area of 0.8 m^2, which provide good access to the plasma for diagnostics and the auxiliary heating sources. Both rf heating and current generation, and neutral injection heating are envisaged. The large ports and the small distance between the outer magnetic surface of the plasma and the vessel walls allow antennas for coupling rf power into the plasma to be conveniently and effectively sited, and can also enhance the neutral injection efficiency by reducing the charge transfer losses. In this tokamak design, an auxiliary power of up to 8 MW can be coupled into the plasma.

The tokamak electromagnetic system has a classical design with both windings located outside the vacuum vessel (Fig. 1). The sixteen, single-turn, D-shaped, toroidal field coils (TF) are made of M1 electrotechnical copper, connected in series to form the toroidal field winding. The coil sections passing through the inner cylinder of the vacuum vessel are fabricated as sixteen mutually insulated segments of thermally quenched bronze. The solenoidal field coil, or central solenoid (CS), is wound onto the central rod formed by these segments. The solenoid is fabricated as a two-ply coil with an overall winding height of 1.3 m. A 66 m long continuous conductor of cold-rolled silver-containing copper having a 20×20 mm rectangular cross section is used for the winding. The nine pairs of poloidal coils are divided into three functional groups, i.e., fringe-field compensating coils (CC) and ''slow'' (PF) and ''fast'' (VFC, HFC) coils to control the plasma position and shape.

Figure 2a shows the operation of the poloidal system of the Globus-M tokamak used to compensate for the fringing fields of the central solenoid in the prebreakdown phase. Figure 2b shows the magnetic configuration in the discharge with two X points (plasma current 0.5 MA, elongation 2, triangularity 0.6).

The tokamak coils are supplied from power transformers connected to a high-voltage 110 kV ac grid using six-phase thyristor rectifiers and fast current inverters at frequencies up to 3 kHz. The power sources are contained in feedback loops and can sustain specific plasma positions, profiles, and currents. The cooling system for the windings in the electromagnetic system ensures that the tokamak can operate at a

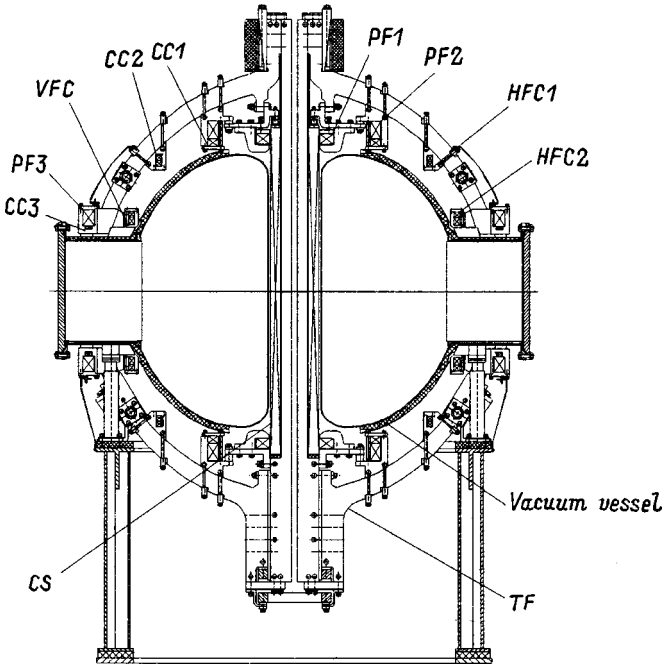

FIG. 1. Cross section of Globus-M tokamak showing: vacuum vessel, TF — toroidal field coils, CS — central solenoid, CC — compensating coils, PF — ''slow'' poloidal field coils, and VFC, HFC — ''fast'' control coils.

frequency of six shots per hour with a discharge duration of 0.3 s and a rated current of 0.5 MA, which is sufficient for relaxation of the current density profile. Discharges at the rated plasma current are provided by the 0.33 V·s capacitance of the poloidal coil system to which the central solenoid contributes 90%. Details of the tokamak design are described in Refs. 2 and 3.

LIMITING REGIMES OF THE GLOBUS-M TOKAMAK

For the chosen scheme of supplying the tokamak from the ac grid, the duration of the working pulse is limited mainly by the heat release in the windings of the electromagnetic system. The discharge duration may be increased to 1 s by increasing the pause between pulses.

An attractive feature of a spherical tokamak is that extremely high currents can be obtained. The modifications introduced in the design increased the current through the central rod to 1.2 MA, which increased the toroidal peak B_T on the plasma axis to 0.65 T and can produce a plasma current I_p up to 0.5 MA. The limiting plasma current while maintaining stability to helical perturbations is extremely high, and for tokamaks such as Globus-M may approach the magnitude of the current through the central rod. Various experiments[4] on the START spherical tokamak have demonstrated that it is possible to operate with a low edge safety factor q_a. For $q_a \approx 3.6$ and $B_T = 0.62$ T the plasma current in Globus-M may reach 1 MA.

Compared with other spherical tokamaks under construction Globus-M can achieve extremely high current densities averaged over the plasma cross section. Even under normal conditions, this density is 1.4 MA/m^2. This circumstance makes Ohmic heating more effective, since the contribution of the specific power is proportional to the square

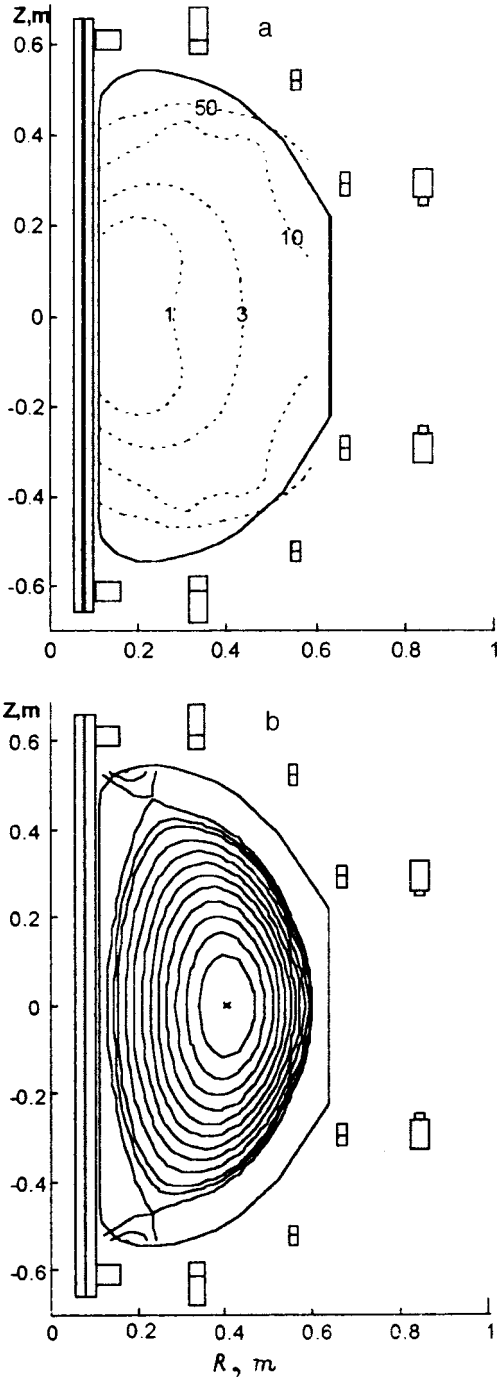

FIG. 2. Lines of constant poloidal component of the magnetic field inside the tokamak vessel: a — prebreakdown phase, the figures give the magnetic field in gauss, b — discharge with two X-points, $I_p = 0.5$ MA, $\beta_T = 0.27$, $I_i = 0.5$, $k = 2$, $\delta = 0.6$, Globus-M. Those windings of the electromagnetic system which are switched on are indicated by shading.

of the current density. The high current density and the high ratio B_T/R of the order of 1.8 T/m can produce a discharge having a high limiting average plasma density of around 2×10^{20} m^{-3} even when the discharge is being supplied with gas. The establishment of regimes with low edge safety factor q_a and high plasma density is an important line of research in the controlled fusion program, since this can enhance the energy yield of the fusion reaction, which is proportional to the square of the plasma density, and can

improve the reactor efficiency because of the high values of the normalized current $I_N = I_p/a \cdot B_T$ in regimes with maximum Troyon coefficients. For low values of q_a Globus-M attains values of $I_N \approx 7$. Calculations of the stability of the plasma column under normal conditions ($q_a \approx 3.3$) in a double X-point configuration with $I_N = 4.4$ indicate that this remains stable with respect to ballooning modes for $\beta_T \leqslant 18\%$ for the first stability zone.[5] In order to attain regimes with a large fraction of noninduction current generated by plasma pressure gradients,[6] we need to establish regimes with high $\beta_p = \langle p \rangle / (B_p^2/8\pi)$, where $\langle p \rangle$ is the volume-averaged plasma pressure and B_p is the magnetic field produced by the plasma current. These regimes can only be attained at maximum plasma pressure by increasing the edge safety factor. In this case, it is important to increase $q^* = L_p B_T / 2\pi R B_p$ (where L_p is the plasma circumference in the poloidal direction), i.e., to reduce the current or magnetic field at a bounded elongation. Reducing both the current and the field in turn leads to a decrease in the maximum plasma density. Globus-M can operate in the range $1.15 < q^* < 7$ at an average plasma density of $0.5\text{–}2.0 \times 10^{20}\ \mathrm{m}^{-3}$ which is substantially higher than that in other spherical tokamaks.

METHODS AND RESEARCH PROGRAM

The first stage of the experimental program, optimizing the parameters of an Ohmic discharge, will be implemented using conventional methods of producing plasmas and controlling their parameters. Compensation for fringing fields down to a few gauss (Fig. 2a) and plasma preionization at 15 GHz can substantially reduce the plasma breakdown voltage and economize on the stored volt-seconds available from the poloidal system. At the initial stage of the experiment, the current rise will be achieved by an inductive method. Calculations indicate that around $0.18\ \mathrm{V \cdot s}$ is required to achieve a plasma current of 0.3 MA and the rate of current rise is then around 5 MA/s. The plasma discharge is stabilized either by currents flowing in the adjacent conducting wall of the vessel or by a feedback control circuit. This can suppress fast vertical plasma surges and lower harmonics of helical modes. Another method of suppressing plasma MHD activity may involve generating a current in the boundary layer by applying voltage to the divertor plates.[7] Since the divertor plate segments in the Globus-M tokamak are insulated from each other and from the vacuum vessel, this experiment can be carried out without making significant modifications to the design.

High-power Ohmic heating combined with auxiliary heating methods leads to increased particle and energy fluxes to the vessel wall. In addition to conventional methods of preparing the surface of the vessel walls, particular attention has been paid to the choice of protective materials and the design of intravessel elements. The shielding elements for the central cylinder and the divertor plates are made of stainless steel, are easily removable, and can be coated with various protective materials such as B, Be, and C up to 1 mm thick. Rapid access to the shielding elements for replacement is provided by two ports, 400 mm in diameter, in the equatorial part of the vessel.

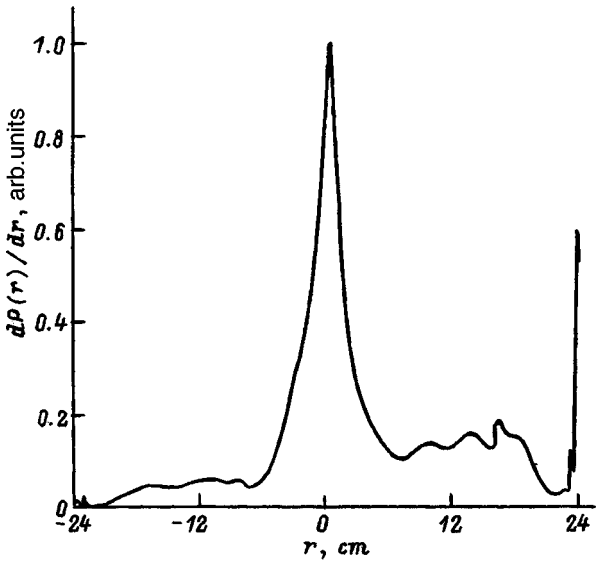

FIG. 3. Spatial profile of the absorption of rf power during ion cyclotron heating of the plasma: $f = 11$ MHz, $B_0 = 0.5$ T, 5%H+95%D.

Work has recently been carried out on the development of rf heating and current generation methods.[8,9] Simulations have been made of the propagation and absorption of magneto-acoustic waves in a quasitoroidal geometry, taking into account the poloidal and paramagnetic components of the field. Calculations were made for the frequency range 6–50 MHz, which covers the region between the fundamental and the eighth harmonic of the ion cyclotron resonance. The code developed was used to solve the one-dimensional wave equation with a ''hot'' permittivity tensor, allowing for corrections of the second order of smallness. The calculations thus allowed for absorption at the fundamental and second cyclotron harmonics and also for Landau damping, magnetic pumping, and interference mechanisms. For the Globus-M tokamak, strong absorption of magnetoacoustic waves was predicted during a single pass (in some cases up to 80%). Figure 3 shows local power absorption profiles integrated over the wave spectrum emitted by the antenna. The absorption zone for the fundamental harmonic of the ion cyclotron frequency is fairly narrow. For higher harmonics of the fast acoustic waves, absorption of energy at 30 MHz provides electron heating in a fairly broad band equal to half the minor radius. This method may be combined with ion heating at the second harmonic of the ion cyclotron resonance. Then electrons (in the broad band) and ions (in the narrow band) are heated simultaneously. This effect may be used to change the plasma pressure profile by tuning the oscillator frequency; this change in frequency has no significant influence on the overall absorption of energy and heating.

The conventional system for generating lower-hybrid waves is ineffective for spherical tokamaks because of the opacity barrier at the plasma edge. In this case, only waves with strong initial damping $N_\parallel = 7\text{–}10$ penetrate inward. However, the strong inhomogeneity of the magnetic field over the poloidal circumference means that a nonstandard system can be used for heating: waves having frequencies

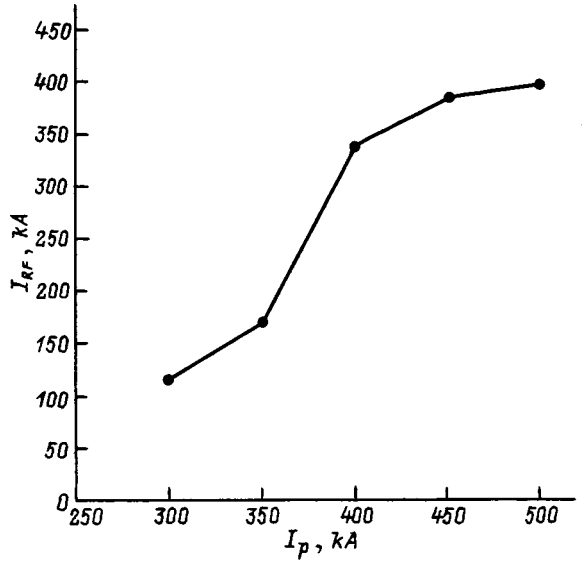

FIG. 4. Current generated by rf waves at 2.45 GHz as a function of plasma current ($n_{e0}=0.5\times10^{20}\,\mathrm{m}^{-3}$).

lower than $(e\,B_{\mathrm{pol}})/(m_e\,c)$, where B_{pol} is the poloidal magnetic field, m_e and e are the electron mass and charge, and c is the velocity of light, and low initial values of $N_{\|}=2-4$, are coupled into the plasma in the poloidal direction near the equator. As these waves penetrate inward, $N_{\|}$ increases as a result of the magnetic field inhomogeneity which leads to effective damping. The matching between a grille positioned in the equatorial plane and turned through 90° and the plasma is fairly good: the reflection coefficient is only 10–15% if the plasma density profile is not excessively steep. Calculations of the ray trajectories showed that depending on the plasma parameters, energy is absorbed on the radius between $0.3a$ and $0.7a$. The lower electron density and temperature at the periphery lead to central absorption. The higher values of the total current create conditions for effective current generation at the boundary which may reach $0.04\times10^{20}\,\mathrm{A/m^2W}$ (Ref. 9). Figure 4 shows the current generated by the waves as a function of the plasma current.

The research program includes the development of a diagnostic complex for the Globus-M. The first stage envisages making magnetic measurements using measuring loops and two sets of probes (36 in each set), using a video camera to monitor the plasma geometry, and also making measurements of various integrated plasma parameters such as the average density, soft and hard x-ray fluxes, and so on. The equipment required is ready for assembly or has already been installed in the tokamak. Further development of the diagnostic complex is directed toward measuring local plasma parameters with good temporal and spatial resolution. So far, draft designs have been completed and work has started on the working design of four innovative diagnostic facilities: a pulsed radar reflectometer, Thomson scattering systems, and x-ray tomography and high-speed video camera systems. The radar reflectometer is a combination of a phase reflectometer with frequency sweeping and a pulsed radar. The Thomson scattering system is based on a neodymium laser, which generates a series of twenty, 20 J pulses during the discharge, and silicon avalanche detectors. The system has an extremely high maximum sensitivity in terms of plasma density, around $1\times10^{-17}\,\mathrm{m}^{-3}$. The x-ray tomography system includes two pinhole cameras to record the soft x-rays and to measure plasma instabilities and its electron temperature. The high-speed video camera uses a CCD and has an automatic system for setting the exposure time during the recording of a frame, and also has an extremely large dynamic range.

Work on preparing the experiment on the Globus-M spherical tokamak has now reached the final stages, and physical commissioning is scheduled for the beginning of 1999.

[1] V. A. Belyakov, V. A. Divavin, N. Ya. Dvorkin *et al.*, *Spherical Tokamak Project Globus-M*, Preprint No. 1629, Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg (1994), 36 pp.
[2] V. E. Golant, V. K. Gusev, V. V. D'yachenko *et al.*, in *Proceedings of the 16th International Conference on Fusion Energy*, Montreal, 1966, paper IAEA-CN-64/GP-15 (IAEA, Vienna, 1997), Vol. 3, pp. 591–599.
[3] V. K. Gusev, N. V. Sakharov, V. V. Spheizman *et al.*, Fusion Technol. **34**(2), 137 (1998).
[4] M. P. Gryaznevich, R. Akers, P. G. Carolan *et al.*, Phys. Rev. Lett. **80**, 3972 (1998).
[5] S. G. Bespoludennov, L. M. Degtyarev, and S. Yu. Medvedev, Spherical Torus'97 **1**, 403 (1997).
[6] J. E. Menard, S. C. Jardin, S. M. Kaye *et al.*, Nucl. Fusion **37**, 595 (1997).
[7] R. J. Goldston, Plasma Phys. Controlled Fusion **21**, 1061 (1997).
[8] M. A. Irzak, E. N. Tregubova, and O. N. Shcherbinin, Fiz. Plazmy **25**(7), 1 (1999).
[9] A. R. Esterkin, E. Z. Gusakov, and M. A. Irzak, Spherical Torus'97 **1**, 265 (1997).

Translated by R. M. Durham

Pellet-injection-based technologies for fusion reactors

B. V. Kuteev

St. Petersburg State Technical University, Technical University Applied Physics (TUAP), Ltd., 195251 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 63–67 (September 1999)

An analysis is made of results obtained in the development of various technologies based on injecting pellets into tokamak and stellarator plasmas. Pellet injection is now used for fuel injection, plasma diagnostics, burn control, mode switching, and conditioning of the first wall. Physical and engineering problems encountered in the development of these technologies and results of recent applications are discussed. © *1999 American Institute of Physics.* [S1063-7842(99)01409-9]

Injection of hydrogen (fuel pellets) and impurity pellets is now widely used for various purposes in plasma magnetic-confinement devices including tokamaks and stellarators. This suggests that new technologies may be developed for fusion reactors based on pellet injection.

The idea of using small pieces of solid hydrogen accelerated to velocities of several kilometers per second to inject fuel into a device first arose in the early 1950s as part of the stellarator program. This method was considered as an alternative to conventional gas puffing, which cannot provide the necessary deep fuel injection under reactor conditions. Recent tokamak and stellarator experiments have demonstrated highly efficient injection of fuel pellets to sustain the discharge in closed magnetic configurations. A fuel pellet injection system has now been incorporated in the design of the International Thermonuclear Experimental Reactor (ITER).[1]

The use of pellets for plasma diagnostics was initially based on the similarity between the processes accompanying the evaporation of pellets in a hot plasma and the injection of barium clouds into the ionosphere. Information on the plasma parameters and transport properties was obtained by observing the evolution of clouds near pellets or the evolution of concentration and temperature perturbations created by the pellet. Both fuel and impurity pellets are used for diagnostic purposes.

Injection of lithium pellets has recently begun to be used to deposit a protective coating on the first wall to prevent heavy impurities from entering the plasma and thereby reducing the effective charge of the plasma. It was also found that since several discharge burning regimes can be achieved for the same external parameters, pellet injection could be used to switch from one mode to another, which considerably extends the scope for controlling the discharge burn.

Finally, one of the critical problems of a fast emergency shutdown of the discharge in tokamaks or quenching of the plasma before a possible disruption may be solved by injecting a fast pellet consisting of inert gases.

Problems of pellet injection in fusion facilities have been discussed in detail in various reviews and original articles.[1–9] The present review is concerned with the latest results, with particular attention being paid to reactor technologies.

FUEL INJECTION

1. Main pellet parameters.

As an example, Table I gives the specifications for the fuel pellet injection system for the ITER reactor.[1] It should be noted that almost all these specifications have been achieved separately in experiments using existing injectors.

2. Penetration of pellets into the plasma.

The evaporation of pellets in a high-temperature plasma is determined by the energy flux reaching the surface. At the high rates of evaporation typical of hydrogen pellets, this flux is predominantly formed by the high-energy tail of the electron distribution function and to a considerably lesser degree by the transverse ion flux. A cloud of neutrals which shields the surface from the incoming thermal fluxes is concentrated near the pellet surface. The electrostatic potential formed at the edge of the cloud to compensate for the difference between the electron and ion currents has a significant effect on the evaporation. There is additional shielding due to the formation of a secondary plasma and to perturbations of the magnetic field by the moving plasma streams.

The models developed so far fairly accurately describe the evaporation processes.[5,6,10–12] As evidence of this, Table II gives two theoretical and one experimental scaling relation obtained using data from the international database IPADBASE.[7] It can be seen that the dependences of the evaporation rate on the electron density n_e, electron tem-

TABLE I. Specifications of fuel injection system for ITER.

Parameter	Value
Pellet size	8–10 mm
Pellet velocity	0.6–1.5 km/s
Injection frequency	1–100 Hz
Injection time	10 000 s
Maximum throughput (D_2)	100 Pa $\cdot$ m^{3}s^{-1}
Maximum throughput (T_2)	50 Pa $\cdot$ m^{3}s^{-1}
Permissible losses of pellet mass	$<$30%
Repeatability	5% of value
Cycle time	$<$1 s

perature T_e, and the pellet radius r_p show good agreement. This confirms that neutral and electrostatic shielding have a substantial influence on the formation of the evaporation rate. In Table II the comparison is made for conditions corresponding to a Maxwellian plasma and no fast ion heating of the plasma. The presence of fast electrons and ions in the plasma may substantially increase the rate of evaporation.[12]

The rate of evaporation is sensitive to the plasma volume interacting with the pellet. In closed plasma configurations this volume has the scale of the volume between two closed magnetic surfaces separated by a distance of around the pellet radius. In open configurations the volume interacting with the pellet is limited by the length of the field line and is thus substantially smaller than that in closed configurations. This implies that the rate of evaporation in open systems may be significantly lower than that in closed ones.

TABLE II. Theoretical and experimental scalings for the rate of evaporation of hydrogen pellets.

Source		Rate of evaporation
Parks	Ref. 10	$\sim n_e^{0.33} T_e^{1.64} r_p^{1.33} M^{-0.33}$
Kuteev	Ref. 5	$\sim n_e^{0.45} T_e^{1.72} r_p^{1.44} M^{-0.283}$
IPADBASE	Ref. 7	$\sim n_e^{0.4} T_e^{1.72} r_p^{1.40}$

3. Pellet formation.

Recently substantial progress has been achieved in the continuous formation of fuel pellets. Several promising methods of pellet formation have been tested on test rigs. A multiextruder method[3] was proposed and developed in research at St. Petersburg State Technical University and Oak Ridge National Laboratory.[13] This technology uses several conventional extruders operating in series. Two or three extruders can simulate a continuous flow; while one extruder is operating, the others are filling with fuel. This technology was initially proposed for the ITER project.[14]

A more advanced method known as hydrogen screw extrusion[15–17] was implemented by St. Petersburg State Technical University and the company Technical University Applied Physics (TUAP), Ltd. A block diagram illustrating the system is shown in Fig. 1. Hydrogen gas is fed into a special unit with a rotating screw auger, maintained at cryogenic temperatures (8–12 K). In the upper part of the extruder the gas condenses into a liquid and then solidifies. The lower part of the screw provides sufficient pressure for extrusion at rates up to 50 mm/s.

A new modification of pellet formation technology directly in a barrel (*in situ* formation) was proposed in Ref. 18. Pellets are formed inside the cylindrical chamber of a barrel

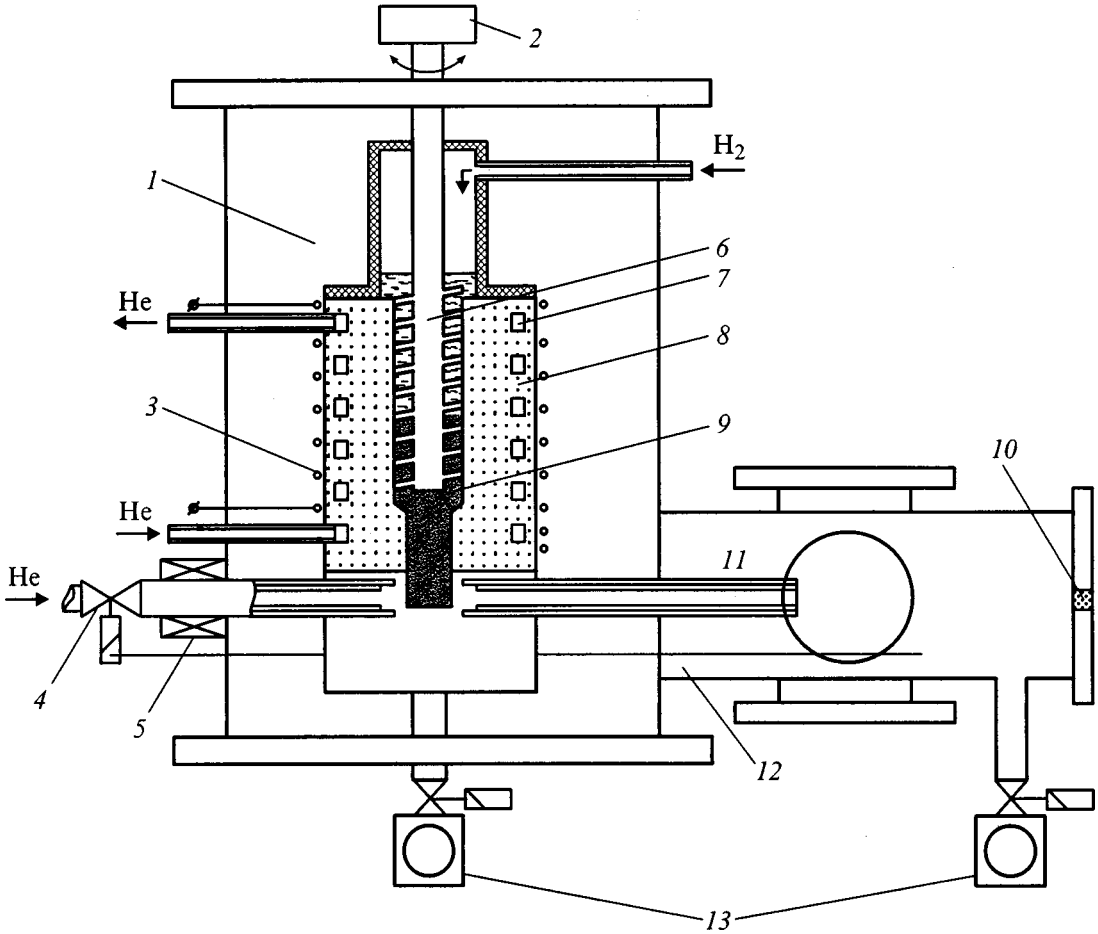

FIG. 1. Block diagram of fuel pellet injector with screw extruder: *1* — vacuum chamber, *2* — motor, *3* — heater, *4* — valve, *5* — electromagnetic drive, *6* — screw auger, *7* — heat exchanger, *8* — extruder, *9* — solid H_2, *10* — microphone, *11* — barrel, *12* — diagnostic chamber, and *13* — roughing pumps.

TABLE III. Comparative characteristics of fuel pellet shapers.

Shaper	Three-piston extruder	Porous shaper	Screw extruder
Tested pellet diameter	Not tested (1–10 mm for single extruder)	2–3 mm	2 mm
Extrusion velocity	50 mm/s (single extruder)	···	20–50 mm/s
Cycle time in continuous regime	Not determined (0.1 s with single extruder)	1 s	0.3 s (0.1 s attainable)
Continuous operating time	Not determined (7000 pellets for single extruder)	60 min	60 min
Velocity	1.4 km/s	1.2 km/s	700 m/s
Tritium technology	Demonstrated	Not checked	Not checked
Stored tritium	20×3 cm^3	1×10 cm^3	~ 10 cm^3
Reliability	Moderate	Good	Good

surrounded by a porous sleeve which is impregnated with solid hydrogen and connected to a gas supply system. The cell is maintained at a temperature of around 10 K. After a shot is fired by high-pressure gas, the cell is heated to a temperature at which the solid hydrogen can melt and diffuse into the barrel through the porous sleeve and refreeze to form a new pellet. This method can produce pellets at a frequency of around 1 Hz and suggests that the ITER parameters can be attained using a system with 10–20 barrels.

These methods of continuous pellet formation are compared in Table III. In view of the lack of return translational motion and the small amount of stored fuel in the shaper, systems using screw extrusion and porous shapers are more reliable than multiextruder ones. At present the ITER design is based on screw extrusion as the main fuel pellet injector. This technology can also be applied to existing tokamaks and accelerators for which long-pulse experiments are scheduled. This particularly applies to Tore-Supra and Large Helical Device.

4. Pellet acceleration.

Methods of pellet acceleration are listed in Table IV. Unfortunately, existing devices cannot achieve the velocities of around 10 km/s required to penetrate into the central regions of ITER-scale facilities. Nevertheless, all the methods of acceleration can ensure that the pellets penetrate beyond the so-called scrape-off layer (SOL). In view of the fact that the depth of penetration is a weak function of the pellet velocity ($\sim v_p^{1/3}$), all the methods listed in Table IV are competitive in terms of velocity. Under these conditions the most important technological parameters are reliability, simplicity, service life, and low load on the pumping system. In the table the methods are listed in order of their suitability for a reactor. Also listed are some typical problems still to be solved.

5. Position of the system on the machine.

Pellets are usually injected into a plasma from the magnetic-field side (from the outside). Recently some experiments have studied pellet injection from the high-magnetic-field side, from the top or bottom of the device in order to determine the optimum injection geometry from the point of view of penetration depth and fuel injection efficiency.[19] Here the efficiency is taken as the percentage of particle trapped by the plasma relative to those injected (see Ref. 20, for instance). The efficiency usually increases with depth of penetration. The search for the optimum direction of injection will clearly continue over the next few years.

DIAGNOSTICS

Both hydrogen and impurity pellets are used for plasma diagnostics.[4] Impurity pellets are usually accelerated in gasdynamic injectors[21] with typical velocities in the range 100–600 m/s. The number of electrons in a completely ionized pellet does not exceed 30% of the total number of electrons in the device, which corresponds to impurity pellets of diameter 0.3–0.6 mm.

Possibilities for the diagnostics of plasma parameters are based on observing the parameters of clouds near evaporating pellets,[22,23] pellet trajectories,[24] and measurements of the rate of evaporation.[4] Since the clouds are elongated in the direction of the magnetic field, high-speed photography of these clouds can be used to determine the direction of the

TABLE IV. Pellet accelerators.

Accelerator	Velocities	Problems
Centrifuge	<1.3 km/s	Difficulties with sizes larger than 2 mm, large pellet expansion angle, matching between shaper and accelerator
Single-stage gun	<1.5 km/s	Increased gas loading on vacuum and tritium systems
Gasdynamic accelerator	<0.6 km/s	High gas load
Two-stage gun	<4.2 km/s	Piston damage, high gas load, problems with periodic regime
Railgun accelerator	<2 km/s	Erosion of electrodes

magnetic field at various points on the pellet trajectory. Recent experiments on the ASDEX-Upgrade tokamak have shown that information on central values of the safety factor can be refined considerably by using this method.

Plasma drift in the poloidal direction leads to asymmetry of the cloud shape. By observing this asymmetry in lines of various ions, it is possible to estimate the radial electric field in the plasma — an important parameter responsible for the plasma transport properties and the type of energy confinement regime. This method has been successfully used in the FT-1 and T-10 tokamaks. Plasma drift velocities in the range 10^5-10^6 cm/s were observed.

In tokamaks, the pellet trajectories are curved because the pellets are accelerated in the toroidal direction under the influence of the plasma current. The curvature of the trajectory is proportional to the local current density so that it can be measured. Such measurements have been made in regimes where electron cyclotron resonance currents were generated.[24]

The rate of evaporation contains information on the localization of magnetic islands and fast particle beams. At the sites of islands, the rate of evaporation decreases as a result of a reduction in the volume of plasma involved in the evaporation. Near beams, the rate of evaporation increases appreciably. Significantly the spatial resolution of the pellet method is limited by the pellet size, which is far superior to other methods of diagnostics.

Pellet injection is accompanied by perturbations of the electron density and temperature which can be used to study transport phenomena in a plasma.[25] Injection of impurities provides information on their transport.

Recently charge transfer of alpha particles at clouds surrounding a lithium pellet was successfully used to determine their parameters in the TFTR tokamak.[26] This method is a unique means of determining the spatial and energy distribution of alpha particles.

TERMINATOR PELLETS

It has been suggested that 1–10 cm pellets consisting of deuterium or inert gases could be injected into a plasma for rapid cooling and current removal in the event of an accident situation, and also to obviate the risk of a major disruption.[4,9,27,28] Numerous experiments have demonstrated the high efficiency of this technology in T-10, JT-60, ASDEX-Upgrade, and DIII-D. Terminator pellet injectors are included in the ITER design. These pellets are presently injected using existing hydrogen pellet injectors, but using inert gases. Specially developed injectors based on light-gas guns are used to inject high-Z solids (KCl, SiC, Ti).

CONDITIONING OF THE FIRST WALL

The struggle to achieve plasma purity in closed plasma configurations has led to the conclusion that low-Z materials are the best solution for the first wall. Even a monolayer of lithium, beryllium, or boron can substantially reduce impurity fluxes and the effective plasma charge. Impurity pellets injected into a plasma evaporate to create a source of atoms which uniformly cover the wall with a low-Z protective layer. Lithium injection experiments revealed a significant improvement in the plasma parameters in the TFTR tokamak.[29,30]

SWITCHING BETWEEN MODES

Discharge switching from the *L* to the *H* mode and back, improved energy confinement, and control of the density profile using pellet injection have been successfully demonstrated in numerous tokamak experiments.[31–33] In the TFTR tokamak for instance, pellet injection was used to transfer the discharge temporarily from the *L* to the *H* mode. The opposite transition was achieved by peripheral injection of deuterium–lithium pellets in Ohmic regimes in the Tuman-3 tokamak. Peripheral injection of hydrogen pellets in the T-10 tokamak in a regime with high-power electron cyclotron heating also caused a transition to the *H* mode. Pellet-stimulated peripheral instabilities (known as ELM) were observed in ASDEX-Upgrade.

CONCLUSIONS

Fusion reactor technologies based on injecting pellets into the plasma are now being successfully developed and widely used in existing tokamak and stellarator experiments. These technologies can be used for internal fuel injection, measurements of plasma parameters, switching between discharge burning regimes, rapid quenching of the plasma and removal of the current in emergency situations, and for conditioning the first wall the chamber. The technical specifications of these injection systems are already close to those required to operate under reactor conditions. These technologies will quite clearly play an important role in the operation of a future fusion reactor.

[1] S. L. Milora, W. A. Houlberg, L. L. Lengyel, and V. Mertens, Nucl. Fusion **35**, 657 (1995).
[2] S. K. Combs, Rev. Sci. Instrum. **64**, 1679 (1993).
[3] B. V. Kuteev, A. P. Umov, I. V. Viniar *et al.*, Plasma Devices Op. **2**, 193 (1994).
[4] B. V. Kuteev, Fusion Technol. **26**, 938 (1994).
[5] B. V. Kuteev, Nucl. Fusion **35**, 431 (1995).
[6] L. Garzotti, B. Pegourie, A. Geraud *et al.*, Nucl. Fusion **37**, 1167 (1997).
[7] L. R. Baylor, A. Geraud, W. A. Houlberg *et al.*, Nucl. Fusion **37**, 445 (1997).
[8] H. Nakamura and ITER JCT, *Fuelling Detail Design Description* (ITER, Garching, 1997).
[9] R. Yoshino, T. Kondoh, Y. Neyatani *et al.*, Plasma Phys. Controlled Fusion **39**, 313 (1997).
[10] P. B. Parks and R. J. Turnbull, Phys. Fluids **21**, 1735 (1978).
[11] W. A. Houlberg, S. L. Milora, and S. E. Attenberger, Nucl. Fusion **28**, 595 (1988).
[12] B. V. Kuteev and A. Yu. Kostrukov, *Europhysics Conference Abstracts*, Vol. 19C. Pt. 1., pp. 1–121 (1995).
[13] S. K. Combs, S. L. Milora, L. R. Baylor *et al.*, Fusion Technol. 1153 (1996).
[14] M. J. Gouge, S. A. Combs, P. W. Fisher *et al.*, *ITER Plasma Fueling Development*, Report MJG-ITER 3-97.
[15] I. V. Vinyar, S. V. Skoblikov and P. Yu. Koblents, Pis'ma Zh. Tekh. Fiz. **23**(20), 43 (1997), [Tech. Phys. Lett. **23**(10), 793 (1997)].
[16] I. V. Viniar, S. V. Skoblikov, P. Yu. Koblents *et al.*, *Fusion Technology*, edited by Varandas and F. Serra (1996), p. 1137; I. V. Viniar, S. V. Skoblikov, and P. Yu. Koblents, Zh. Tekh. Fiz. **68**(5), 117 (1998) [Tech. Phys. **43**, 588 (1998)].
[17] I. V. Viniar, S. Sudo, S. V. Skoblikov *et al.*, J. Plasma Fusion Res. **1**, 414 (1998).

[18] I. V. Viniar and S. Sudo, Rev. Sci. Instrum. **68**, 1444 (1997).
[19] P. T. Lang, K. Buechl, M. Kaufmann *et al.*, *Europhysics Conference Abstracts*, Vol. 21A, p. 1481 (1997).
[20] L. R. Baylor, T. C. Jernigan, T. Carstrom *et al.*, *Europhysics Conference Abstracts*, Vol. 19C. Part 1, pp. 1–113 (1995).
[21] P. T. Lang, P. Cierpka, S. M. Egorov *et al.*, Fusion Technol. **27**, 653 (1995).
[22] H. W. Mueller, P. T. Lang, K. Buechl *et al.*, Rev. Sci. Instrum. **68**, 4051 (1997).
[23] V. Yu. Sergeev, A. Yu. Kostrukov, and S. A. Shibaev, Fusion Eng. Des. **34**, 323 (1997).
[24] V. Yu. Sergeev and D. A. Polivaev, Fusion Eng. Des. **34**, 215 (1997).
[25] P. Mantica, M. R. De Baar, J. de Kloe *et al.*, *Europhysics Conference Abstracts*, Vol. 21A, p. 1853 (1997).
[26] M. H. Redi, S. H. Batha, R. V. Budny *et al.*, *Europhysics Conference Abstracts*, Vol. 21A, p. 1069 (1997).
[27] B. V. Kuteev, V. Yu. Sergeev, and S. Sudo, Nucl. Fusion **35**, 1167 (1995).
[28] V. Yu. Sergeev, B. V. Kuteev, S. G. Kalmykov *et al.*, *Europhysics Conference Abstracts*, Vol. 19C, Part 3, p. IV-057 (1995).
[29] J. D. Strachan, M. Bell, M. Bitter *et al.*, Nucl. Fusion **33**, 991 (1993).
[30] K. Kondo, J. Nucl. Mater. **241**, 956 (1997).
[31] L. G. Askinazi, V. E. Golant, S. V. Lebedev *et al.*, Phys. Fluids B **5**, 2420 (1993).
[32] V. G. Kapralov, K. V. Khlopenkov, and V. A. Rozhansky, Pis'ma Zh. Tekh. Fiz. **21**(2), 31 (1995) [Tech. Phys. Lett. **21**, 57 (1995)].
[33] V. G. Kapralov, V. A. Rozhansky, and K. V. Khlopenkov, *Europhysics Conference Abstracts*, Vol. 19C, Part 1, pp. 1–117 (1995)

Translated by R. M. Durham

Visualization of the atomic structure of the subsurface region of a solid

I. I. Pronin, M. V. Gromoyunova, N. S. Faradzhev, and D. A. Valdaĭtsev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 68–71 (September 1999)

A new variant is proposed for a method of visualizing the atomic structure of the subsurface region of a solid, based on the focusing of medium-energy reflected electrons and suitable for studying atomic processes at the surface in real time. An original device developed for this purpose is briefly described. The possibilities of the method are illustrated by examples of studying the atomic structure of epitaxial silver films grown on the surface of an Si(100) single crystal and also in the subsurface region of a layered $VSe_2(0001)$ crystal. © *1999 American Institute of Physics.* [S1063-7842(99)01509-3]

Various surface effects, widely used in technology, are of enormous importance in the modern world. The processes taking place at surfaces are highly complex, and information on the microscopic level is required to understand them.[1] The main methods used for diagnostics of surface atomic structure are low-energy ($\sim$100 eV) electron diffraction (LEED), high-energy ($\sim$50 keV, glancing incidence) electron diffraction (HEED), and also scanning tunneling microscopy (STM). Virtually no studies of adsorption systems or the initial stages of thin film growth could be made without these. In contrast to these methods, medium-energy electron diffraction (MEED) at average energies of $\sim$1 keV had until recently been very little used in the structural analysis of surfaces. This was because the mechanism for the formation of the patterns had not been studied in depth and also because, in this energy range, particular attention had been focused on problems of quantitative elemental analysis of surfaces using methods of electron spectroscopy.

Considerable progress in understanding the mechanism responsible for the diffraction scattering of medium-energy electrons was achieved with the development of x-ray photoelectron and Auger electron diffraction techniques.[2–4] Focusing of electrons emitted by ''internal'' sources was shown to play an important role here. It is also well known that when the electron kinetic energies are similar, the diffraction patterns produced by reflected electrons, Auger electrons, and photoelectrons are very similar.[5,6] It was therefore natural to assume that this effect also predominates in the diffraction of reflected electrons. In order to check out this hypothesis, we carried out a wide range of experiments on samples and compared the results with model calculations.[7–9] It was ultimately found that this effect not only predominates in the formation of the diffraction patterns of quasielastically scattered electrons but the patterns themselves also allow a very simple interpretation: their principal peaks occur when electrons are emitted along close-packed strings of atoms. This specific feature of the spatial distributions of the electrons led to the development of a new method of visualizing the atomic structure of a thin subsurface region of a solid, which has been successfully tested on various adsorption systems and multicomponent compounds.[10–12] One of its main advantages is that atomic processes on a surface can be visualized in real time. Here we report the first results obtained using an original device developed by us for these investigations.

The experiments were carried out in ultrahigh vacuum using samples having an atomically clean surface. Unlike the initial variant of the method,[7] where the diffraction patterns were recorded by scanning the entire electron reflection hemisphere using a narrow-aperture energy analyzer, here we use a small electron-diffraction camera fitted with a source of 1.5–3 keV electrons, which are incident on the crystal surface at a glancing angle. The electrons bombarding the sample are scattered intensively in the subsurface layer and are partially reflected, undergoing energy losses. The contrast interference pattern formed only by the fast reflected electrons is obtained using a retarding-field energy analyzer with spherical grids. The flux of transmitted electrons is amplified by a microchannel plate and causes the phosphor to luminesce, which is observed through an optical window in the vacuum chamber. The electron diffraction camera collects electrons in a fairly wide solid angle (in a cone with an opening half angle of 57°). It is thus possible to observe the important region near the normal to the surface of the sample, which is inaccessible to the commercial devices used in low-energy electron diffraction. The diffraction patterns are recorded using a video camera combined with a computer which can write them to hard disk at a rate of up to 50 frames per second and then process them using standard application programs.

We shall illustrate the possibilities of this new variant by means of two examples. The first is a further development of research on the interaction between silver and a silicon surface. Figure 1 shows several diffraction patterns obtained during the deposition of silver on an Si(100) 2×1 single crystal heated to 100 °C. The data are represented as two-dimensional distribution patterns of the intensity of electron reflection over the polar and azimuthal angles of emission, constructed using a stereographic projection. The light areas correspond to maxima in the angular distributions and the black areas correspond to minima. A comparison of Figs. 1a and 1e shows that the diffraction pattern from the initial sili-

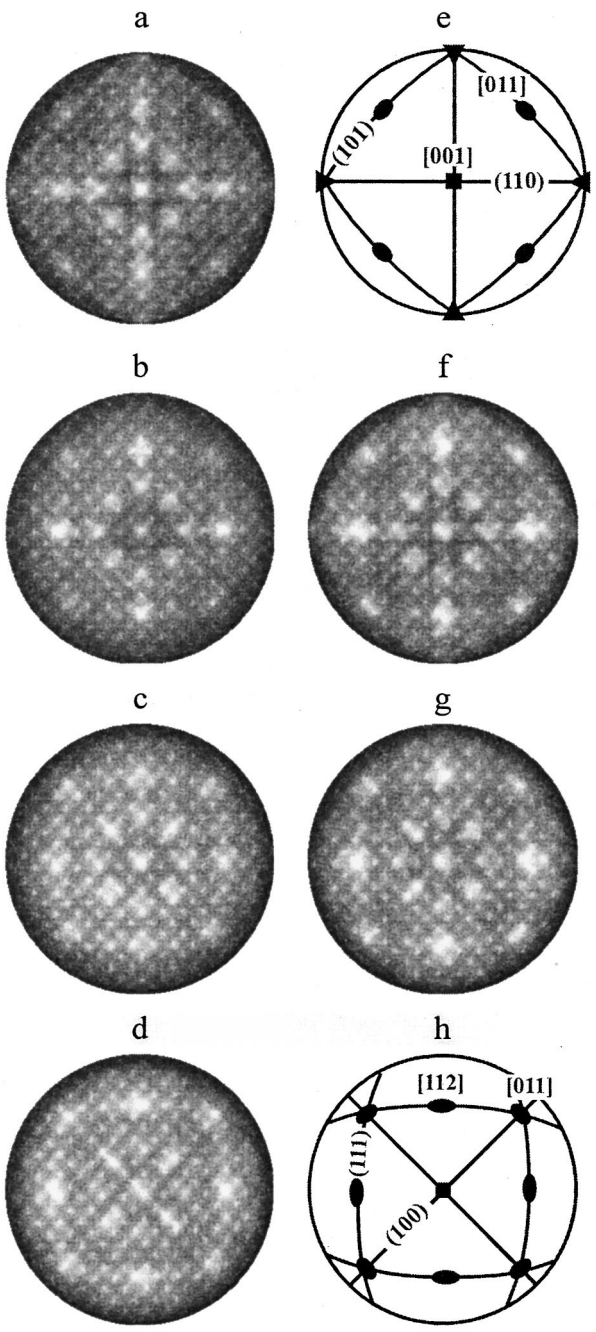

FIG. 1. a–d — Diffraction patterns of 2 keV reflected electrons observed during deposition of silver on the Si(100) face: a — silicon with a clean surface, b–d — silver films having 20%, 40%, and 90% continuity, respectively, e,h — diagrams of most closely packed planes and directions of the (001) faces of silicon and silver; f,g — results of a numerical modeling of the patterns in Figs. 1b and 1c.

con sample with a clean surface exhibits symmetry characteristic of the (100) face. Its principal maxima are in fact observed when the electrons are emitted along close-packed strings of silicon crystal atoms. For instance, the bright spot at the center of the pattern is caused by focusing of electrons propagating in the $\langle 100 \rangle$ direction. Also clearly visible are extended regions of increased electron emission (Kikuchi bands) along the projections of these close-packed {100} planes in the silicon. They are made up of numerous focusing maxima along the interatomic directions assigned to these planes.[9] Hence, this pattern actually reflects in real space the relative position of the atoms in the subsurface layer, whose thickness is equal to the average depth of emission of the reflected electrons which in this case is 15 Å.

The growth process of the deposited film is seen clearly from the diffraction patterns (Figs. 1b–1d). As the coverage increases, changes gradually appear in this pattern and are seen particularly clearly when the data are compared for the substrate and a comparatively thick film (Fig. 1d) of average thickness 20 Å, when the electrons are predominantly scattered in the layer of condensed silver.[1)] Deposition does not change the fourfold symmetry of the pattern but causes appreciable redistribution of its intensity in various regions. Whereas for Si(100) the Kikuchi bands along the projections of the {100} planes are seen most clearly, in the pattern from the deposited film, the bands along the {100} and {111} planes predominate (Fig. 1h). Since these atomic planes are the most closely packed in the face-centered crystal lattice of silver, we can immediately conclude that the as-grown film consists of crystallites having a face-centered crystal structure with the Ag(100) face emerging on their surface. Moreover, the azimuthal orientation of the crystallites is strictly matched with the substrate $\langle 100 \rangle_{\mathrm{Ag}} \| \langle 100 \rangle_{\mathrm{Si}}$. The validity of this conclusion is confirmed by the results of modeling the pattern in Fig. 1d, made using the single-scattering cluster approximation.[8] It should be stressed that this pattern differs qualitatively from that observed by us previously when visualizing the structure of a silver film grown under the same conditions on the other Si(111) 7×7 silicon face. In an Ag/Si(111) system two types of Ag(111) domains, azimuthally misoriented by 180°, coexist on the surface.[12] Thus, a comparison between the results obtained for the two principal silicon faces clearly shows that the atomic geometry of the substrate surface has a decisive influence on the epitaxial growth of the silicon film.

Returning to the patterns observed at the intermediate stages of deposition (Figs. 1b and 1c), we must note that since the film growth in this system takes place in an island mode,[13] both the substrate and the various silver islands contribute to these images. Consequently, each of these is a superposition of the patterns in Figs. 1a and 1d taken with weighting factors which reflect the condensate coverage of the silicon. The validity of this interpretation of the diffraction patterns is convincingly demonstrated by the results of computer modeling using a method similar to that described in Ref. 14. The intensity of the electron reflection in an arbitrary direction was calculated by simply summing the intensities of the reflection from the clean substrate and the thick silver film. The results of the calculations are plotted in Figs. 1f and 1g and these show very good agreement with the experiment. Note that this method can also be used to determine the fraction of the substrate surface coated with the deposited silver film. This can be accomplished by varying the coverage and minimizing the difference between the calculated and the observed pattern. This technique can naturally also be applied to other systems. Hence, numerical modeling of diffraction patterns measured in real time is a

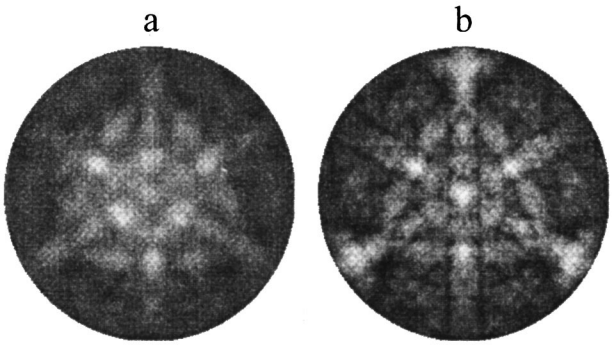

FIG. 2. Diffraction patterns of a $VSe_2(0001)$ single crystal: a — 2 keV experiment; b — calculations using single-scattering cluster approximation for an interplanar distance of 1.57 Å between the layers of V and Se atoms.

highly promising method of determining the phase composition of island films.

A second example illustrating the possibilities of this method of visualization involves studying transition metal dichalcogenides. A typical representative of this type of compound is VSe_2 which possesses clearly defined anisotropic properties and is of particular interest from the point of view of understanding their electronic structure.[15]

The starting point to calculate the energy band structure of a material is to define the relative position of the atoms in the crystal unit cell. The lattice constants of the VSe_2 crystal are known, but the atomic structure of the subsurface region of this compound has not been sufficiently studied. In particular, the distance between the layers in which the V and Se atoms are localized is unknown. A simple experiment was carried out to determine this unknown. A single-crystal $VSe_2(0001)$ sample was cleaved in the vacuum chamber of the device and the diffraction pattern was obtained almost immediately afterward (Fig. 2a). The measurements were made under the same conditions as for the silver films on silicon considered above. In this case, a distinct pattern appears, but already exhibits threefold symmetry, typical of the (0001) face of the CdI_2 type crystal structure in which VSe_2 crystallizes. This reaffirms that the method directly visualizes the crystal structure of the subsurface region of the sample.

The pattern observed was then compared with the results of model calculations made in the single scattering approximation for VSe_2 clusters with different structural parameters. By using *R*-factors which describe the deviation of the calculated patterns from the measured one, we obtained a fairly accurate determination of the unknown distance *d* between the V and Se layers. The best agreement between the results of the modeling and the experiment was observed for the case $d = 1.57$ Å. In this case, the lattice constants *a* and *c* were as in the bulk: 3.35 and 6.12 Å. Figure 2b shows the calculated pattern for these values of the parameters. This pattern faithfully reproduces all the main features of the observed pattern and the value of the R_1 factor is 0.095.

To conclude, these results show that this method of visualizing the atomic structure is not only suitable for qualitative investigations but can also be used to obtain quantitative information on objects. This method can be used to study in real time the atomic processes taking place in the subsurface region of a solid. It has the advantages that data can be obtained in direct space (rather than in inverse space as in conventional methods of high-energy and low-energy electron diffraction) and objects which do not possess far-range order can be investigated.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-16909) and by the Russian State Program ''Surface Atomic Structures'' (Project No. 95-1.21).

[1)]Results obtained using Auger spectroscopy showed that this film is not continuous and occupies around 90% of the substrate area.

[1] C. B. Duce, *Surface Science: The First Thirty Years* (North-Holland, Amsterdam, 1994), 1064 pp.
[2] W. F. Egelhoff Jr., Crit. Rev. Solid State Mater. Sci. **16**(3), 213 (1990).
[3] C. S. Fadley, in *Synchrotron Radiation Research*, Advances in Surface and Interface Science, Vol. 1, Techniques, edited by R. Z. Bachrach (Plenum Press, New York, 1992), pp. 421–518.
[4] S. A. Chambers, Surf. Sci. Rep. **16**(6), 261 (1992).
[5] H. Hilferink, E. Lang, and K. Heinz, Surf. Sci. **93**, 398 (1980).
[6] M. V. Gomoyunova and I. I. Pronin, Poverkhnost', No. 7, 44 (1982).
[7] N. S. Faradzhev, M. V. Gomoyunova, and I. I. Pronin, Phys. Low-Dimens. Semicond. Struct. No. 9, 11 (1994).
[8] M. V. Gomoyunova, I. I. Pronin, and N. S. Faradzhev, Zh. Éksp. Teor. Fiz. **110**, 311 (1996) [JETP **83**, 168 (1996)].
[9] I. I. Pronin, N. S. Faradzhev, and M. V. Gomoyunova, Fiz. Tverd. Tela (St. Petersburg) **39**, 752 (1997) [Phys. Solid State **39**, 666 (1997)].
[10] M. Erbudak, M. Hochstrasser, and E. Wetli, Mod. Phys. Lett. B **8**, 1759 (1994).
[11] N. S. Faradzhev, M. V. Gomoyunova, J. Osterwalder *et al.*, Surf. Sci. **331–333**, 1446 (1995).
[12] N. S. Faradzhev, M. V. Gomoyunova, and I. I. Pronin, Phys. Low-Dimens. Semicond. Struct. **3**, 93 (1997).
[13] V. G. Lifshits *et al.*, *Surface Phases on Silicon. Preparation, Structure and Properties* (Wiley, New York, 1994).
[14] M. V. Gomoyunova, I. I. Pronin, and N. S. Faradzhev, Zh. Tekh. Fiz. **67**(12), 62 (1997) [Tech. Phys. **42**, 1381 (1997)].
[15] V. N. Strocov, H. I. Starnberg, P. O. Nilsson, and L. J. Holleboom, J. Phys. **8**, 7549 (1996).

Translated by R. M. Durham

Intercalation of two-dimensional graphite films on metals by atoms and molecules

N. R. Gall', E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 72–75 (September 1999)

An analysis is made of some general laws governing a new physical effect, i.e., the spontaneous penetration of particles (atoms, C_{60} molecules) adsorbed on a two-dimensional graphite film on a metal (Ir, Re, Pt, Mo, ...) to beneath the graphite film (intercalation). It is shown that atoms having low ionization potentials (Cs, K, Na) intercalate a two-dimensional graphite film on iridium at $T=300-400$ K with an efficiency $\varkappa\approx 0.5$, accumulating beneath the film to a concentration of up to a monolayer. Atoms having high ionization potentials (Si, Pt, Ni, C, Mo, etc.) intercalate a two-dimensional graphite film on iridium at $T\approx 1000$ K with an efficiency $\varkappa\approx 1$, forming beneath the film a thick intercalate layer which is strongly bonded chemically to the metal substrate but is probably weakly bonded to the graphite monolayer by van der Waals forces. The presence of a graphite ''lid'' impeding the escape of atoms from the intercalated state up to record high temperatures $T\sim 2000$ K leads to superefficient diffusion (with an efficiency close to one) of various atoms (Cs, K) into the bulk of the substrate (Re, Ir). © *1999 American Institute of Physics.* [S1063-7842(99)01609-8]

Two-dimensional graphite films on the surface of metals are frequently encountered in various fields of science and technology. These comprise a carbon mesh of strongly bound carbon atoms, structurally equivalent to a single layer in a graphite single crystal.[1] It has now been established that such a film formed on many metals (Ir, Re, Pt, Ni, Rh, Ru, Mo, Pd) of various types and crystal geometry is extremely nonreactive and is bound to the metal surface only by weak van der Waals forces.[2–4] Its work function is $e_\varphi=4.45$ eV and does not depend on the nature and crystal geometry of the substrate.[4]

The formation of two-dimensional graphite films on metals leads to various remarkable effects. For instance, these films completely poison metal catalysts in dissociative heterogeneous catalysis.[2] Even at room temperature, molecules of various gases (CO_2, H_2, N_2, O_2) and even polyvalent atoms of transition metals (Au, Cu, Pt, Ni) are readily desorbed from the surface.[2–4]

Thus two-dimensional graphite films exhibit unique adsorption, chemical, and catalytic passivity combined with extremely high thermal stability. For instance, on (111)Ir these films break down at $T\geqslant 1900$ K (Ref. 5), and on (1010)Re they break down at $T=1600-1900$ K, depending on the conditions of formation.[4]

In the present paper we consider the laws governing another very interesting physical effect which is a consequence of the weak van der Waals coupling between the graphite film and the metal surface. This is intercalation, i.e., the spontaneous penetration of particles adsorbed on a two-dimensional graphite film to beneath the graphite film. Intercalation was observed in work carried out by our group at the A. F. Ioffe Institute,[6,7] and fundamental studies were then reported.[2–4]

EXPERIMENTAL METHOD

We used a combination of various complementary surface diagnostic techniques: high-resolution Auger electron spectroscopy ($\Delta E/E\sim 0.1\%$), thermal desorption spectroscopy, measurement of the work function using surface ionization and thermionic emission, and molecular beam probing of the surface. The samples were thin metal ribbons (Ir, Re, Mo, Pt, etc.) measuring $40\times 1\times 0.02$ mm which were textured and cleaned of impurities by a standard technique.[5] As a result, in practice a single corresponding face: (111)Ir, (1010)Re, (100)Mo, (111)Pt appeared at the surface of the ribbons. An Auger spectrometer with a prism energy analyzer recorded the Auger spectra directly at the strongly heated metal ribbons ($T\leqslant 2500$ K).

A two-dimensional graphite film was formed on the metal surface by holding the heated ribbons in benzene (C_6H_6) vapor.[4] Films obtained in this way ''automatically'' turn out to be two-dimensional and continuous, since on the graphite islands the benzene molecules do not dissociate (they are desorbed) but on the active metal surface they dissociate completely, with the carbon accumulating in the adlayer and the hydrogen being desorbed.

Various experimental evidence suggests that these graphite films are in fact two-dimensional, continuous, and uniform in terms of work function.[8,9] In reality the graphite monolayer consists of a set of coalesced graphite islands. Their concentration for a monolayer coating on (111) iridium is estimated as $\sim 5\times 10^7$ per cm^{-2} and the average island radius is $\sim 10^4$ Å (Ref. 3).

It is reasonable to expect that the edges of the graphite islands form a surface defect. Their relative area turned out to be extremely small, accounting for $\sim 0.1\%$ of the total surface area.[9]

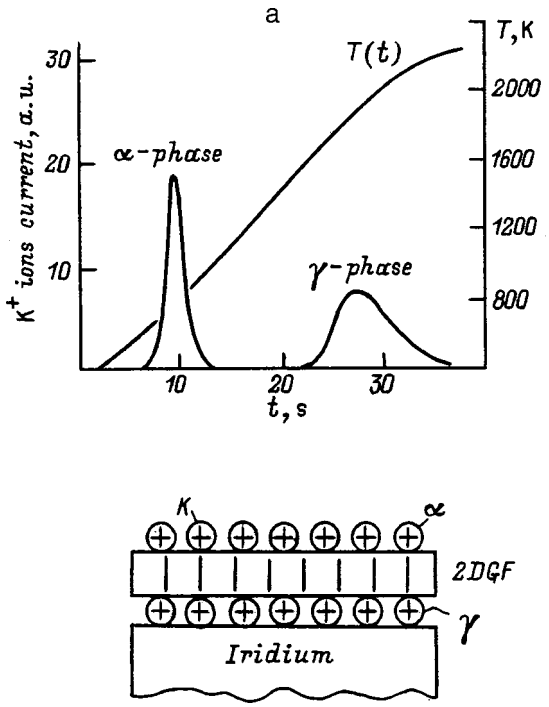

FIG. 1. a — Thermal desorption spectra of potassium ions after adsorption of potassium: flux $\nu_K = 1\times10^{11}\,\mathrm{cm^{-2}s^{-1}}$ for 40 s, α and γ desorption phases; b — structure of intercalated film when alkali metal atoms are adsorbed to saturation on a two-dimensional graphite film on iridium: K^+ — positively charged alkali metal adatom; *2*DGF — two-dimensional graphite film.

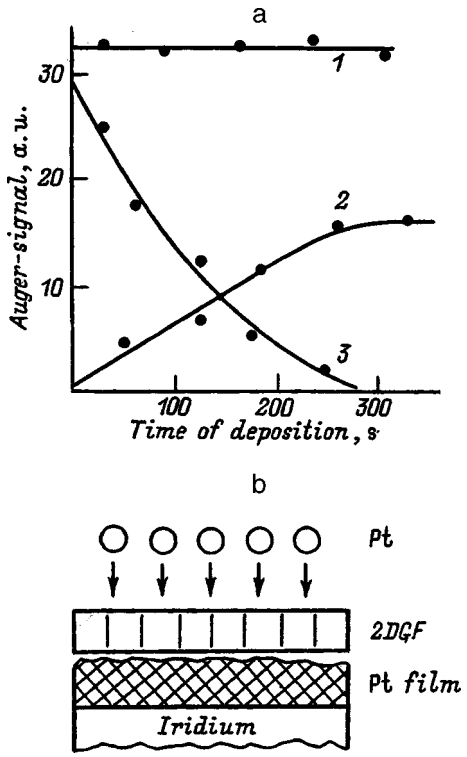

FIG. 2. a — Intensities of the Auger peaks of carbon (*1*), platinum (*2*), and iridium (*3*) as a function of the time of deposition of Pt atoms on a two-dimensional graphite film on (111) iridium at $T = 1200$ K with the flux $\nu_{Pt} = 1\times10^{13}\,\mathrm{cm^{-2}s^{-1}}$; b — structure of intercalated film when Pt atoms are adsorbed on a two-dimensional graphite film on iridium: Pt — platinum atom, Pt film — thick platinum film.

RESULTS AND DISCUSSION

1. Intercalation by low-ionization-potential atoms such as Cs, K, Na (Refs. 3 and 4)

Intercalation beneath two-dimensional graphite films on metals by alkali metal atoms is observed at $T \leqslant 700$ K with a maximum diffusion efficiency of $\sim 50\%$ when these atoms are deposited in the temperature range 300–400 K. Coulomb repulsion of the adatoms plays a decisive role in the intercalation behavior of these atoms which, on a graphite monolayer, are positively charged and strongly bound to the layer by image forces.[10]

If alkali metal atoms (such as K) are adsorbed to saturation on a two-dimensional graphite film on (111)Ir at $T = 300$ K and a temperature flash is applied, two desorption phases are observed (Fig. 1a). The low-temperature α-phase ($T \sim 800$ K) corresponds to the desorption of potassium adatoms from the graphite surface (Figs. 1a and 1b). The second, completely unusual, γ-phase corresponds to potassium desorbed at record high temperatures $T \geqslant 1900$ K. The high-temperature γ-phase is formed by potassium located beneath the two-dimensional graphite film in the intercalated state (Fig. 1b). The intercalated potassium (and also sodium and cesium) is only released when damage occurs at the edges of the graphite islands and for (111)Ir this is at ~ 1800–1900 K. Note that the maximum concentrations of alkali metal adatoms in the α and γ phases are approximately the same and correspond to a near-monolayer concentration: for example, when cesium is adsorbed on a two-dimensional graphite film on (111)Ir, this concentration is $N_\alpha \approx N_\gamma \approx 3-4\times10^{14}\,\mathrm{atom/cm^2}$ (Ref. 11).

2. Intercalation by high-ionization-potential atoms such as Pt, Ni, C, Mo, Si, etc. [3,4]

For atoms having high ionization potentials, which are not charged on the graphite surface and are weakly bound to it by van der Waals polarization forces, the temperature range of efficient intercalation is shifted toward moderate and high temperatures ($T \geqslant$ 900–1500 K). At lower temperatures $T < 900$ K, polyvalent atoms readily form two- or three-dimensional islands on the graphite film, impeding the penetration of incoming atoms beneath the graphite layer and the adsorbate islands act as traps.

By way of example we shall consider the adsorption of Pt atoms on a two-dimensional graphite film on (111)Ir (Figs. 2a and 2b).[12] At $T = 1000$ K the diffusion efficiency of Pt atoms reaching the surface beneath the graphite film is $\varkappa \approx 1$. This implies that at these temperatures each atom impinging on the graphite monolayer penetrates beneath it into the intercalated state. At $T > 1000$ K some of the Pt adatoms are desorbed from the film, the intercalation efficiency drops ($\varkappa < 1$) and at $T = 1500$ K is $\varkappa \approx 0.1$.

The condition $\varkappa = 1$ implies that before it ha time to desorb, a Pt adatom can migrate from the center of a graphite island to its edge where it is transferred to the intercalated state.

It is extremely interesting to note that, unlike alkali metal atoms which only form a monolayer film in the intercalated state, atoms with high ionization potentials form a thick film beneath the two-dimensional graphite film, which is strongly bound to the metal substrate; the other side of the film is weakly bound to the graphite monolayer, probably by van der Waals forces (Fig. 2b). The presence of a thick adsorbate film is confirmed by electron Auger spectroscopy. In the above example, the Auger signal intensity of the iridium substrate falls to the noise level as a result of its screening by the growing platinum film, while the Auger signal intensity of the carbon remains constant (Fig. 2a).

It was established experimentally that not all polyvalent atoms form a ''pure'' thick intercalate film beneath a two-dimensional graphite film. For example, when Si atoms are adsorbed beneath a two-dimensional graphite film at $T \approx 900-1300$ K, bulk metal silicides are formed with the graphite monolayer remaining on their surface.[12]

3. Intercalation by C_{60} molecules.

Experiments have shown that two-dimensional carbon films on metals efficiently intercalate not only many atoms but also C_{60} fullerene molecules. A transition takes place beneath the graphite layer when a thick (3–4 ML) film of C_{60} molecules deposited at $T \sim 300$ K on a two-dimensional graphite film on (111)Ir is annealed at $T \sim$ 800–900 K (Ref. 13). For lower concentrations of initially deposited fullerene molecules, no intercalation takes place. The limiting concentration of C_{60} molecules observed in the intercalated state is close 1 ML ($N_{C_{60}} \sim 2 \times 10^{14}$ mol/cm^2).

4. Superefficient diffusion

The existence of a graphite lid impeding the escape of atoms from the intercalated state leads to another extremely interesting effect. Intercalate atoms in close contact with the metal substrate for a long time at high T may undergo superefficient diffusion inside the metal. Conditions may be established when, for example, atoms such as K and Cs diffuse into the bulk of rhenium and iridium with 30% efficiency.[14,15] Without this graphite lid the diffusion efficiency for these atoms falls by approximately ten orders of magnitude and it is impossible for them to penetrate into the metal.

5. Mechanism for intercalation of a two-dimensional graphite film

A thermal ''bubble'' mechanism was proposed in Ref. 16 to explain the intercalation of two-dimensional graphite films by atoms. If a graphite island on a metal is bombarded with particles, these migrate over the top of the island and are adsorbed at its edge in the starting position for intercalation. Intercalation begins when a bubble forms at this point on the graphite island. This bubble formation is based on allowance for the random thermal motion of the atoms when each atom of the graphite island, undergoing thermal motion may at some time acquire sufficient kinetic energy at its bond with the metal substrate, such that as it moves away from the metal surface it can break this bond and those of various neighboring carbon atoms. This results in the formation of a bubble at this point on the graphite island. Bubbles form at any point on the island, have a certain lifetime, and then decay. The surface of the graphite layer can be likened to the surface of the sea in stormy weather. We stress that the transport of the intercalated particles can be reduced to their migration inside bubbles over the metal and the two-dimensional graphite film.

CONCLUSIONS

Two-dimensional graphite films on metals behave as anisotropic selective diffusion barriers transmitting or not transmitting specific atomic particles depending on the deposition conditions and subsequent heat treatment. It is highly probable that two-dimensional graphite films on metals intercalated with atoms and molecules may prove to be some of the most interesting nanometer-size objects for modern micro- and optoelectronics devices.

This work was supported by the Russian Federation State Program ''Physics of Solid-State Nanostructures,'' Project No. 2-025/4.

[1] A. R. Ubbelohde and F. A. Lewis, *Graphite and Crystal Compounds* (Clarendon Press, Oxford, 1960), p. 264.
[2] A. Ya. Tontegode, Prog. Surf. Sci. **38**, 201 (1991).
[3] A. Ya. Tontegode and E. V. Rut'kov, Usp. Fiz. Nauk **163**(11), 57 (1993).
[4] N. R. Gall, E. V. Rut'kov, and A. Ya. Tontegode, Int. J. Mod. Phys. B **11**, 1865 (1997).
[5] N. R. Gall', M. S. Kobrin, S. N. Mikhaĭlov, E. V. Rut'kov, and A. Ya. Tontegode, Poverkhnost', No. 9, 21 (1984).
[6] E. V. Rut'kov and A. Ya. Tontegode, Pis'ma Zh. Tekh. Fiz. **7**, 1122 (1981) [Sov. Tech. Phys. Lett. **7**, 480 (1981)].
[7] E. V. Rut'kov and A. Ya. Tontegode, Zh. Tekh. Fiz. **52**, 921 (1982) [Sov. Phys. Tech. Phys. **27**, 589 (1982)].
[8] N. A. Kholin, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **139**, 155 (1984).
[9] E. V. Rut'kov and A. Ya. Tontegode, Surf. Sci. **161**, 373 (1985).
[10] N. R. Gall, S. N. Mikhailov, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **276**, 381 (1990).
[11] N. R. Gall, S. N. Mikhailov, E. V. Rut'kov, and A. Ya. Tontegode, Synth. Met. **34**, 447 (1989).
[12] N. R. Gall', E. V. Rut'kov, and A. Ya. Tondegode, Pis'ma Zh. Tekh. Fiz. **14**, 527 (1988) [Sov. Tech. Phys. Lett. **14**, 235 (1988)].
[13] E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Phys. Rev. Lett. **74**, 758 (1995).
[14] A. Ya. Tontegode and F. K. Yusifov, Fiz. Tverd. Tela (St. Petersburg) **35**, 987 (1993) [Phys. Solid State **35**, 506 (1993)].
[15] A. Ya. Tontegode and F. K. Yusifov, Appl. Surf. Sci. **90**, 185 (1995).
[16] A. Ya. Tontegode, Pis'ma Zh. Tekh. Fiz. **15**(7), 57 (1989) [Sov. Tech. Phys. Lett. **15**(4), 271 (1989)].

Translated by R. M. Durham

Nonlinear energy-selective nanoscale modifications of materials and dynamics in metals and semiconductors

S. Marka, C. Parks Cheney, W. Wang, G. Lüpke, J. Gilligan, Y. Yao, and N. H. Tolk

Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 76–80 (September 1999)

Studies of nonlinear, energy-selective material interactions localized at surfaces, heterointerfaces, impurities, and defects in solids are described. Particular reference is made to a new molecular interaction effect caused by transfer of surface energy by low-energy collisions, a new noncontact nonlinear optical method of studying electron and hole dynamics at a heterointerface, and a new approach using a free-electron laser developed at Vanderbilt University to activate hydrogen-passivated dopants in silicon. In each case the unique characteristics of particle and photon beams, optimized for the technology, were used to extend the range of applications of these new energy-selective techniques to solve fundamental and applied problems. © *1999 American Institute of Physics.* [S1063-7842(99)01709-2]

INTRODUCTION

Studies of dynamic processes at surfaces, heterointerfaces, and in the bulk have achieved considerable progress at the end of the second millennium following the appearance of new, and technologically more refined, nanosize probes and material processing technologies. This research frequently leads to new physical discoveries and useful applications. In particular in the present paper we report results of research which extends the range of sputtering yield measurements toward very low threshold energies and provides new information on processes of nonlinear energy transfer from molecular and atomic ions to surfaces. We also show that the nondestructive method of optical second-harmonic generation is an important tool for studying Si/SiO_2 heterointerfaces, providing new information on the dynamics of electrons and holes at a heterointerface. Finally, high-intensity, tunable, infrared radiation obtained using the W. M. Keck free-electron laser at Vanderbilt University was used for wavelength-selective excitation of vibrational modes localized at specific impurities and defects. This method offers a fundamentally new approach to resonant nanosize modification of materials, which is important from both the theoretical and practical point of view.

NEW MOLECULAR INTERACTION EFFECT OBSERVED IN THE SPUTTERING YIELD NEAR THRESHOLD ENERGIES

Measurements of the sputtering yield provide exceptional possibilities for studying the interaction between primary particles and a surface, for example during epitaxial growth or etching. We recently observed a new molecular effect which provides a deeper understanding of processes of energy transfer to a surface at low projectile kinetic energies.[1] A comparison of the sputtering yield per incident atom for low-energy molecular projectiles (such as N_2^+ and O_2^+) relative to the corresponding atomic primary particles (N^+ and O^+) revealed that at energies below the characteristic initial energy of the molecular projectiles, the sputtering yield per incident atom increases appreciably for incoming molecular ions. Figure 1 gives the sputtering yield per atom for N_2^+ and N^+ ions. All the experiments were carried out in an ultrahigh-vacuum chamber and the sputtering yield was obtained by measuring the number of incoming ions (from the current at an 80% transmission mesh) and measuring the mass change of a polycrystalline gold sample (using a quartz microbalance).[2] Figure 1 shows that the sputtering yield for N_2^+ increases appreciably at low kinetic energies, being approximately four times higher when the kinetic energy of the projectiles is 50 eV. We attribute this increase to the larger quantity of energy transferred by the molecular projectiles per atom compared with atomic projectiles. Assuming that the collision is entirely inelastic, we can easily show that the maximum energy transferred to the target is considerably higher for molecules of mass $2m$ compared with two atoms each of mass m (this value is 1.76 for nitrogen molecules and a gold target). This reasoning only holds for low projectile kinetic energies, when the molecules have sufficient time to

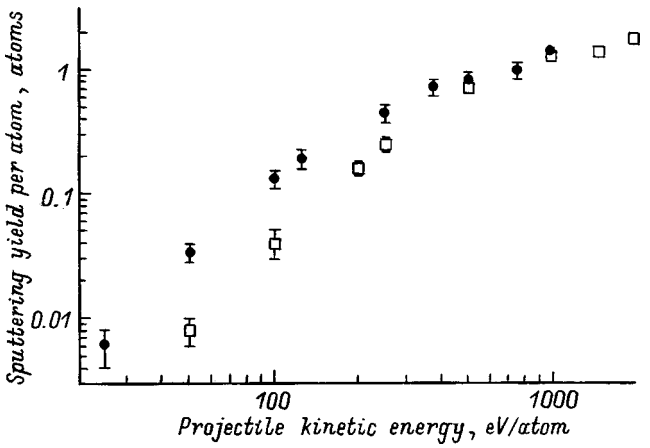

FIG. 1. Sputtering yield per atom for N^+ (□) and N_2^+ (●) ions on a polycrystalline gold target as a function of the primary particle kinetic energy.

interact with the surface. During this collisional interaction time the molecules may undergo several vibrations, whereas at high kinetic energies the collision time becomes so small compared with the period of the vibrations that the molecule interacts with the surface as if it consisted of two separate atoms. Consequently, this new molecular effect is only observed at a specific threshold kinetic energy corresponding to the case when the collision time and the vibrational period are equal. For N_2 (we assume that the ions are neutralized before they impact with the surface) the calculated vibrational period and collision time become equal at a kinetic energy of 224 eV, which shows reasonable agreement with the experimentally observed value of ~300 eV. The relationship between these parameters explains why the target ''senses'' the molecular nature of the diatomic low-energy projectile, whereas at high energy the surface interacts with the molecule as if it consisted of two separate atoms. The importance of the relationship between the collision time and the vibrational period was noted earlier by Landau and Teller to explain the mechanisms of energy transfer by molecular collisions in the gas phase,[3] although this relationship is used here for the first time to explain low-energy molecular scattering processes.

CARRIER DYNAMICS AT AN ULTRATHIN OXIDE LAYER–SILICON HETEROINTERFACE

The physical characteristics of the Si/SiO_2 heterointerface have played a vital role in establishing the dominance of silicon in semiconductor technology. As the dimensions of MOS devices continue to decrease, it becomes increasingly important to understand the influence of the thin oxide layer on the carrier dynamics at a buried heterointerface, which may, for instance, lead to breakdown and injection of hot carriers. A recently developed noncontact method of studying dynamic processes at an Si/SiO_2 heterointerface is based on monitoring the second-harmonic generation (SHG) signal. This method can ''sense'' the electric field (or charge distribution) at a buried heterointerface. All the experiments were carried out using a Ti-sapphire laser with a 150 fs pulse duration; the wavelength was varied between 7100 and 9100 Å, the average power was of the order of 300 W, and the repetition frequency 80 MHz. The beam was focused into a spot of around 10 μm diameter on the sample and the reflected SHG signal was measured using a photomultiplier and a photon counter with a time resolution of 0.5 s.

For all the thermal oxide samples we observed a sharp increase in the SHG signal over the first few hundred seconds of irradiation, followed by gradual saturation. This agrees with the results of observations made by other groups,[4–6] although our results (Fig. 2) also demonstrate a new effect: an appreciable increase in the SHG signal after blocking the exciting beam ($\lambda = 7900$ Å, maximum intensity 10 GW/cm^2) for several seconds and then unblocking it. After unblocking the laser beam, the SHG signal dropped rapidly to its saturation level. This characteristic was observed for all oxide layers more than 40 Å thick and at photon energies exceeding the 1.57 eV threshold value.[7] The inset to Fig. 2 shows the SHG signal response for a thin oxide layer

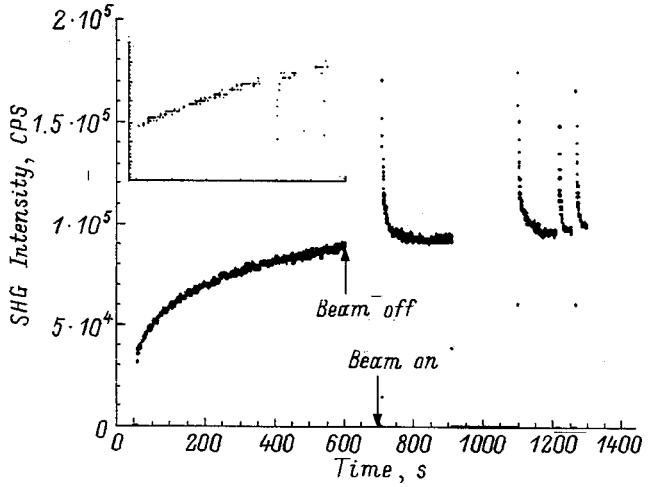

FIG. 2. Time dependence of the SHG signal from a thermal oxide layer 40 Å thick grown on a *p*-type Si(001) substrate. The inset shows the SHG signal after etching the oxide layer to 10 Å.

(around 10 Å), which shows no increase in field as a result of blocking the beam. In this case the field at the heterointerface decreased appreciably.

This effect cannot be attributed merely to multiphoton excitation of electrons in the conduction band of SiO_2 (as has been suggested[4,5] for the initial increase in the SHG signal). We postulate that the holes also play an important role in the charge separation dynamics.[7] In accordance with numerous studies made by other groups the band offset for oxide is between 4.35 and 4.9 eV (Refs. 8 and 9). Bearing in mind that the silicon band gap is 1.1 eV and the SiO_2 band gap is approximately 9 eV at room temperature, and also that the excited beam should have a photon energy higher than 1.52 eV, we conclude that four photons (6.08 eV) are required to create holes in the oxide valence band, whereas three photons of energy exceeding 1.36 eV (4.1 eV) are required to excite electrons from the silicon valence band to the oxide conduction band. Since the transition probability is considerably higher for the three-photon process compared with the four-photon, photoinjection of hot electrons causes an initial increase in the SHG signal although holes excited in the oxide layer also exist. Since the electrons possess a normal mobility of around 0.002 m^2/(V·s) (Ref. 10), they migrate very rapidly (within a few picoseconds) from the oxide layer, either into the silicon or to the outer surface of the oxide. In contrast the holes possess very low mobility, of the order of 10^{-10} m^2/(V·s) (Ref. 10). Thus, the holes passing through the heterointerface into the oxide will remain in the vicinity. For samples with an oxide layer more than 30 Å thick, holes near or at the heterointerface can easily migrate back into the silicon when the beam is switched off, whereas most of the electrons trapped at the surface remain there. The different behavior of the holes and electrons leads to increased charge separation, which is responsible for the increased dark field in the SHG signal. When the beam is then switched on, the hot holes will migrate across the heterointerface into the oxide, whereas the injection of electrons is severely impeded by the strong space charge field created by the electrons trapped at the SiO_2 outer surface. Conse-

quently, predominantly holes will migrate into the oxide, reducing the SHG signal. For thinner layers of oxide, electrons (like the holes) can tunnel back into the silicon. As a result, no increase in field is observed as a result of switching off the beam. This new effect of increasing field when the beam is switched off, which depends strongly on the oxide thickness and the photon energy, provides a deeper understanding of the role of trapping, generation from traps, and relaxation of holes at the Si/SiO_2 heterointerface, which is extremely important for developing a new generation of microelectronics devices.

WAVELENGTH-SELECTIVE MODIFICATIONS OF AN *a*:Si–H–P COMPLEX

The W. M. Keck free-electron laser (FEL) at Vanderbilt University is a powerful and unique instrument for studying semiconductors. In particular, it can be used to test the possibility of wavelength-selective modifications of materials. Here we demonstrate its scope for depassivating dopants in hydrogenated semiconductors. The free-electron laser delivers macropulses of 4 μs rated duration and energies between 20 and 60 J at a repetition frequency up to 30 Hz; each macropulse consists of approximately 10^4 pulses of 1 ps duration at 350 ps intervals. The FEL wavelength can be tuned continuously in the range 2.0–9.5 μm. This range, which is usually increased by frequency multiplication or by using higher harmonics, covers the band-gap energies of most technologically important semiconductors. The high average power and complex time structure of the FEL pulses are used for various spectroscopic studies of semiconductors and for modifying materials. Preliminary work with FELs has already demonstrated that structure-selective modifications can be made to materials.[11–16] The high FEL pulse energy means that energy can be absorbed at specific bonds, leading to localized modifications of materials. These processes depend strongly on the excited-state lifetimes. The intrinsic lifetime, stimulated lifetime, and atomic configuration determine the absorption energy and the suitability of high-intensity infrared radiation for nonthermal wavelength-selective modification of a material.

Hydrogen plays a key role in amorphous silicon, passivating dangling bonds to eliminate defects. However, it also changes the doping properties of *a*-Si–H, binding the dopants and lowering the conductivity. The dopant is passivated by forming a complex with a silicon atom and a hydrogen atom, as shown in Fig. 3. This complex can be seen at $1555\,cm^{-1}$ in the vibrational absorption spectra for phosphorus impurities and at $1875\,cm-1$ for boron impurities.[17]

We investigated the possibility of using high-intensity FEL laser radiation for selective activation of dopants in *a*-Si–H. By resonantly exciting the hydrogen doping complex, it is possible to remove the hydrogen, leaving the hydrogen at the sites of dangling bonds. Figure 3 shows vibrational absorption spectra for a phosphorus-doped *a*-Si–H film before and after FEL irradiation at $1555\,cm^{-1}$, which corresponds to the vibrational mode of the Si–H–P complex. It can be clearly seen that after FEL irradiation, the shoulder on the low-energy side of the strong absorption line at

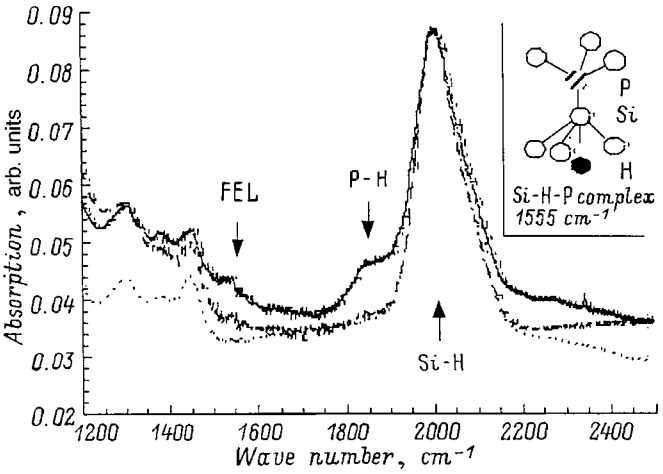

FIG. 3. Comparison between the vibrational absorption spectra before (solid curve) and after FEL irradiation (dotted and dashed curves) at $1555\,cm^{-1}$. The inset shows a schematic of the passivation of donors and acceptors in *n*-type *a*-Si–H.

$2000\,cm^{-1}$ is reduced appreciably. On comparing the vibrational absorption spectra for *a*-Si–H samples with various dopant concentrations, we find that the intensity of this shoulder correlates with the dopant concentration and we ascribe this resonance to the P–H vibrational mode. Preliminary electrical measurements showed an increase of approximately two orders of magnitude in the surface conductivity of phosphorus-doped *a*-Si–H films, which provides further evidence of the coupling between this mode and the H-passivated donors. It should also be noted that the FEL radiation does not influence the intensity of the Si–H peak at $2000\,cm^{-1}$. These results confirm that dopants could be selectively activated without removing hydrogen from dangling Si–H bonds.

CONCLUSIONS

We have described three new, fundamental, energy-selective nonlinear effects, relating to the ion–surface interaction, carrier dynamics at an Si/SiO_2 heterointerface, and structure-selective modification of a material. We have observed that energy transfer is considerably greater for molecular ions compared with atomic ions. These results can be used in growth processes, such as molecular beam epitaxy. Second-harmonic generation, applied to study Si/SiO_2 heterointerfaces, has provided a deeper understanding of the nanoscale dynamics of bound electrons and holes, including trapping and generation processes in an ultrathin oxide film. Finally, we have achieved selective activation of dopants which improved the conductivity by two orders of magnitude without removing hydrogen from dangling bonds. This work offers a fundamentally new energy- and site-selective approach which is likely to produce a paradigm shift in the technology of nanodimensional modifications to materials.

This work was supported by the Office of Naval Research under Grants Nos. 00014-94-11023 and 00014-94-10995.

The authors thank colleagues at the FEL center for their tireless efforts and support. G. Lüpke also acknowledges financial support from the Alexander von Humboldt foundation.

[1] Y. Yao, Z. Hargitai, M. Albert *et al.*, Phys. Rev. Lett. **81**, 550 (1998).
[2] Y. Yao, PhD dissertation, Vanderbilt University (1997).
[3] L. D. Landau and E. Teller, Phys. Z. Sowjetunion **10**, 34 (1936).
[4] J. G. Mihaychuk, J. Bloch, and H. M. van Driel, Opt. Lett. **20**, 2063 (1995).
[5] J. Bloch, J. G. Mihaychuk, and H. M. van Driel, Phys. Rev. Lett. **77**, 920 (1996).
[6] N. Shamir, J. G. Mihaychuk, and H. M. van Driel, J. Vac. Sci. Technol. A **15**, 2081 (1997).
[7] W. Wang, G. Lüpke, M. Di Ventra *et al.*, Phys. Rev. Lett. **81**, 4224 (1998).
[8] J. L. Alay and M. Hirose, Appl. Phys. Lett. **81**, 1606 (1997).
[9] P. Perfetti, C. Quaresima, C. Coluzza *et al.*, Phys. Rev. Lett. **57**, 2065 (1986).
[10] *The Si–SiO_2 System*, 1st ed., edited by P. Balk (Elsevier, Amsterdam, 1988).
[11] A. Ueda, J. T. McKinley, R. G. Albridge *et al.*, Mater. Res. Soc. Symp. Proc. **285**, 215 (1993).
[12] J. T. McKinley, R. G. Albridge, A. V. Barnes *et al.*, Nucl. Instrum. Methods Phys. Res. A **341**, 156 (1994).
[13] J. T. McKinley, R. G. Albridge, A. U. Barnes *et al.*, J. Vac. Sci. Technol. A **12**, 2323 (1994).
[14] G. Margaritondo, C. Coluzza, J. L. Staehli *et al.*, J. Phys. (Paris) III **4**, Colloque C9, Suppl. (1994).
[15] A. Ueda, R. G. Albridge, A. V. Barnes *et al.*, Nucl. Instrum. Methods Phys. Res. B **100**, 427 (1995).
[16] N. H. Talk, R. G. Albridge, A. V. Barnes *et al.*, Appl. Surf. Sci. **106**, 205 (1996).
[17] K. Bergman, M. Stavola, S. J. Pearton, and J. Lopata, Phys. Rev. B **37**, 2770 (1988).

Translated by R. M. Durham

Mass spectrometric analysis of clusters formed in laser ablation of a sample

B. N. Kozlov and B. A. Mamyrin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 81–84 (September 1999)

Experimental data are used as a basis for discussion of the principal methods of cluster formation in the laser ablation of targets: condensation during expansion of the cloud of evaporated material, clustering at the surface accompanying redeposition of material back on the target, and emission of entire nanoblocks from the target. Methods of distinguishing between these processes are discussed. [S1063-7842(99)01809-7]

INTRODUCTION

The development of laser ablation mass spectrometry[1,2] has been stimulated first, by the requirements for optimizing processes in laser technologies such as laser deposition of thin films and second, by the possibility of identifying large nonvolatile formations such as clusters, polymers, and biomolecules. The composition of the particles emitted by laser ablation differs radically from the products of thermal emission or secondary ion emission. In thermal emission, large structural units of the target usually break down and atoms, simple molecules, small clusters, and, very rarely, ions are observed in the gas phase. When a target is bombarded with ions or fast atoms, predominantly atoms and extremely simple compounds, albeit ionized to an appreciable extent, are emitted from the target surface as a result of a cascade of collisions; almost no large particles are observed. At this point, it should be acknowledged that an exception to this is the interaction of ultrahigh-energy (megaelectron-volt) ions with a target; these interact with the target electrons rather than the atoms.[3] Sputtering of a target by these ions has many features in common with laser ablation.[4,5] The processes leading to the appearance of large gas-phase clusters and macromolecules as a result of laser ablation are not yet fully understood. A special method of dissolution in a highly volatile matrix, known as matrix-assisted laser desorption ionization (MALDI)[6] is used to reliably identify these objects. In this case, the laser energy is absorbed by the matrix and the complex component simply does not have time to break down. It is entrained into the gas phase by the evaporating matrix and is rapidly cooled by adiabatic expansion of the cloud of matrix molecules. However, we know that without preliminary dissolution in a highly volatile matrix, polyatomic cluster ions are found as a result of the laser ablation of samples. In the present paper we consider typical examples of their formation, their properties, and methods of identifying their origin. The experimental results taken as examples were obtained by analyzing the composition and particle velocities in a laser jet using a time-of-flight mass spectrometer. The experimental apparatus (plasma expansion from a target in a field-free space and its analysis at a distance of 40–120 mm) is similar to that described earlier.[7]

CONDENSATION IN AN EXPANDING LASER JET

The best-understood process is the formation of gas-phase clusters as a cloud of evaporated materials cools down. For instance, the formation of fullerenes was observed for the first time[8] as a result of the laser evaporation of carbon in a chamber filled with inert gases. However, condensation also occurs during evaporation in vacuum conditions. Cooling takes place as a result of the isentropic expansion of the cloud of evaporated material. A fairly high intermolecular collision efficiency is required for condensation in vacuum. In most cases, the particle concentration above the surface under pulsed laser evaporation is high enough: 10^{17}–10^{20} $cm^{-3}\,s$ (evaporation of 1–1000 atomic layers within approximately $10^{-8}\,s$ at a cloud expansion velocity of 10^5–10^6 cm/s). Characteristic features of these fluxes and the condensation taking place in them have been quite well studied experimentally for gases expanding into empty space.[9] Model representations developed to calculate these fluxes can also be applied to describe condensation in a laser jet. For instance, it is possible to estimate the temperatures to which the cloud would be cooled in the absence of condensation, the average number of collisions for each particle during expansion, and the cloud volume at which the interparticle collisions terminate. For example, in the evaporation of 1000 atomic layers having an area of $1\,mm^2$ the average number of collisions may be of the order of 10^5. Note that the observed micron-size particles (of the order of 10^{10} atoms) cannot have a condensational origin. Even if there is a 100% probability of the particles combining in each collision, the center of the resulting particle size distribution should not differ substantially from the calculated number of collisions in the approximation of independently expanding particles (10^5 atoms form a droplet of around 10 nm). However, the reliability of these calculations is very poor because of the indeterminacy of the initial expansion conditions of the cloud, its composition, particle energies, and interaction cross sections. Thermal evaporation is usually assumed when estimating the initial parameters, approximating the equation for the particle thermal flux (generally atoms) from the surface at high temperatures.

We shall now consider the experimentally most studied case of laser evaporation, viz., the ablation of YBaCuO ce-

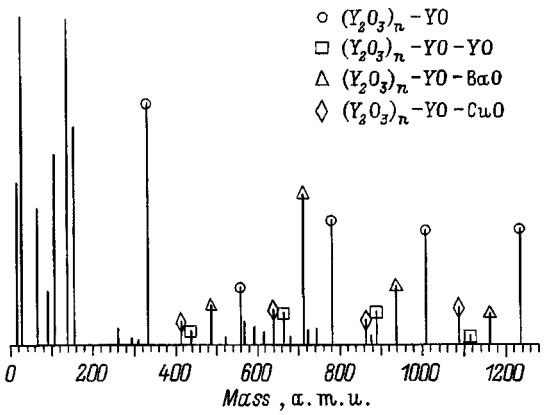

FIG. 1. Yttrium oxide clusters formed by condensation during expansion of laser-ablated target material in vacuum.[7] Target — YBaCuO ceramic.

ramic. Clusters consisting predominantly of yttrium oxide were observed in the jet of evaporated material even under vacuum conditions.[7,10] Figure 1 shows the mass spectrum of these clusters. Since the yttrium atoms in the sample are separated by more numerous atoms of other elements, the origin of these clusters can be attributed to condensation, predominantly of yttrium oxide as the least volatile component in the gas phase. It should be noted that the most favorable conditions for condensation in the case of laser ablation in vacuum exist when the material to be formed has a higher binding energy than the material of the initial sample and whose particles have internal degrees of freedom (vibrational and rotational) for damping of the collision energy, and when a large sample volume is evaporated in each pulse. Important experimental evidence for the condensational origin of the clusters during the ablation of YBaCuO ceramic is provided by the disproportionately rapid decay and complete disappearance of the cluster peaks as the area of the irradiated spot on the surface decreases.[7]

Note that the quantity of evaporable material generally increases substantially for targets with smaller ordered structures (microcrystals), such as sintered powders, soot, and so on. This may be attributed to the reduced thermal conductivity of the target and to the detachment of entire blocks and their evaporation during the emission process.

FORMATION OF CLUSTERS BY REDEPOSITION OF MATERIAL FROM THE LASER JET BACK ONTO THE TARGET

Deposition of material back onto the target and the formation of clusters as a result of particles migrating over the surface and coalescing is very rarely taken into account when identifying clusters during the laser ablation of targets. It has been noted that as they expand, the emitted particles undergo numerous collisions and that even under vacuum conditions a considerable proportion of the ablated material is deposited back onto the target. Usually the entire target surface except for the irradiated part is passivated or often simply coated with a layer of molecules or radicals (water or vacuum-pump oil) that are not taken into account. The redeposited particles are not usually incorporated back into the target structure (their energy is low and the target is only heated locally), and

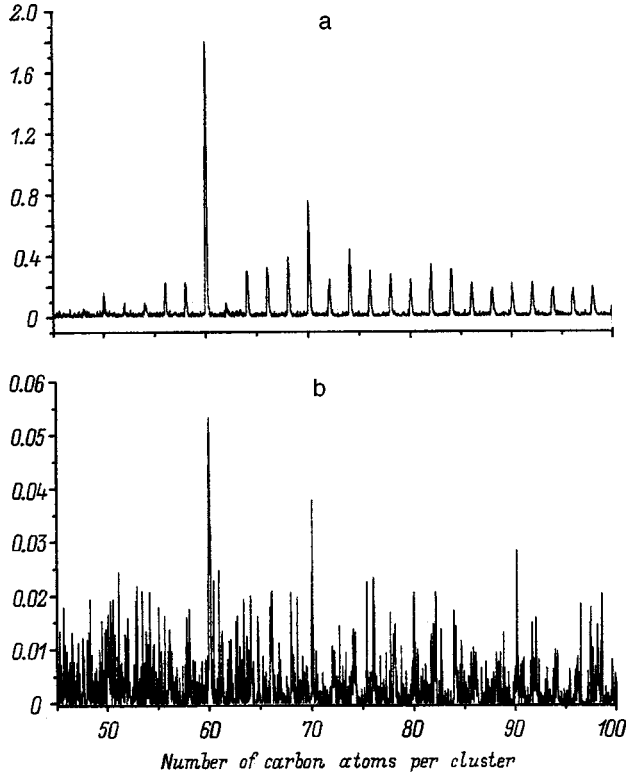

FIG. 2. Formation of fullerenes by redeposition of evaporated carbon back onto the target. Target — graphite, initially containing no fullerenes: a — signal obtained for constant displacement of the laser beam over the target, b — under the same conditions when the same point on the target is irradiated. The signal is thirty times smaller.

they form clusters weakly bound to the target. If upon reaching a certain size these clusters become quite inert or if their surface is covered with radicals of background substances between laser shots, it will be the number rather than the size of the clusters that increases from shot to shot. When such a region of redeposition is exposed to low-power radiation (near the evaporation threshold of the target), those clusters not bound to the target are readily emitted, with almost no breakup. An example of this process is the formation of fullerenes as a result of the laser evaporation of various carbonaceous samples in vacuum, especially pyrolytic graphite. Even the first reports of the formation of fullerenes under vacuum conditions mentioned that they appeared only after a sufficient preliminary number of laser pulses.[11] Granted, this has been ascribed from the start to the need to form a crater to provide a denser flux of evaporated material, but later it was shown that the redeposited phase plays a decisive role.[12] Figure 2 shows fragments of the mass spectra of carbon clusters formed as a result of the laser evaporation of graphite. The signal shown in the upper part is accumulated over 3000 pulses when a laser spot of area $0.1\,mm^2$ is continually shifted over a $20\,mm^2$ region of the target. The signal shown in the lower part was obtained for 3000 laser shots at a single point. In the second case the signal is thirty times weaker! In the first shots using a fresh sample, no fullerenes are observed, i.e., the criteria for the identification of cluster formation by redeposition of sputtered material is the appearance and growth in the number of clusters as the sample

evaporates and, most importantly, a substantial increase in the number of these clusters as the laser beam moves to a region adjacent to that which was irradiated previously.

EMISSION OF FRAGMENTS AND ENTIRE BLOCKS FROM THE TARGET

There are many hypotheses on the mechanisms for the removal of large clusters from a target under the action of a laser pulse. Hydrodynamic mechanisms for the formation of droplets and cleaving of blocks as a result of steep thermal gradients have been discussed.[13,14] Overheating of the subsurface layers of the target and release of a gas phase in the bulk may also play a significant role.[15] Quite clearly, defects and boundaries between microcrystals in the target play a major role in the breakup of the target. A simplified case of the detachment of blocks from a polycrystal is the emission of clusters weakly bound to the surface, which was considered in the previous section. The kinetic energy of these clusters is orders of magnitude higher than the thermal energy. For instance, the energy of desorbed fullerenes is up to 50 eV (Ref. 12). It is true that their velocity is approximately the same as that of the remaining particles in the jet and can be attributed to acceleration in the entire flux. However, even when isolated clusters are emitted, their energy is also high. Gold clusters, desorbed after strong excitation of their electron system by the propagation of a high-energy ion, leave the substrate with an energy of around 100 eV (Ref. 5). In order to explain these high energies, the possibility is considered that the electron excitation at the cluster boundaries is converted directly into kinetic energy of the entire particle, i.e., the electron excitation does not have time to be transferred to the atomic core, and the collective excitation of the electrons at the boundary does not simply destroy the bond with the surface, but also leads to repulsion of the entire particle. It is postulated that in this case, the kinetic energy approximately corresponds to the binding energy of the particle with the surface. Similar models were considered earlier for the emission of atoms and simple molecules having anomalously high kinetic energy from a surface.[4,13,16] In a solid target, release of electron energy at grain boundaries and defects accumulated under intensive irradiation must also be considered. Thus, entire fragments may become detached from the target if it does not melt. At near-threshold radiation energies, graphite is ablated to a considerable extent, giving clusters of 7–27 atoms comprising fragments of molecular graphite layers.[12] We attribute the very substantial variation of the amplitude and peak ratio in the mass spectrum of the graphite clusters from one pulse to another to the sparse information on the evaporation of microcrystals.[17] Graphite microcrystals have not been detected by mass spectrometry because of their very large mass, but we have observed fragments of porous graphite. The initial sample was obtained by etching silicon from SiC. The mass spectrum near clusters of 7–27 atoms is very stable for these samples. For clusters of 300–1300 atoms, the signal could not be resolved but is reliably observed. As the cluster mass increases, their velocity decreases. Figure 3 shows sections of mass spectra obtained at a distance of 45 mm from the target

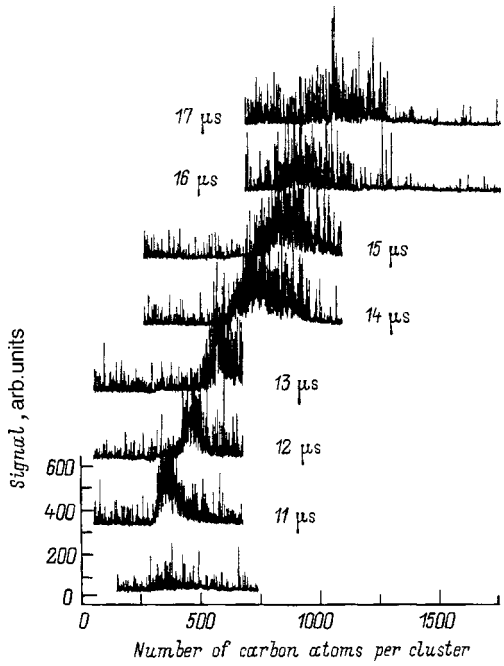

FIG. 3. Emission of nanoclusters from porous graphite. The point of analysis is 45 mm from the target. With increasing delay after the laser pulse the masses of the detected clusters increase.

for various delays after the laser pulse. The masses of these clusters correspond approximately to the postulated sizes of the interpore barriers. The kinetic energy and therefore the spread of the initial kinetic energies of these particles is higher than 400 eV, which severely impedes mass spectrometric analysis. The criterion for cluster detachment from the target rather cluster formation during expansion as a result of the condensation of small fragments is that the number of clusters in the jets produced in different pulses can vary considerably but the mass distribution of the particles is conserved.

It can be seen from these examples that when laser mass spectrometry is used to analyze structural units of complex targets, the results must be evaluated with extreme caution. The simple reasoning described above can give preference to a particular hypothesis on the mechanism of cluster formation using a laser beam. At the same time, this reasoning is important for the directed development of cluster fluxes in vacuum.

The authors thank A. M. Danishevskiĭ for supplying the porous graphite samples and discussing their properties.

The authors are grateful to The Foundation for Intellectual Collaboration and the Program ''Fullerenes and Atomic Clusters'' for supporting this work.

[1] *Laser Ionization Mass Analysis*, edited by A. K. Vertes, R. Gijbels, and F. Adams (Wiley, New York, 1994).
[2] Yu. A. Bykovskiĭ and V. N. Nevolin, *Laser Mass Spectrometry* [in Russian], Énergoatomizdat, Moscow (1985), 129 pp.
[3] R. D. MacFarlane and D. F. Torgerson, Int. J. Mass Spectrom. Ion Phys. **21**, 81 (1976).
[4] F. R. Krueger, Z. Naturforsch. **38**, 385 (1983).
[5] I. Baranov, S. Bogdanov, A. Novikov *et al.*, Nucl. Instrum. Methods Phys. Res. B **122**, 329 (1997).

[6] M. Karas, D. Bachmann, U. Bahr, and F. Hillenkamp, Int. J. Mass Spectrom. Ion Processes **78**, 53 (1987).
[7] B. N. Kozlov, I. I. Pilyugin, V. G. Shchebelin *et al.* Zh. Tekh. Fiz. **64**(9), 155 (1994) [Tech. Phys. **39**, 944 (1994)]; B. N. Kozlov, I. I. Pilyugin, V. G. Shchebelin *et al.*, Mikrochim. Acta **120**, 111 (1995).
[8] H. W. Kroto, J. R. Heath, S. C. O'Brien *et al.*, Nature (London) **318**, 152 (1985).
[9] O. F. Hagena, Surf. Sci. **106**(1), 101 (1981).
[10] C. H. Becker and J. B. Pallix, J. Appl. Phys. **64**, 5152 (1989).
[11] A. O'Keefe, M. M. Rass, and A. P. Baranovski, Chem. Phys. Lett. **130**, 17 (1986).
[12] B. N. Kozlov, S. N. Kirillov, and B. A. Mamyrin, Proc. SPIE **3093**, 233 (1997).
[13] R. Kelly and A. Miotello, *Pulsed Laser Deposition of Thin Films*, edited by D. B. Chrisey and G. K. Huber (Wiley, New York, 1994), pp. 55–87.
[14] R. Kelley, J. J. Cuomo, P. A. Learly *et al.*, Nucl. Instrum. Methods Phys. Res. B **9**, 329 (1985); R. Kelly and J. E. Rothenberg, Nucl. Instrum. Methods **7/8**, 755 (1985).
[15] E. N. Sobol and N. G. Petrovskaya, Supercond. Sci. Technol. **6**, 67 (1993); V. N. Bagratashvili, E. N. Antonov, E. N. Sobol *et al.*, Appl. Phys. Lett. **66**, 2451 (1995).
[16] R. W. Dreifus, R. Kelly, and R. E. Walkup, Appl. Phys. Lett. **49**, 1478 (1986).
[17] B. N. Kozlov, V. P. Bul'chenko, and B. A. Mamyrin, Mol. Mater. **7**, 99 (1996).

Translated by R. M. Durham

Electron spectroscopy of objects for nanoelectronics

V. M. Mikushkin

A. F. Ioffe Physicotechnical Institute, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 85–89 (September 1999)

An electron-spectroscopic analysis is made of layered nanostructures and clusters at the surface and in the bulk of a solid. A new method of forming metal/insulator/semiconductor (superconductor) nanostructures is proposed based on ion-stimulated metal segregation effects at the surface of low-temperature gallium arsenide and a 123 high-temperature superconductor. The geometric parameters and electronic structure of these nano-objects are studied. It is shown that their electronic properties can be controllably varied *in situ* by acting on the surface. The dimensional transformation of the electronic properties of metal clusters is studied for clusters in the insulator SiO_2, in the superconductor LTMBE-GaAs, and on silicon and graphite surfaces. The nature of this transformation is clarified. A diagnostics for cluster ensembles is developed by which one can determine the parameters needed to describe single-electron transport: the average number of atoms per cluster, the average distance between clusters and isolated atoms, and the chemical state of the atoms. Ensembles of silver clusters with specified parameters are obtained on a silicon surface. It is shown that these ensembles are potentially useful for developing single-electron devices.
[S1063-7842(99)01909-1]

Films and layered structures of nanometer thickness, quantum wires and clusters, are some of the ''hottest'' objects in modern physics. These objects can be used to develop fundamental ideas on an intermediate state of matter situated between the atomic and the solid state, and offer fundamentally new possibilities in electronics. Thus modern physics and nanoelectronics require the development of diagnostic methods to determine the electronic structure and the elemental and chemical composition of nano-objects. A promising basis for these diagnostic techniques are various modifications of electron spectroscopy, characterized by nanometer resolution over the depth of a solid-state object. In the present paper we give examples of the controlled development and investigation of unique layered nanostructures and cluster ensembles, which demonstrate the capabilities of electron-spectroscopic methods of studying these objects.

Some of the most promising materials in electronics are III–V semiconductors such as GaAs and InP and 123 high-temperature superconductors (HTSCs) such as $Y(Dy)Ba_2Cu_3O_{7-\delta}$. The potential usefulness of these materials derives from their inherent high carrier mobility which, in principle, can increase the speed of devices by one or two orders of magnitude compared with silicon devices. However, the fabrication of device nanostructures from metals, insulators, semiconductors, and superconductors (MIS structures) is impeded by the effective interdiffusion of materials in the conventional processes used for the layer-by-layer growth of structures, such as molecular beam epitaxy. For instance, the typical broadening of HTSC/insulator interfaces is some tens of nanometers. It is therefore important to observe and study physical processes which could suggest fundamentally new methods of fabricating thermodynamically-stable MIS structures. One such process may be ion-beam-stimulated segregation of metals, which we have observed at the surface of new materials: gallium arsenide grown by molecular beam epitaxy at low temperature (LTMBE-GaAs) and the HTSC $DyBa_2Cu_3O_{7-\delta}$ (Refs. 1–3). The MIS nanostructures are formed during ion bombardment of the surface of these materials (Fig. 1).

Segregation of arsenic and dysprosium at the surface of LTMBE-GaAs and $DyBa_2Cu_3O_{7-\delta}$ was observed as a result of analyzing the photoelectrons emitted by deep and shallow core levels. These photoelectrons have different mean free paths λ and carry information on the elemental composition of a thin (5–7 Å) surface layer and a considerably thicker (22–27 Å) ''bulk'' layer. The intensity ratio of the surface photoelectron line to the bulk line in samples exposed to ion bombardment is considerably higher than the same ratio in reference materials, normal gallium arsenide and metallic dysprosium. This at least indicates that the surface is enriched in these elements. A detailed quantitative analysis of the relative intensities revealed that a layer of pure arsenic of thickness $l=8$ Å forms on the LTMBE-GaAs surface, whereas a monatomic or diatomic layer of dysprosium or an ensemble of dysprosium clusters forms on the $DyBa_2Cu_3O_{7-\delta}$ surface, depending on the ion energy.

The reasons for the segregation of arsenic and dysprosium differ. The samples inserted in the spectrometer are coated with a layer of native oxides. After this layer has been removed by ions, the excess arsenic contained in the LTMBE-GaAs simply ''pours'' onto the freshly prepared surface. The interaction between the ions and HTSC $DyBa_2Cu_3O_{7-\delta}$ is rather more complex: initially chemical conversions take place, resulting in the formation of an insulator layer containing elemental dysprosium, and then the dysprosium diffuses toward the surface. Studies of the spec-

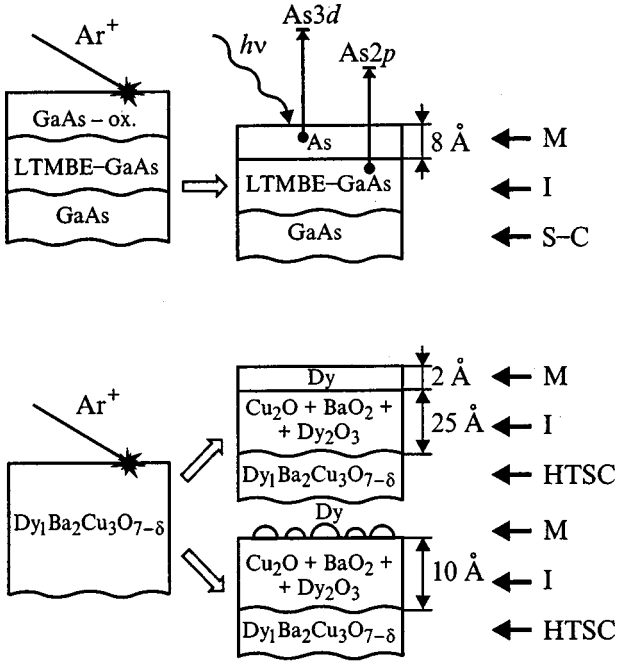

FIG. 1. Schematic of ion-stimulated formation of MIS nanostructures.

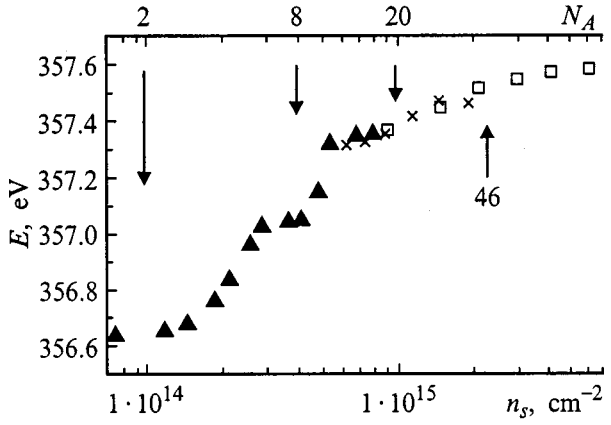

FIG. 2. Dependence of the Auger electron energy of $AgM_{4,5}$ *VV* on the dose n_s of deposited silver and the average number of atoms N_{at} in silver clusters on a graphite surface.

tra of the core and valence photoelectrons have established that the segregate layers are of a metallic nature and have determined the band offsets at the interfaces, the band bending, and the density of states, Ultimately the electronic structure was reconstructed in its entirety.[1,3] It was also observed that external influences on the segregate such as dosed oxidation can change the parameters of the electronic structure, i.e., the electrophysical properties of the structures. Thus, MIS structures with nano- and subnanometer layers were obtained by a fundamentally new method under conditions of *in situ* monitoring of their electronic properties.

As we know, the properties of clusters depend on their sizes. Electron spectroscopy provides information averaged over an ensemble of different-sized clusters at the surface and in the bulk of a solid. Thus in order to improve the accuracy and reliability of the information on clusters, we need to produce ensembles with a small variance in cluster size. We solved this problem for ensembles of metal clusters at the surface of silicon and graphite.[4–6] It was shown that these clusters are planar. The nature of the so-called size shifts of the Auger and photoelectron lines was identified, i.e., changes in the energies of the electrons emitted by the clusters as the cluster sizes vary. For clusters consisting of several atoms these shifts are caused by a change in the binding energy as a result of a change in the internuclear distance and the increasing importance of quasimolecular orbitals. For larger clusters these size shifts are mainly caused by a change in the relaxation energy of the electron subsystem as a result of electron collectivization and an increase in the density of states near the Fermi level. The established dependences can be used to link the dose scale of the deposited metal (n_s) with the average number of atoms in the clusters (N_{at}) and to propose a method of determining this number.[7] In this way we solved the problem of the controlled formation of a cluster ensemble with given parameters and known properties.

The proposed method of determining the number N_{at} revealed new dependences in the formation of the electronic structure of clusters at a surface.[6,8] For example, Fig. 2 gives the size dependence of the Auger electron energy over a wide range of N_{at}; as we have said, this dependence is almost completely described by the size dependence of the relaxation energy *R*. The position of the observed step characteristics in $R(N_{at})$ correlates with the known ''magic'' numbers for free spherical clusters: 2, 8, 20, and 46, which regulate the formation of the electron shells of the cluster. It was thus concluded that the electronic structure of planar clusters on a surface is formed in accordance with the same relationships as for free spherical clusters. This conclusion is quite unexpected, since the symmetry of the clusters under comparison is altogether different. This result also confirms the reliability of the method of determining N_{at}. The abruptness of the steps indicates the small spread of the cluster sizes in our experiments.

Another interesting result obtained on the basis of an accurate determination of cluster size is that the spin–orbit splitting of the level (Δ_{sl}) varies nonmonotonically during growth of a cluster on a surface.[9] The splittings of the $3d$ and $4d$ levels of silver shown in Fig. 3 were obtained from the Auger and photoelectron spectra. It can be seen that Δ_{sl} increases abruptly, then decreases, and oscillates with decaying amplitude and frequency as the atoms merge to form diatomic, triatomic, and polyatomic clusters. This effect is evidently caused by distortion of the electron shells as a result of the polarization of the atoms by the surface, which may vary appreciably at the initial stage of cluster formation. Polarization of the shell increases the effective orbital angular momentum and thus the spin–orbit interaction ($\langle sl \rangle$). The outer $4d$ shell is considerably more strongly polarized that the inner $3d$ shell. Thus, the observed effect is considerably stronger for the $4d$ shell, although the absolute value of the splitting is an order of magnitude lower.

Polarization of the clusters by the surface is clearly responsible for the anomalous behavior of the probability of

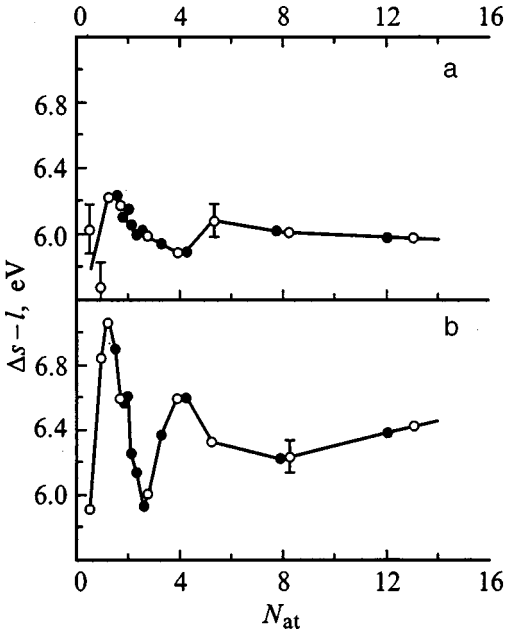

FIG. 3. Spin–orbit splitting of Ag$3d_{3/2,5/2}$ (a) and Ag$4d_{3/2,5/2}$ (b) levels as a function of the average number of atoms N_{at} in silver clusters on a graphite surface. Filled and unfilled circles — results of two experiments.

excitation of collective oscillations of the electron subsystem of an isolated cluster (local plasmon) by a photoelectron emitted in that cluster.[7] The size dependence of this probability is shown in Fig. 4b, and the analogous dependence of the probability of excitation of a local plasmon in a detected event in which an external electron is scattered by the cluster (a). It can be seen that these probabilities do not depend on the cluster size for particles containing more than 12 atoms. For smaller clusters the dependences are directly the opposite: the probability of excitation of a local plasmon by an external electron decreases while that for a photoelectron increases drastically. The first dependence is expected and correlates well with the size dependence of the relaxation energy: the lower relaxation energy is caused by a lower degree of electron collectivization in the cluster, which is also responsible for the lower probability of excitation of collective oscillations. The anomalous increase in the probability of excitation of a plasmon by a photoelectron can be explained qualitatively by polarization of the outer $5s$ shell of the silver atom. Since plasmons in silver exhibit hybrid oscillations of the $5s$ free-electron gas with correlated excitation of $4d$ electrons, the plasmons can be excited via the excitation of $4d$ electrons. The probability of this process depends on the overlap of the wave functions of the $4d$, $5s$, and photoionizing $3d$ electrons. For small clusters the overlap should be greater because of the stronger interaction of the atoms with the surface and thus the stronger polarization of the valence $5s$ shell. For larger clusters, the role of intracluster interaction increases and the polarization of the atoms decreases.

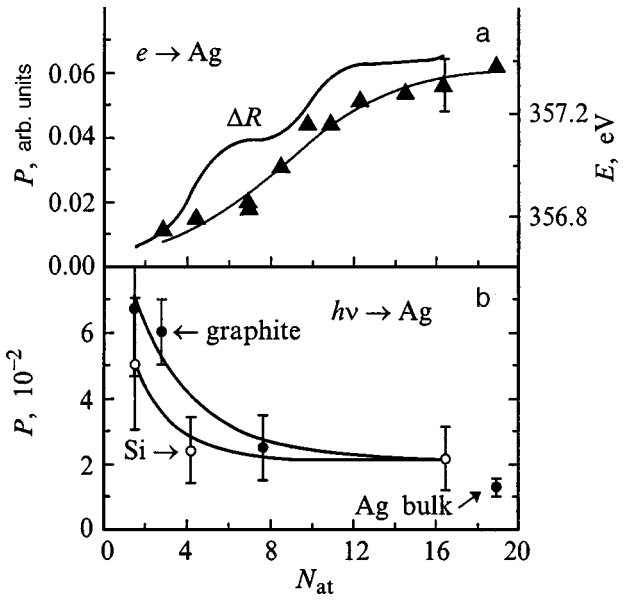

FIG. 4. Size dependences of the probability of excitation of a local silver-cluster plasmon in a detected event in which an external electron is scattered by the cluster (a) and in the emission of a $3d$ photoelectron by the cluster (b). Size dependence of the relaxation energy ΔR (a).

Figure 4a illustrates an important result: local collective oscillations are excited in clusters consisting of only a few atoms (between two and four).[10] This implies that plasma oscillations can be excited in systems consisting of only a few electrons. We observed low-frequency (''red'') local plasmon modes and established that they are excited considerably more effectively by photoelectrons than by external electrons.[7] It was shown that the size dependence of the local-plasmon energy red shift due to these red modes can also be used to determine the average particle size.

Electron-spectroscopic analyses have been used to develop the technology for the controlled production of ensembles of homogeneous subnanometer clusters on a surface, having a given average size and intercluster distance. Systems showing Coulomb blockade effects in the transport of individual electrons were obtained, and it was demonstrated that these ensembles are potentially useful for developing fundamentally new single-electron devices.

Changes in the electronic properties of copper in an SiO_2 matrix have been studied for the transition from an atom to a cluster and then to the solid metal.[9] Figure 5 shows size dependences of the electron binding energy in the atom $E_B(N_{at})$, the energy levels $\varepsilon(N_{at})$, the Auger electron energies $E_A(N_{at})$, and the relaxation $R(N_{at})$. It can be seen from Fig. 5 that the relaxation changes during cluster growth are responsible for the enormous shifts of the Auger lines, which reach 5 eV. It was noted that these large shifts can be used to determine cluster sizes of less than 10 Å in an insulator matrix with reasonable accuracy. It was established that for a transition from an atom to a cluster, the energy levels decrease rather than increase as they do for clusters at the surface, and that this behavior is caused by charge transfer from the matrix to the atom and not by any change in the quasimolecular orbitals.

These investigations were used as the basis to develop electron-spectroscopic diagnostics for copper and arsenic clusters in new composite materials Cu/SiO_2 and LTMBE-GaAs which are also extremely promising for fabricating single-electron devices.[6,11] This technique can provide the information needed to describe the transport of isolated elec-

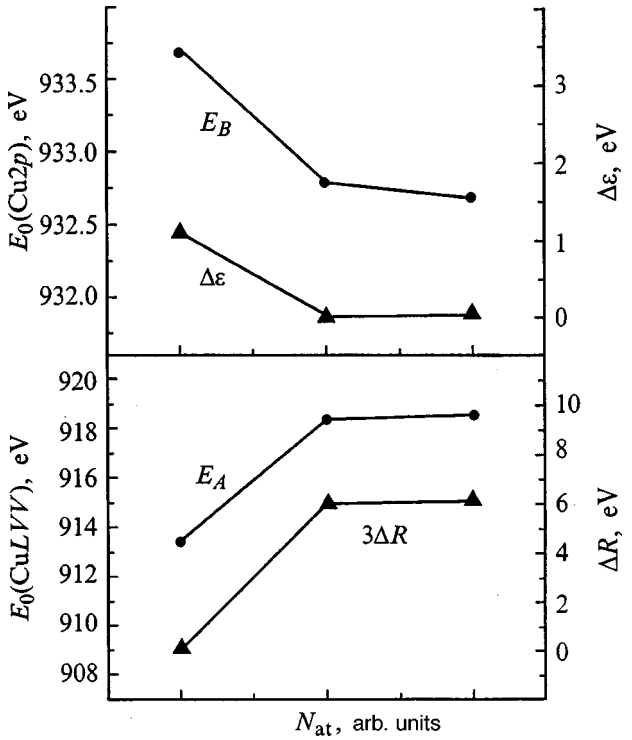

FIG. 5. Binding energy, Auger electron energy, level energy, and change in the relaxation energy ΔR in an atom–cluster–solid transition.

trons in a composite material, i.e., the average cluster diameter, the distances between clusters and atoms not combined to form clusters, their concentrations, and also the chemical state of the atoms.[12]

To sum up, we have made electron-spectroscopic analyses of nanometer and subnanometer layered structures and clusters at the surface and in the bulk of a solid. We observed ion-stimulated segregation of metallic subnanofilms and clusters at the surface of semiconductors and superconductors and ascertained the size dependence of the energies of the electrons emitted by the clusters. We have devised diagnostic techniques for determining the parameters of their electron structure and the geometric parameters of cluster ensembles and developed a technology for the controlled production of cluster ensembles having the properties required to fabricate single-electron devices.

This work was supported financially by the Russian Fund for Fundamental Research (Project No. 96-02-17966-a), the Russian MNTP (under the program ''Physics of Solid-State Nanostrutures'' Project No. 97-3-005) and also with the collaboration of the St. Petersburg Joint Center for Shared Utilization.

[1] Yu. S. Gordeev, V. M. Mikushkin *et al.*, Fiz. Tverd. Tela (St. Petersburg) **38**, 3299 (1996) [Phys. Solid State **38**, 1800 (1996)].
[2] Yu. S. Gordeev, V. V. Mamutin, V. M. Mikoushkin, and S. F. Sysoev, *Abstracts of MRS-Spring Symposium of Materials Research Society*, San Francisco, 1995, Vol. K5, p. 30.
[3] Yu. S. Gordeev, V. M. Mikushkin, S. Yu. Nikonov, and S. E. Sysoev, Poverkhnost' No. 8, 91 (1998).
[4] Yu. S. Gordeev, M. V. Gomoyunova, V. M. Mikushkin *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 1777 (1994) [Phys. Solid State **36**, 973 (1994)].
[5] M. V. Gomoyunova, Yu. S. Gordeev, V. M. Mikoushkin *et al.*, Phys. Low-Dim. Struct. **4**, 11 (1996).
[6] Yu. S. Gordeev, Poverkhnost', No. 8, 40 (1998).
[7] V. M. Mikushkin and S. E. Sysoev, Fiz. Tverd. Tela (St. Petersburg) **38**, 558 (1996) [Phys. Solid State **38**, 305 (1996)].
[8] Yu. S. Gordeev, M. V. Gomoyunova, V. M. Mikoushkin *et al.*, in *Abstracts of Contributed Papers of the 20th International Conference on the Physics of Electronic and Atomic Collisions (XX ICPEAC)*, Vienna, 1997, Paper FR 107.
[9] V. M. Mikoushkin, *Contributed Papers of the Sixth EPS Conference on Atomic and Molecular Physics*, Siena, Italy, 1998, Inv 41.
[10] Yu. S. Gordeev, M. V. Gomoyunova, A. K. Grigor'ev *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 2388 (1994) [Phys. Solid State **36**, 1298 (1994)].
[11] S. A. Gurevich, T. A. Zaraĭskaya, and V. M. Mikushkin, Fiz. Tverd. Tela (St. Petersburg) **39**, 1889 (1997) [Phys. Solid State **39**, 1691 (1997)].
[12] Yu. S. Gordeev, S. A. Gurevich, T. A. Zaraiskaya *et al.*, in *Abstracts of Invited Lectures and Contributed Papers, International Symposium ''Nanostructures: Physics and Technology''*, St. Petersburg, 1998. p. 436.

Translated by R. M. Durham

Paleoastrophysics: progress and prospects

G. E. Kocharov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 90–93 (September 1999)

Results of highly accurate measurements of the radiocarbon concentration in the annual growth rings of trees over the last 400 years are presented. The temporal behavior of the intensity of galactic cosmic rays is reconstructed for the first time for the periods before, during, and after a deep and extended solar activity minimum — the Maunder minimum (1645–1715). It was established that even during the epoch of a deep solar minimum, the intensity of the galactic cosmic rays experienced solar modulation. The time profile of the generation of high-energy gamma radiation from the supernova explosion of 1006 is established. It is shown that processes of particle acceleration and generation of high-energy gamma rays take place with a three-year delay relative to the onset of the optical flare. The time profile of the generation of solar cosmic rays was obtained for the first time over the last thirty cycles of solar activity by precision measurements of the nitrate content in the polar ice cap. It is shown that solar flares accompanied by the generation of cosmic rays occur during the growth and decay phase of solar activity (measured by the Wolf numbers). Research prospects in the field of experimental paleoastrophysics are discussed. [S1063-7842(99)02009-7]

INTRODUCTION

The following definition may be proposed for paleoastrophysics as a scientific field: the study of astrophysical phenomena whose signals reached the solar system before the birth of instrumental astronomy. Instrumental astronomy dates back to the beginning of the seventeenth century as a result of the systematic study of the skies made by Galileo Galilei. Using a small optical telescope made with his own hands, he discovered the satellites of Jupiter and observed the phases of the Moon. He officially reported the observation of sunspots in Padua (Italy) in 1610.

Konstantinov and Kocharov[1,2] have set forth the fundamental ideas of a new scientific field — paleoastrophysics. They showed that by using natural archives, it is possible to study quantitatively a whole range of astrophysical phenomena that occurred in the distant past. Some of these include: long-lived variations in the intensity of galactic and solar cosmic rays, flare and modulation activity of the Sun, catastrophic events in the past, amplitude–time characteristics of supernova explosions, and so on.

NATURAL ARCHIVES OF COSMIC PARTICLES AND RADIATION

The Earth's crust is a permanent detector of cosmic particles and radiation. Wolfsberg and Kocharov[3] showed that only the crust offers the fundamental possibility of determining the dynamics of the generation of thermonuclear energy in the Sun's interior on a time scale of the last tens of millions of years. The principles of neutrino paleoastrophysics were developed in Ref. 3. It was shown that under the action of solar neutrinos, specific isotopes of lead and technetium are generated in the Earth's crust, and by measuring their content it is possible to determine the fundamental importance of the energy release dynamics in the deep interior of the Sun. This experiment is extremely complex and expensive and requires the collaboration of scientists from different countries.

The annual growth rings of trees have become a traditional source of quantitative data on the time variation in the intensity of galactic cosmic rays on the time scale between the present day and ten thousand years ago. The width of the annual rings and their isotopic content contain information on solar activity and climatic effects over a large time scale in the past. It was recently established[4] that the dynamics of solar activity up to 25 million years ago can also be reconstructed from the width of these rings.

The polar ice cap is an important source of quantitative information on solar flare activity, supernova explosions, and climatic effects. The frequency and amplitude of solar-flare protons are determined from the nitrate concentration in dated layers of the polar ice cap. The amplitude-time characteristics of supernova explosions are established by measuring the time profile of the concentration of the cosmogenic isotopes ^{14}C, ^{10}Be, and ^{36}Cl in independently dated samples of polar ice. These isotopes are formed in nuclear reactions in the Earth's atmosphere under the action of galactic cosmic rays assumed to originate from supernova explosions. These isotopes are also generated under the influence of gamma rays of cosmic origin. According to current thinking on supernova explosions, both high-energy protons and hard gamma rays are formed.

The time scale for obtaining astrophysical information using the polar ice cap is now hundreds of thousands of years.

It has been shown[5] that the outer centimeter layer of lunar soil contains quantitative information on the time profile and power of solar-flare protons on a time scale of the last few hundred thousand years. Another possible method of studying paleosolar flares also exists.[6] We know that various

isotopes are generated in the Sun's atmosphere under the action of particles accelerated in solar flares. These transmutations are taking place on the Sun at the present time and also occurred in the past. This implies that the chemical and isotopic composition of the solar matter vary with time and this variation should in turn affect the Sun's atmosphere. Consequently, an experimental determination of the isotopic composition of the solar wind can be used to establish the intensity of the solar flares and nuclear reactions in the past, the mechanism for the formation of the solar wind, the dynamic characteristics of physical processes in various regions of the solar atmosphere, in the subphotosphere layers, and so on. Comprehensive information is contained in the nuclear gamma lines from the quiet Sun, which can be recorded after increasing the sensitivity by some tens of times over the present level.

The idea of experimental paleoastrophysics, formulated in 1965 (Ref. 1), has been developed rapidly over the last few decades as a result of the efforts of scientists from many countries, and can now be divided into three fields in terms of aims and experimental methods.

Paleoastrophysics of cosmogenic isotopes is based on the following: Cosmic rays generated in natural particle accelerators in space continuously bombard the Earth's atmosphere, initiating various nuclear reactions. Radioactive nuclei such as ^{14}C,^{10}Be, and ^{36}Cl are generated in these reactions. For these isotopes there are natural archives with good retention and accurate time markers. Tree rings are the best for radiocarbon, and the polar ice cap is best for ^{10}Be, and ^{36}Cl. By measuring the content of cosmogenic isotopes in independently dated samples, it is possible to reconstruct the cosmic ray intensity over a long time interval, and obtain information on the source of the cosmic rays (supernova explosions, solar flares) and on the characteristics of the interstellar medium, solar activity, and the entire heliosphere.

Paleoastrophysics of nitrates is based on highly accurate measurements of the nitrate content in the polar ice cap. Independent methods have been developed for dating ice, and yearly nitrate measurements are already available for the last 400 years. Nitrates in the polar ice cap are not only a unique archive for the study of astrophysical phenomena in real time but also in the distant past (supernova explosions, solar flare activity, and deep solar minima). Unlike cosmogenic isotopes for which nuclear interactions are typical, nitrates are sensitive to the low-energy range: atomic collisions, temperature effects, changes in the ambient medium, and so on.

THE PROBLEM OF DEEP EXTENDED SOLAR ACTIVITY MINIMA

According to historical data, in the past history of the Sun there were long time intervals when not a single sunspot was observed on the surface of the Sun. The nearest deep extended solar minimum to the present day occurred over the period 1640–1715. This minimum is called the Maunder minimum, after the English scientist who published a work on the existence of this minimum in 1921.

Over the last thirty years, since the idea of the specific possibilities of paleoastrophysics was first formulated, an extensive series of highly accurate measurements have been made of the radiocarbon concentration in the annual growth rings of trees over the last 400 years in order to determine the nature of the deep solar minimum and the nature of the solar modulation of the intensity of galactic cosmic rays before, during, and after the Maunder minimum. In order to study solar flare activity over the same time interval, yearly measurements have been made of the nitrate content in the polar ice cap. Nitrates are a unique source of yearly variation in the intensity of solar cosmic rays which provide experimental data on the power, frequency, and time of flare generation of the cosmic rays.

Radiocarbon studies have been carried out by groups in Russia, Lithuania, Ukraine, and Georgia under the scientific guidance of the present author. Highly accurate yearly measurements of the nitrate content in the polar ice cap were made jointly by the Ioffe Institute and the University of Kansas, and the results are presented in Refs. 7 and 8. Here we merely describe the main features of the Maunder minimum.

1. The radiocarbon content in the Earth's atmosphere in the epoch of a deep minimum is $2.3 \pm 0.3\%$ higher than that before and after the minimum. This increase was predicted, and the magnitude of the effect was established for the first time.

2. The transition of the solar activity from the normal level to a deep minimum and back takes place over a period of no more than a year.

3. The main characteristics of the eleven-year cycle of solar activity before and after a deep minimum were the same as those over the last fifty years, according to the results of direct measurements.

Results of measurements of the nitrate content in the polar ice cap over the last 400 years reliably established the following.

1. No solar flare activity was observed during the Maunder minimum.

2. In the epochs before and after the Maunder minimum, solar cosmic rays were reliably recorded. Of fundamental importance is that the cosmic rays are generated during the growth and decay in the numbers of sunspots. This effect was observed in direct experiments over the last few decades. Increasing the time span to 400 years (around 40 cycles of solar activity) provides evidence to support the conclusion that solar cosmic rays are generated during the growth and decay of sunspot-forming solar activity.

The clearest, and most unexpected, experimental observation is that the intensity of the galactic cosmic rays varies in the epoch of the Maunder minimum, although the nature of this variation differs from the well-known eleven-year cycle of solar activity. A twenty-two year cycle was observed during the Maunder minimum. This corresponds to the period of polarity reversal of the Sun's general magnetic field. A theory for this effect has not yet been developed. Bearing in mind the importance of this effect and the fact that the duration of the Maunder minimum was only seventy years (three periods of twenty-two year cycles), we need to study the time variations for the Spörer minimum, which was of relatively long duration (1416–1534), i.e., more than five

twenty-two year cycles. The problem of the Spörer minimum is now the central issue for radiocarbon and nitrate measurements.

THE ORIGIN OF COSMIC RAYS

The origin of cosmic rays has been one of the key issues in high-energy astrophysics for several decades. Fundamental and guiding ideas on this subject were put forward by Ginzburg (see Ref. 9 and the literature cited therein). The fundamental idea proposed by Ginzburg that the recorded cosmic rays are generated in our Galaxy has already been confirmed experimentally. According to Ginzburg,[9] these mainly originate from supernova explosions. This idea cannot be checked out experimentally in real time, because the required duration of an experiment to record cosmic rays from a supernova explosion is tens of thousands of years. The only possibility is to use paleoastrophysics methods,[10] such as recording high-energy gamma rays from the explosions of historic supernovas and high-energy protons from a close supernova (less than a few hundred light years away).

The idea[1] was to make precision measurements of the radiocarbon content in the annual growth rings of trees over a time interval embracing the explosions of two supernovas: the supernovas of 1004 and 1572 (Tycho Brahe). In both cases the observed effect is caused by high-energy γ-rays, generated in nuclear reactions initiated by protons accelerated in the explosions. The time profile and total explosion energy were determined.[10,11] The rise time of the γ-ray fluxes is 3–5 years and the decay time ~ 10 years. The total explosion energy is 10^{50} erg.

These results fully support the Ginzburg's idea that supernova explosions in our Galaxy are the main source of galactic cosmic rays.

CONCLUDING REMARKS

1. The problem of deep solar minima

It is extremely important to make highly accurate yearly measurements of the concentration of cosmogenic isotopes for a deep and extended solar activity minimum — the Spörer minimum (1450–1550). If it is established that in this case the modulation of the cosmic rays also has a twenty-two year cycle, the problem of modulation will undoubtedly become one of the central issues in the astrophysics of cosmic rays and solar physics.

2. The origin of cosmic rays

It is impossible to establish the time profile and determine the total proton energy from the explosion of historic supernovas, since the protons from these have not yet reached the Earth. Moreover, when they do reach the Earth, it will not be possible to identify them against the existing background. Thus, we are only concerned with supernovas which have exploded in the far distant past, close to the Earth.

An analysis of experimental data on the concentration of the cosmogenic isotopes ^{14}C and ^{10}Be in dated natural archives led the author to conclude[12] in 1982 that 35 000 years ago there was a supernova explosion at a distance of 160 light years from the Earth. The total energy of the cosmic rays is 10^{50} erg and shows good agreement with the theory. However, it must be borne in mind that the enhanced concentration of cosmogenic isotopes established may be a consequence of the reduced geomagnetic level in the same epoch, with a corresponding time profile. We suggested[10] a specific possibility for determining the nature of this phenomenon. The Farmington meteorite has a cosmic age of 40 000 years, so that it should have been irradiated by cosmic rays from the supernova explosion being discussed. As a result of this irradiation, ^{81}Kr atoms should have accumulated in the meteorite as a result of nuclear reactions. Hence, determining the content of these nuclei in the Farmington meteorite should answer the question as to the nature of this phenomenon. In order to achieve this, the sensitivity of counting the ^{81}Kr atoms must be improved tenfold. Another possible method of recording cosmic rays from a supernova explosion also exists. Here we are talking of the unique Geminga object, which is 300 000 years old and 300 light years away. Geminga is a powerful source of high-energy gamma rays, so this unique object should be a strong source of cosmic rays. Highly accurate measurements must be made of the cosmogenic isotope ^{10}Be in the polar ice cap over the last few thousand years. This is a difficult task but is possible in principle.

To conclude, the author would like to stress that the possibilities of experimental paleoastrophysics are enormous and embrace a wide range of problems in physics and astrophysics.

The author would also to specially mention the initiating role of B. P. Konstantinov in creating this new, now widespread field of high-energy astrophysics.

This work was supported by the Russian Fund for Fundamental Research, Grant No. 990218398.

[1] B. P. Konstantinov and G. E. Kocharov, Dokl. Akad. Nauk **165**, 63 (1965).
[2] B. P. Konstantinov and G. E. Kocharov, Preprint No. 64 [in Russian], Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad (1967), 54 pp.
[3] K. Wolfsberg and G. E. Kocharov, in *The Sun in Time*, edited by C. P. Sonett, M. S. Giampapa, and M. S. Matthews (Tucson 1991), pp. 288–313.
[4] S. Cecchini, M. Galli, T. Nanni, and L. Ruggiero, Nuovo Cimento C **19**, 527 (1996).
[5] R. C. Reedy and K. Marti, in *The Sun in Time*, edited by C. P. Sonett, M. S. Giampapa, and M. S. Matthews (Tucson 1991), pp. 260–287.
[6] G. E. Kocharov, Izv. Ross. Akad. Nauk, Ser. Fiz. **60**(8), 112 (1996).
[7] G. E. Kosharov, V. M. Ostryakov, As. N. Peristykh, and V. A. Vasiliev, Sol. Phys. **159**, 381 (1995).
[8] G. E. Kocharov and M. G. Ogurtsov, *Contemporary Problems in Solar Cyclicity* [in Russian], GAO, St. Petersburg (1997).
[9] V. L. Ginzburg, Usp. Fiz. Nauk **166**(2), 169 (1996).
[10] G. E. Kocharov, Nuovo Cimento C **19**, 883 (1996).
[11] P. Damon, D. Kaimei, G. Kocharov *et al.*, Radiocarbon **37**, 599 (1995).
[12] G. E. Kocharov, in *Proceedings of the 12th Leningrad Seminar on Cosmic Physics* [in Russian], Physicotechnical Institute, Russian Academy of Sciences (1982), pp. 203–207.

Translated by R. M. Durham

Investigation of heavy cosmic-ray nuclei by solid-state track detectors on orbiting platforms

D. G. Baranov, Yu. F. Gagarin, V. A. Dergachev, and E. A. Yakubovskiĭ

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 94–98 (September 1999)

Results are presented from investigations of Fe-group nuclei in galactic and solar cosmic rays in the energy ranges 30–210 MeV/n and 7–210 MeV/n in various phases of the solar cycles. Spectra of Fe particles have been obtained with high energy resolution as a result of exposing PLATAN chambers, made up of layers of a polyethyelene terephthalate solid-state track detector, for between one and three years on the Salyut-6, Salyut-7, and Mir space stations, orbiting at an inclination of 51.6° to the plane of the equator and at an altitude of 350–400 km. Measurements were made of the energy spectra of Fe particles from a unique event, the series of solar flares of September 29 and October 19–29, 1989, which is the most powerful of those recorded and measured over the entire history of cosmic ray research. A modern model of particle penetration inside the Earth's magnetosphere, developed at the Institute of Nuclear Physics at Moscow State University, is used to compare the measured spectrum with measurements made using the solid-state track detector in the HIIS experiment on the LDEF station, and with extramagnetospheric measurements made using electronic equipment on the IMP-8 satellite and the Galileo space station. It is shown that the solid-state track detector technique has advantages for obtaining the characteristics of the energy spectrum. [S1063-7842(99)02109-1]

INTRODUCTION

Until recently, the low-energy region of heavy nuclei in cosmic rays between tens and several hundreds of megaelectron-volts per nucleus had been the least well studied and the most difficult to interpret. This can be attributed to the relatively weak particle fluxes and the broad spectrum of sources of these nuclei, including galactic, solar, and anomalous sources, and, for measurements within the magnetosphere, trapped particles. Outside the Earth's magnetosphere, only the integral particle flux had been measured for galactic nuclei in this energy range and no energy spectra had been obtained, while the energy spectra of Fe particles of solar origin were limited to a maximum energy of 20–30 MeV/n until the end of 1996. Nowadays, measurements of the energy spectra of heavy nuclei inside the magnetosphere, made using long exposures of solid-state track detectors on orbiting space stations, can provide new and more detailed information on heavy nuclei, including the profile of the energy spectra of galactic cosmic rays (GCRs), its variation during a solar cycle, and determination of the degree of ionization of heavy particles in solar cosmic rays (SCRs). These data can be used to check the accuracy of various models of particle penetration inside the magnetosphere by comparing these with measurements outside the magnetosphere and also inside the magnetosphere for various orbit inclinations.

METHOD

Research on a solid-state track detector was started some twenty-five years ago in the Cosmic Ray Laboratory of the A. F. Ioffe Physicotechnical Institute. This resulted in the development of the technology and procedures for a Russian solid-state track detector using polyethyelene terephthalate, specially designed to record and identify $Z \geqslant 20$ heavy nuclei stopped in the detector layers.[1,2] This technology includes the conditions for exposure to the detector by ultraviolet radiation after prolonged exposure in outer space in order to restore and enhance the detector sensitivity, and conditions for etching the detector in an alkali solution to identify particle tracks. Second, we developed the PLATAN-1, -2, -3, -4, and -5 units, with geometric factors of 0.25–1 m^2sr, which were successfully deployed by the crews of nine missions in which the apparatus was mounted and exposed on the exterior surface of the Salyut-6, Salyut-7, and Mir space stations and then demounted. The nuclear charge was determined by the LR method, using the dependence of the lengths L_i of the track cones etched on the upper and lower surfaces of the detector layer on the range R_i measured from the particle stopping point in the detector. Figure 1 gives these dependences for SCR particles having charges between 18 and 28. The most common element, iron, can be identified from the cluster of points on the diagram. It can be seen that elements between Ca and Fe are almost absent. The charge distributions for GCRs (Ref. 3) and SCRs (Ref.4) are plotted in Figs. 2a and 2b, respectively. The charge resolution achieved for GCR and SCR Fe nuclei is 0.25 and 0.35 charge units, which is substantially better than the resolution of 0.45 and 1 charge unit achieved in the more sensitive solid-state track

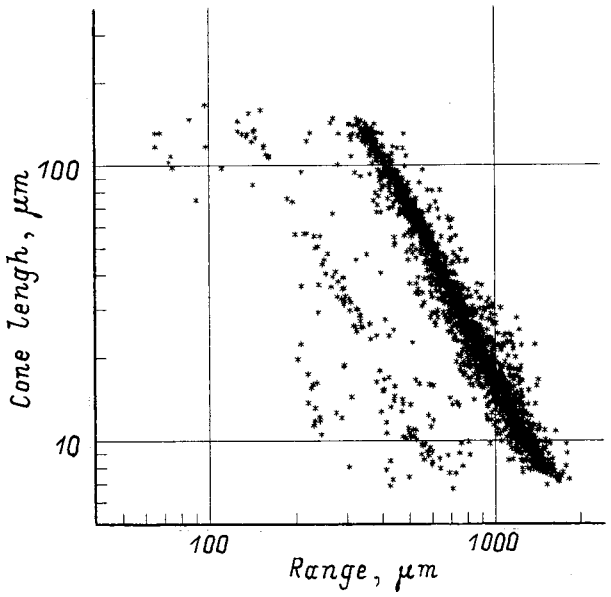

FIG. 1. Cone lengths of SCR heavy particle tracks in detector layers as a function of residual range.

detectors made up of a CR detector and cellulose nitrate, and is the same as the resolution of the electron detector on the Galileo station.

RESULTS AND DISCUSSION

Figure 3 gives the energy spectrum of GCR Fe nuclei in the energy range 40–210 MeV/n measured for the first time inside the Earth's magnetosphere over the period 1984–1985 in the PLATAN-2 experiment. It is observed that the flux increases by almost an order of magnitude as the energy increases. The increase in flux observed with decreasing energy in the range 30–40 MeV/n is determined by the contribution of Fe particles from solar flares in April and July 1985. For comparison Fig. 3 gives the energy spectrum of GCR Fe nuclei measured on Voyager during the maximum GCR intensity in 1986–1987 at 28 AU (curve *1*), and calculated curves *2* and *3* of GCR Fe fluxes outside the Earth's magnetosphere for April 1985 and October 1984. Spectra *2* and *3* were obtained indirectly using measured spectra of helium and oxygen, with allowance for the relative Fe/He and Fe/O abundance. Note that the absolute particle fluxes measured within the magnetosphere are twenty times lower than those measured on Voyager. Despite the large difference between the fluxes, the energy resolution of the spectra in our experiment was comparable with the Voyager data.

Preliminary results of measurements of the Fe nuclei flux from the PLATAN-4 chamber in 1994–1995 near the GCR intensity maximum, indicate that at energies of 100–150 MeV/n the flux increases by almost an order of magnitude. A comparison between the experimental spectrum of GCR Fe nuclei in 1984–1985 and calculations made by various authors (see the literature cited in Ref. 5), which show good agreement among themselves, reveals that the experimental values are almost twice the calculated ones (Fig. 3, curve *4*) in the energy range 100–150 MeV/n. This difference could be attributed to the incomplete ionization of low-energy GCR Fe particles as a result of which this flux is attenuated in the magnetosphere considerably less than a flux of completely ionized particles. However, recent calculations[6] using the model developed at the Institute of Nuclear Physics at Moscow State University, which more accurately allows for the change in the particle transmission coefficient inside the magnetosphere and the average geomagnetic disturbance over the 1984–1985 year of exposure, agree with the experimental data. A detailed analysis of the experimental data for GCR heavy nuclei at energies up to several hundred mega-electron volts per nucleus showed that the fraction of incompletely ionized heavy particles in GCRs at these energies is negligible.[7]

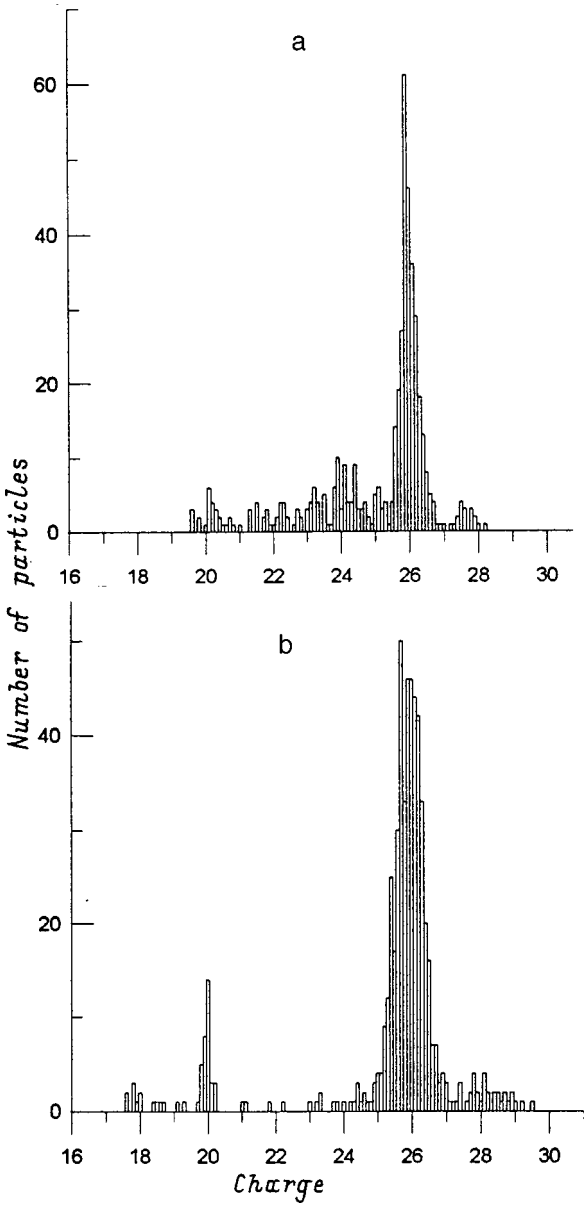

FIG. 2. Measured charge distributions of heavy nuclei: a — GCR particles, b — SCR particles.

In the first three experiments (PLATAN-1, -2, and -3) we observed heavy nuclei from solar flares in September

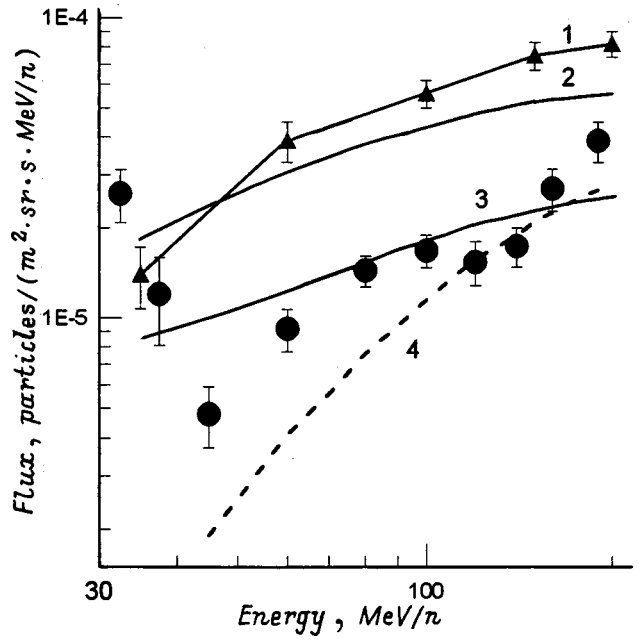

FIG. 3. Energy spectrum of Fe nuclei: on the orbit of Salyut-7 in 1984–1985 (●), Voyager in 1986–1988 at 28 AU (▲ and *1*), calculations for May 1985 and October 1984 (*2* and *3*), and calculations for Salyut-7 orbit (*4*). For curves *1–3* the fluxes were reduced tenfold.

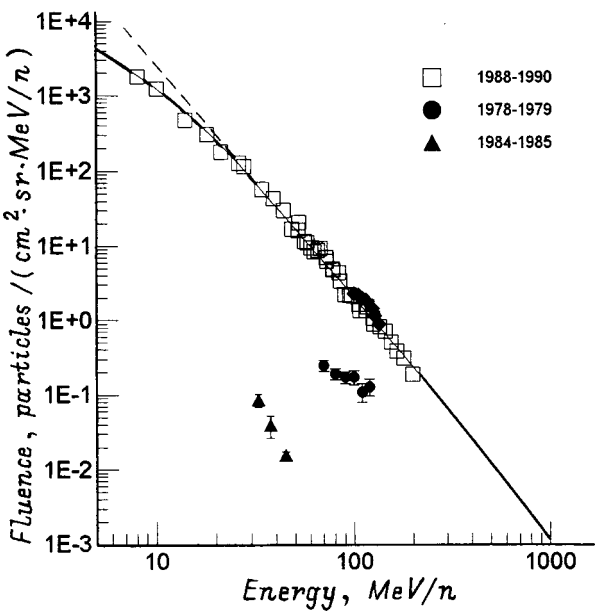

FIG. 4. Energy spectra of SCR Fe particle fluences for three PLATAN exposures. Solid curve — approximation.

1978, April and July 1985, and from a series of flares in September–October 1989 (Fig. 4).

During the 1988–1990 period of exposure, the PLATAN-3 chamber registered a unique phenomenon: powerful fluxes of Fe particles from a series of solar flares on September 29–October 4 and October 19–29, 1989. The particle fluxes are at least an order of magnitude higher than those measured earlier in similar events. The background of GCR Fe nuclei in this exposure is almost negligible. The Fe particle fluence at the maximum energy in the spectrum, 200 MeV/n, is an order of magnitude higher than the total two-year background of GCR nuclei. At low energies this difference increases by a further three orders of magnitude. In order to determine the orientation of the PLATAN-3 chamber relative to the recorded particle flux during the time of the flares, we used data on the angular distribution of particles entering the chamber in the total solid angle 4π. The working surface of the chamber was perpendicular to the longitudinal axis of the station. The counterflux of Fe particles entering through the bottom of the chamber was equal to the direct flux in the range 100–130 MeV/n (Fig. 4), which indicates that the irradiation was isotropic during this series of flares.

We used the *LR* method to construct the energy spectrum in the range 45–210 MeV/n to determine the particle charge. A fast method of estimating particle charge was used to increase the rate of processing in the energy range 7–130 MeV/n. This method can separate $Z \geqslant 20$ particles from lighter $Z \leqslant 16$ particles according to the number of etched-through detector layers without making detailed measurements. The energy spectrum of the Fe particles was obtained by subtracting the fraction of Fe particles from the flux of $Z \geqslant 20$ particles on the basis of the measured particle charge distribution (Fig. 2b). The complete agreement between the Fe particle spectra in the energy range 45–130 MeV/n obtained by the two different methods indicates that this fast method of obtaining the spectrum of SCR Fe particles is reliable and efficient.

On the basis of our measurements, the energy spectrum of the Fe particles reveals a kink at ~30 MeV/n. In the range 30–100 MeV/n the spectrum is approximated by a straight line (dashed line in Fig. 4) which is extrapolated to lower energies. Below 26–30 MeV/n the spectrum departs from a straight line. For particles with $Z = 12$–16 a sharper kink is observed in the spectrum at this energy (not shown in Fig. 4. This difference in the behavior of the spectra of light and heavy particles within the Earth's magnetosphere can be attributed to their different degrees of ionization: light SCR particles are almost completely ionized whereas Fe particles have charges of ~14–15. The spectrum of heavy particles becomes less deformed in the magnetosphere since these particles penetrate as far as the station orbit over a wider range of latitudes.

Note that for light SCR particles the kink in the spectrum was observed earlier since the spectra of light particles are measured in a larger energy range. The kink in the Fe particle spectrum observed for the first time at 30 MeV/n is first, indicative of a general feature in the profile of the energy spectra which does not depend on the particle charge, suggesting that these spectra have a common mechanism of formation, and second indicates that the spectral profile does not depend on the power of the solar flare. At the preliminary stage of research using the Institute of Nuclear Physics model, we compared the Fe particle spectrum obtained in the energy range 30–130 MeV/n with that measured using the solid-state track detector in the HIIS experiment on the LDEF station with its orbit inclined at 28° (Fig. 5, Ref. 8). We used our experimental data to reconstruct the extramag-

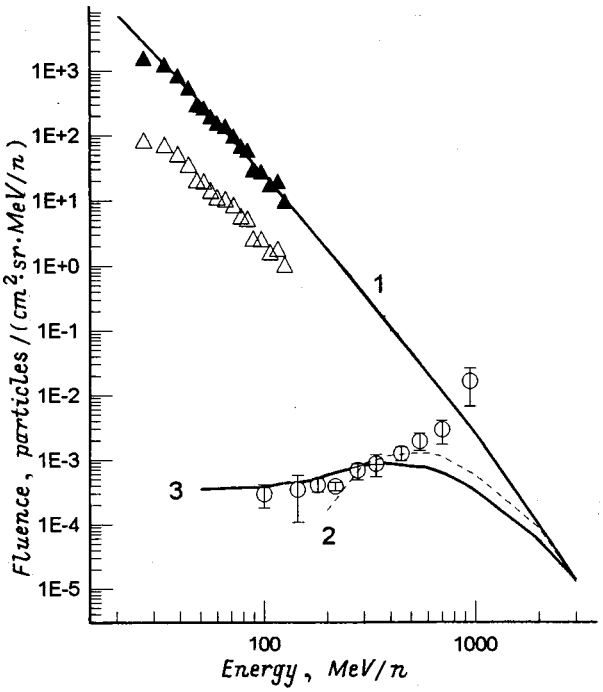

FIG. 5. Energy spectra of Fe particles from a series of solar flares in September–October 1989: PLATAN experiment (51.6° orbit) (△), HIIS experiment (28° orbit) (○), calculated spectrum outside Earth's magnetosphere (▲), extrapolation of calculated spectrum (*1*), and calculated spectra inside the Earth's magnetosphere for a 28° orbit (*2* and *3*).

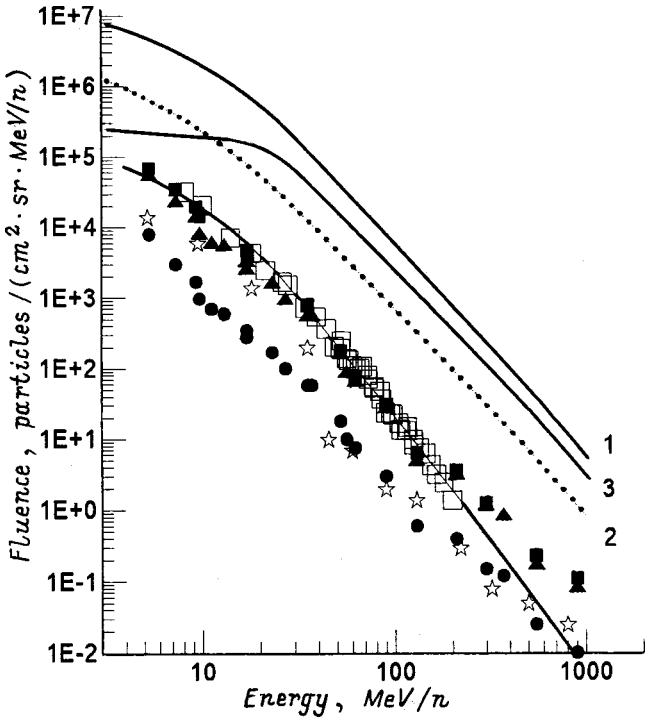

FIG. 6. Energy spectra of Fe and protons outside the Earth's magnetosphere for events in September–October 1989. Curves — proton fluences (reduced by a factor of 100): *1* — October 19–27, *2* — October 24–27, *3* — September 29–October 6, □ — PLATAN-3 experimental data converted to extramagnetospheric space, *4* — approximation, ● — event of October 24–27, ▲ — sum of events of October 19–27, ☆ — September event, and ■ — sum of September and October events.

netospheric spectrum of Fe particles and then extrapolated to the high energy range (curve *1* in Fig. 5). By allowing for the transformation of this spectrum as the particles penetrated to an orbit inclined at 28°, we were able to make a comparison with the results of measurements made on the LDEF station. The calculations made for various modifications of the model (curves *2* and *3* in Fig. 5) show satisfactory agreement with the HIIS experimental data. This analysis allowed us to estimate the degree of ionization of SCR Fe particles at energies of 50 and 500 MeV/n which varies between 14.0 ± 2.6 and 15.7 ± 3.1. The possibility of matching data obtained under such different experimental conditions indicates that the model used is quite adequate.

Equally interesting is a comparison of the experimental data on particle fluxes obtained inside and outside the magnetosphere. Unfortunately Fe particle fluxes outside the magnetosphere have only been measured using electronic apparatus for the flare dated September 29–October 4, 1989 using two devices on the IMP-8 satellite in the energy ranges 5–50 and 50–800 MeV/n. For the event dated October 24–27 the spectrum was also measured using equipment on Galileo in the range 7–8 MeV/n as well as these two devices (Fig. 6). The spectrum measured in the PLATAN-3 chamber covers the energy range of these three electronic devices, and the energy resolution achieved in the 30–80 and 50–210 MeV/n ranges is four and six times higher. Figure 6 shows the extramagnetospheric spectrum reconstructed using our data. The calculations allowed for the exact movement of the station over the orbit during the solar flares and the variation in the geomagnetic disturbances which influence the degree of deformation of the measured spectra. The total Fe particle flux from the entire series of flares outside the magnetosphere was obtained using data from IMP-8 and Galileo and proton fluxes (data from the HEAO satellite, curves *1–3* in Fig. 6). Assuming that the Fe/*p* ratio for the entire series of flares in October 1989 was the same as that for the event of October 24, we reconstructed to the total Fe particle flux. A comparison between the characteristics of the spectra of Fe particles and protons in Fig. 6 shows that for SCR flares the ratio Fe/*p* decreases with increasing energy. The deformation of the initial energy spectrum of the SCR Fe particles as the particles penetrate inside the magnetosphere to an orbit inclined at 51.6° reduces the flux by a factor of 15 at 8 MeV/n and by a factor of 6 at 200 MeV/n.

New experiments using the PLATAN apparatus are planned in the Russian section of the International Space Station to continue our systematic studies of Fe-group GCR and SCR heavy nuclei and the search for low-energy superheavy GCR and SCR nuclei.

This work was supported by the RFBR (Grants Nos. 94-02-06387 and 96-02-17965).

[1] Yu. F. Gagarin, I. V. Gordeev, N. S. Ivanova *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **47**, 1844 (1983).
[2] D. G. Baranov, V. A. Dergachev, Yu. F. Gagarin *et al.*, Radiat. Meas. **25**, 301 (1995).
[3] D. G. Baranov, V. A. Dergachev, Yu. F. Gagarin *et al.*, in *Proceedings of the 24th International Cosmic Ray Conference*, Rome, 1995, Vol. 4, pp. 1051–1054.

[4] D. G. Baranov, Yu. F. Gagarin, V. A. Dergachev *et al.*, in *Proceedings of the Conference on ''Contemporary Problems of Solar Cyclicity''* [in Russian], St. Petersburg, 1997, pp. 255–257.
[5] D. G. Baranov, Yu. F. Gagarin, V. A. Dergachev *et al.*, Geomagn. Aeron. **35**(4), 144 (1995).
[6] R. A. Nymmik, Doctoral Dissertation [in Russian], Moscow (1998).
[7] Yu. F. Gagarin, R. A. Nymmik, and M. I. Panasyuk, in *Proceedings of 25th International Cosmic Ray Conference*, Durban, 1997, Vol. 3, pp. 293–296.
[8] Yu. F. Gagarin, A. M. Marenny, R. A. Nymmik, and M. I. Panasyuk, Adv. Space Res. **21**, 1699 (1998).

Translated by R. M. Durham

The search for antimatter in cosmic rays

É. A. Bogomolov, G. I. Vasil'ev, S. Yu. Krut'kov, V. A. Romanov, S. V. Stepanov, and M. S. Shulakova

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted January 14, 1999)
Zh. Tekh. Fiz. **69**, 99–103 (September 1999)

The search for antimatter in the universe is a page in the history of the Ioffe Physicotechnical Institute (IPTI). Experiments on spacecraft and high-altitude balloons, begun in the 1960s, yielded information on to the presence or absence of antimatter stars or galaxies according to evidence arising in explosive processes in these objects. Antiprotons with energies of 2–5 GeV in galactic cosmic rays were observed at the end of the 1970s in balloon experiments by the Cosmic Spectrometry Laboratory at the IPTI. These studies were done using a magnetic spectrometer at altitudes with a residual pressure of $10\,g/cm^2$ with a threshold geomagnetic rigidity of 3 GV. High-latitude experiments in the 1980s, yielding the first measurements of the flux of galactic antiprotons with energies of 0.2–2 GeV, gave some indication of the mechanism by which they are generated. The measured ratios of the fluxes of antiprotons and protons in the cosmic rays are $2.4^{+2.4}_{-1.3}\times10^{-4}$ and $6^{+14}_{-5}\times10^{-5}$ at energies of 2–5 and 0.2–2 GeV, respectively. Subsequent balloon-borne experiments employing magnetic spectrometers by groups from the USA and Japan have confirmed the results obtained by the IPTI. Experimental and theoretical work on the search for antiparticles in cosmic rays is summarized and the astrophysical consequences of this research are discussed. Experimental data on the detection of antiparticles in galactic cosmic rays indicate that there are no objects made of antimatter within the local group of galaxies. [S1063-7842(99)02209-6]

INTRODUCTION

The search for antimatter in the universe, one of the most interesting areas of astrophysics, was begun at the A. F. Ioffe Physicotechnical Institute (IPTI) in the early 1960s at the initiative of B. P. Konstantinov. Data from elementary particle physics suggested a baryon symmetry in the universe and the possible existence of antimatter objects in the universe in the form of antistars, antigalaxies, or aggregations of them. Explosive processes in these hypothetical objects should result in the ejection of antimatter into space in the form of antiparticles, dust, or comets; hence, the major effort in the experiments on the antimatter problem at the IPTI was to search for antinuclei and antiprotons in cosmic rays and to detect the possible entry of antimeteors into the Earth's atmosphere. The experiments were conducted using high altitude balloons and spacecraft. A search for antiparticles in cosmic rays using balloons was also begun in the early 1960s by groups from the USA, Japan, and India.

The use of spacecraft ensures greater exposure than balloon studies. For example, the search for antimatter carried out by the IPTI in the mid-1970s on the Zond-5 and -7 spacecraft and the Kosmos and Soyuz series satellites by means of emulsion chambers yielded a better upper limit for the ratio of antinuclei to nuclei with charges exceeding 3, at a level of 2.7×10^{-4} for energies below 1.2 GeV/nucl.[1] This work also laid the foundation for subsequent studies at the IPTI of nucleus–nucleus interactions at high energies in space.

Experiments to detect the entry of micrometeors into the Earth's atmosphere in the early 1960s during meteor showers on airplanes at altitudes of 13–18 km revealed an enhancement in the intensity of the hard gamma radiation and neutrons that correlated with the formation of a meteor wake at a height of about 100 km as recorded by radar. The effect was roughly 2% above background and exceeded the statistical error by a factor of 6.[2] Of the 20 meteor shower surges in the period 1961–1964 a positive effect was observed in 18.

Experiments to observe the entry of meteor showers into the Earth's atmosphere in 1966–1967 in the orbital path of the scintillation gamma spectrometer on the Kosmos-135 satellite recorded an $\sim$50% increase in the intensity of the 511 keV annihilation line (at a level of 7.5 standard deviations) during the active period of two of the three showers that were studied.[3] The experiments on the satellite and in airplanes had a high statistical reliability but have not been explained yet. The orbital experiments to detect gamma radiation during the entry of micrometeors laid the foundation for future successful work at the IPTI on x-ray and gamma astronomy and on the problems of the dust component in space.

Balloon experiments to search for antiprotons in galactic cosmic rays begin in the end of the 1960s detected the first antiprotons with energies of 2–5 GeV at the end of the 1970s.

ANTIPROTONS IN COSMIC RAYS

A balloon borne magnetic spectrometer, which first flew in the autumn of 1969, was used to search for galactic antiprotons. This was the second spectrometer lifted into the stratosphere, following an American spectrometer built for measuring a positron spectrum. The measurements in its first flight yielded an upper limit for the ratio of the fluxes of antiprotons and protons in cosmic rays at a level of 10^{-2} (Ref. 4).

The magnetic spectrometer consisted of: a monitor telescope of scintillation detectors which determined the solid angle of the instrument for detecting particles and made it possible to distinguish singly charged particles from the amplitude of the signals in the scintillators; systems of spark chambers for determining the trajectories of particles through the device; a permanent magnet which deflected charged particles and, together with the trajectory measurements, made it possible to measure the rigidity spectrum of the particles and to separate the particles according to the sign of the charge; a threshold gas Čerenkov detector which permitted velocity selection of the particles and, together with data on the rigidity of the detected particles, made it possible to separate particles of the same charge by mass; and, a directional solid state Čerenkov detector that was used for shielding against albedo particles (later, this problem was solved using time-of-flight analysis). Figure 1 shows a conceptual diagram of the magnetic spectrometer. The apparatus was calibrated using the natural background of atmospheric muons, which could be used to measure and monitor the required parameters of the device.

The balloon experiments were done at altitudes with a residual atmospheric pressure of 10–11 g/cm^2 in a region with a geomagnetic cutoff rigidity of 3.2 GV. During the flights, the spectrum of deflections of the singly charged particles in the magnetic field of the spectrometer was recorded (here the deflection is the reciprocal of the rigidity). The threshold Lorentz factor of the gas Čerenkov detector operating in an anticoincidence mode was chosen to be 6.1, which corresponded to threshold rigidities for particle detection of 5.6, 0.8, 0.6, and 0.003 GV for protons, pions, muons, and electrons. Thus, within the deflection region for positively charged particles, primary protons with rigidities of 3.2–5.6 GV, secondary protons from the residual atmosphere at 0.1–5.6 GV, positive atmospheric muons at 0.1–0.6 GV, and positive pions from nuclear interactions in the device at 0.1–0.8 GV were detected. Within the deflection region for negatively charged particles, muons and pions were detected within 0.1–0.8 GV. The range 3.2–5.6 GV of rigidities corresponded to the expected region for detection of galactic antiprotons, while within the range 0.8–5.6 GV, antiprotons from the residual atmosphere could be detected.

The balloon experiments of 1972, 1974, and 1977 recorded events that satisfied all the criteria for sampling of primary antiprotons. The measured ratio antiproton and proton fluxes over 2–5 GeV was estimated to be at a level of $6\pm4\times10^{-4}$ (Ref. 5). A similar result was obtained in a balloon experiment with a magnetic spectrometer by an American collaboration between NASA and the University of New

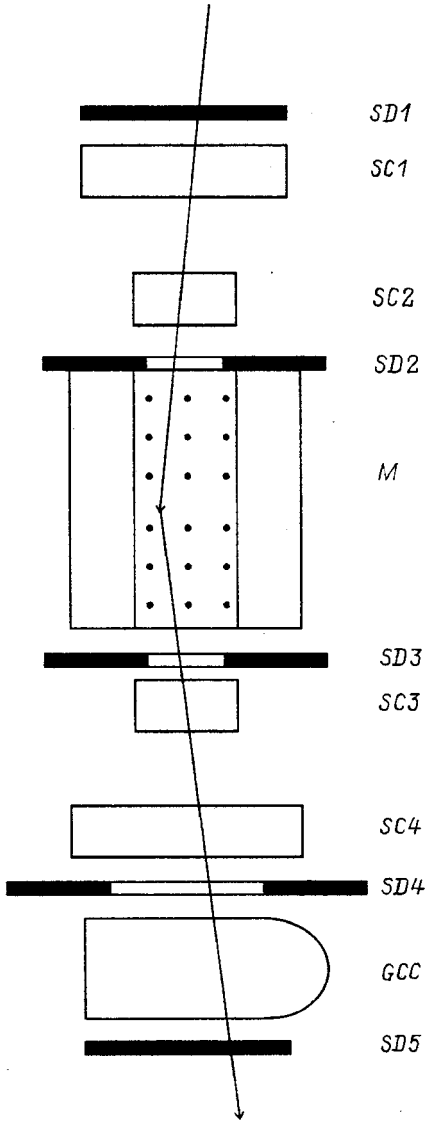

FIG. 1. Balloon-borne magnetic spectrometer from the A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences: SD1–SD5 scintillation detectors, GCC gas Čerenkov counter, SC1–SC4 spark chambers, M magnet.

Mexico in June of 1979. The measured ratio of the antiproton and proton fluxes over energies of 4.7–11.6 GeV was $5.2\pm1.5\times10^{-4}$ (Ref. 6). The results of these first experiments were roughly 5 times the antiproton fluxes expected as a result of interactions of cosmic rays with the interstellar medium and suggested the existence of additional sources of antiprotons in outer space. The situation was further escalated after a balloon-borne experiment in 1980 by a group from the University of California who measured a ratio of the antiproton and proton fluxes of $2.2\pm0.6\times10^{-4}$ at energies of 0.13–0.32 GeV using a device which measured antiproton annihilation events.[7] American measurements of antiproton fluxes at high energies of 10–50 GeV in the same period also yielded positron fluxes an order of magnitude higher than expected from interactions in the interstellar medium. The processes responsible for generating positrons in this energy region could be associated in energy with the creation of antiprotons at energies of the order of a few GeV.

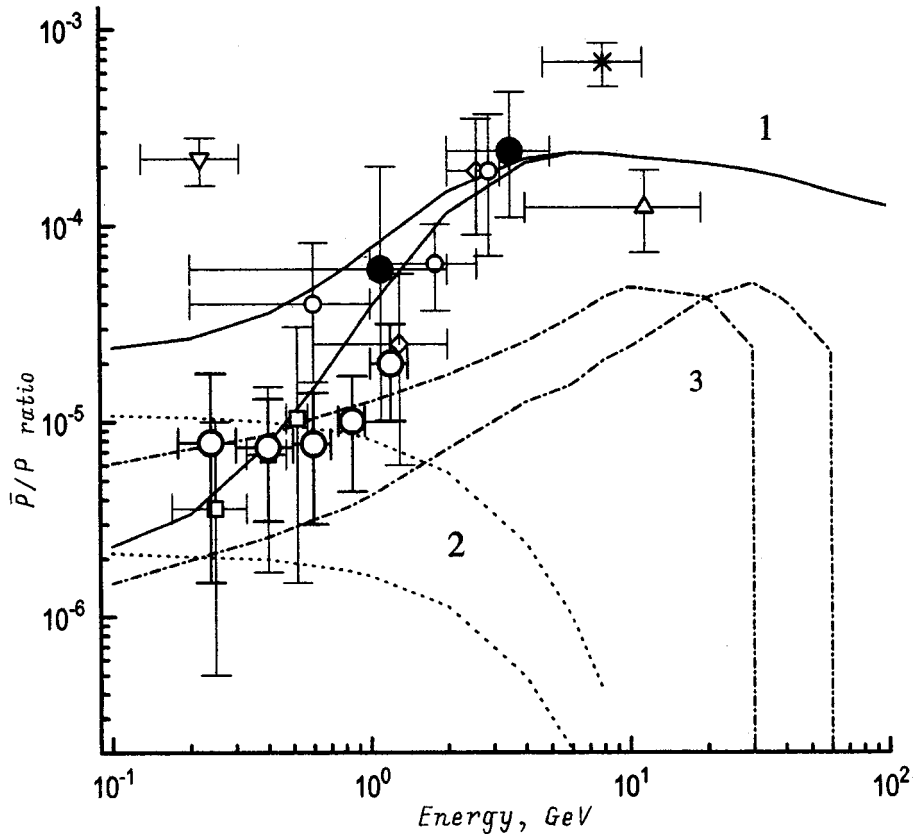

FIG. 2. Experimental and theoretical values of the ratio of the fluxes of antiprotons and protons in galactic cosmic rays. Experiments: ▽ — Buffington,[7] ● — Bogomolov,[8,12] ∗ — Golden,[14] □ — BESS'93, 94,[10] ○ — IMAX,[9] ◇ — CAPRICE,[11] △ — MASS2,[13] ○ — BESS'95.[15] Theory: *1* — interaction of cosmic rays with interstellar material (periods of solar maximum and minimum);[16] *2* —evaporation of original black holes, $(0.2–1)\times10^{-2}$ $ps^{-3}\,yr^{-1}$ (Ref. 17); *3* — annihilation of supersymmetric particles with masses of 30–60 GeV in the dark halo of the galaxy.[18]

The experiments stimulated a series of theoretical studies related to possible sources for the antiprotons. Detailed analyses were done of various models for the creation of antiprotons in interstellar space, clouds of molecular hydrogen, metagalactic sources made of antimatter, the generation of antiprotons in primary black holes, neutron oscillations, ''dark matter'' (photino, higgsino) in the galactic halo, and relativistic astrophysical objects.

In 1983–1985, high latitude balloon experiments were done at the IPTI using magnetic spectrometers in the energy range 0.2–2 GeV. These experiments yielded a ratio of the antiproton and proton fluxes of $0.6^{+1.4}_{-0.5}\times10^{-4}$,[8] which cast doubt on the measurements of the University of California group and indicated the existence of a kinematic cutoff typical of a secondary origin for the antiprotons in cosmic rays. This conclusion was subsequently confirmed in measurements by American and Japanese groups in 1992–1994 in the IMAX, BESS, and CAPRICE experiments.[9–11]

Measurements of antiproton fluxes at energies of 2–5 GeV were extended in 1986–1989, when magnetic spectrometers from the IPTI were carried aloft in flights lasting many days off the Kamchatka Peninsula in order to refine the previous measurements. The resulting refined value of the ratio of the antiproton and proton fluxes was $2.4^{+2.4}_{-1.3}\times10^{-4}$ (Ref. 12). In a 1991 balloon flight, a University of New Mexico group obtained a ratio of the antiproton and proton fluxes for energies of 4–19 GeV of $1.24^{+0.68}_{-0.51}\times10^{-4}$,[13] substantially lower than that measured previously by this group.

CONCLUSION

Balloon studies of antiprotons in primary cosmic radiation have yielded a world statistics of roughly 90 antiprotons distributed over energies of 0.1–19 GeV. The energy distribution of the ratios of the antiproton and protons fluxes suggests that, most likely, the antiprotons detected in these experiments originate in the galaxy in the course of nuclear interactions of cosmic rays with the interstellar medium. The results of experiments and calculations for the relative fluxes of antiprotons born in outer space in nuclear interactions during possible evaporation of primary black holes, during annihilation of supersymmetric particles in the dark halo surrounding the galaxy are shown in Fig. 2 The observations of galactic antiprotons can, in particular, yield an experimental limit on the number of evaporations of primary black holes in the cosmos of less than 3×10^{-3} $pc^{-3}yr^{-1}$ and an experimental limit on the mass of the Higgs particles of greater than 50 GeV.

The prospects for searching for antimatter in the universe using antiproton spectrum measurements are limited by a background level of 2×10^{-6} at energies of hundreds of MeV and 2×10^{-4} at energies above a few GeV owing to the birth of antiprotons in nuclear interactions in interstellar space. The possibility of solving the problem by searching for antinuclei is much better, since the creation of antihelium nuclei in nuclear interactions of cosmic rays is expected to occur with a ratio of the fluxes of antihelium-3 and helium of the order of 10^{-12}, and the experimental limits are associated

with the level of separation of the particles in the device and the number of detected events.

The upper limit for the ratio of the fluxes of antihelium and helium nuclei at energies of 0.1–8.6 GeV/nucl obtained in recent years in balloon experiments by American and Japanese groups is 8×10^{-6} (Ref. 19). The minimum measured upper limit for the ratio of the fluxes of antinuclei and nuclei with a charge exceeding 3 is 8×10^{-5} within the energy range 1–15 GeV/nucl.[20] These results indicate that there are no antimatter objects within a radius of the order of 1 Mps. On the whole, there is a restriction on the search for antimatter using cosmic ray observations, probably to a distance of the order of 10 Mps, owing to the escape of cosmic rays from sources, diffusion in magnetic fields, and the penetration of extragalactic cosmic rays into the galaxy. Gamma-ray astronomy permits observation of annihilation processes at great distances, but the results are negative, so far.

Recent U. S. balloon experiments on high energy positrons in cosmic rays yield a ratio of the fluxes of positrons and electrons that is close to that expected from nuclear interactions in interstellar space.[21]

Thus, decades of experimental searches for antimatter using cosmic rays have yielded a negative result up to now. During this period, changes have occurred in our theoretical concepts of the baryon symmetry of the universe. Whereas it was assumed before the Grand Unification theory that the baryon number is exactly conserved in any reactions, according to current ideas there are reactions which violate the conservation of baryon number. These reactions involve superheavy Higgs and gauge particles, so that reactions involving a change in the baryon number can occur effectively only at very high energies. At temperatures above 10^{28} K and times shorter than 10^{-35} s the universe contained an equal number of particles and antiparticles in thermodynamic equilibrium, and the baryon number was equal to zero. A baryon charge develops when the temperature falls below 10^{28} K and the rate of processes with superheavy particles and antiparticles turns out to be slower than the rate of expansion of the universe. In the subsequent decay and annihilation processes up to a time of 10^{-6} s, an excess of baryons develops, and a ratio of the number of protons and photons of 10^{-10}–10^{-9} is maintained thereafter. Thus the universe may be asymmetric in terms of the baryon charge.

Further studies of antiparticles in cosmic rays will require the organization of long (of the order of a month) balloon flights of high throughput magnetic spectrometers and the planning of experiments on spacecraft. In the Russian–Italian space project with the ''Pamela'' magnetic spectrometer, in which the IPTI is participating, plans are to measure antiproton and positron spectra at energies of 0.1–100 GeV beginning in 2001 and search for antinuclei at an antihelium to helium ratio of the order of 10^{-7} (Ref. 22). In the international space station ''Alpha,'' an American–European collaboration is planned for early in the next millennium on an experiment to search for antinuclei with the aid of an even more sensitive AMS magnetic spectrometer.[23]

[1] N. S. Ivanova, D. G. Baranov, and E. A. Yakubovsky, in *Proceedings of the 14th ICRC*, Munich (1975), Vol. 1, pp. 300–304.
[2] B. P. Konstantinov, M. M. Bredov, A. I. Belyaevskiĭ, and I. A. Sokolov, Kosmicheskie Issledovaniya **4**, 66–73 (1966).
[3] B. P. Konstantinov, M. M. Bredov, S. V. Golenetskiĭ, E. P. Mazets *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **33**, 1820 (1969).
[4] E. A. Bogomolov, N. D. Lubyanaya, and V. A. Romanov, in *Proceedings of the 12th ICRC*, Hobart (1971), Vol. 5, pp. 1730–1739.
[5] E. A. Bogomolov *et al.*, in *Proceedings of the 16th ICRC*, Kyoto (1979), Vol. 1, pp. 330–335.
[6] R. L. Golden *et al.*, Phys. Rev. Lett. **43**, 1196 (1979).
[7] A. Buffington *et al.*, Astrophys. J. **248**, 1179 (1981).
[8] E. A. Bogomolov *et al.*, in *Proceedings of the 20th ICRC*, Moscow, (1987), Vol. 2, pp. 72–75.
[9] A. W. Labrador *et al.*, in *Proceedings of the 24th ICRC*, Rome (1995), Vol. 3, pp. 64–67.
[10] S. Orito *et al.*, in *Proceedings of the 24th ICRC*, Rome (1995), Vol. 3, pp. 76–79.
[11] G. Barbiellini *et al.*, in *Proceedings of the 25th ICRC*, Durban, (1997) Vol. 4, pp. 217–220.
[12] E. A. Bogomolov *et al.*, in *Proceedings of the 24th ICRC*, Adelaide (1990), Vol. 3, pp. 288–290.
[13] M. Hof *et al.*, Astrophys. J. **467**, L33 (1996).
[14] R. L. Golden *et al.*, Astrophys. J. Lett. **24**, L75 (1984).
[15] H. Matsunaga *et al.*, Phys. Rev. Lett. **81**, 4052 (1998).
[16] A. W. Labrador and R. A. Mewaldt, Astrophys. J. **480**, 371 (1997).
[17] K. Maki, T. Mitsui, and S. Orito, Phys. Rev. Lett. **76**, 3474 (1996).
[18] T. Mitsui, K. Maki, and S. Orito, Phys. Rev. Lett. **389**, 169 (1996).
[19] J. F. Ormes *et al.*, Astrophys. J. **482**, L187 (1997).
[20] G. F. Smoot, A. Buffington, and C. D. Orth, Phys. Rev. Lett. **35**, 258 (1975).
[21] D. Muller *et al.*, in *Proceedings of the 24th ICRC*, Rome (1995), Vol. 3, pp. 13–16.
[22] S. A. Voronov *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 623 (1999).
[23] S. Ahlen *et al.*, Nucl. Instrum. Methods Phys. Res. A **350**, 351 (1994).

Translated by D. H. McNeill

Production of fullerenes in gas discharge plasmas. I. Kinetics of fullerene formation from polycyclic structures

N. I. Alekseev and G. A. Dyuzhev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted September 3, 1998)
Zh. Tekh. Fiz. **69**, 104–109 (September 1999)

A model is proposed for the formation of fullerenes from three-ring structures. It is shown that under arc discharge conditions, fullerenes are more efficiently formed through the capture of C_2 molecules from the surrounding space by a growing cluster. A system of equations is written down for calculating the kinetics of fullerene structure formation under these assumptions and is then used in some calculations. © *1999 American Institute of Physics.* [S1063-7842(99)02309-0]

When analyzing the process of fullerene synthesis, most researchers view an arc as an efficient source of carbon vapor. The differences in the properties of the charged and uncharged clusters which precede the formation of a fullerene, as well as the reactions specific to an arc plasma, lie beyond the scope of the investigation. The purpose of this cycle of papers is to fill this gap. Naturally, it is necessary to begin with an examination of the kinetics of fullerene formation.

The soundest scheme for fullerene formation is the evolution of carbon along the following path: atoms →molecules→chains→two- and three-ring systems→ fullerenes. This path is based on studies of the mobility and annealing of carbon clusters formed during laser vaporization of graphite.[1] The least comprehensible and most complicated modeling step in this carbon evolution scheme is the last transition from three-ring systems to fullerenes.

A model for this kind of transition has been proposed[2] in a study of the evolution of three-ring clusters (Fig. 1a and b) as a result of a sequence of internal transitions. It was assumed[2] that a cluster develops in such a way that an already assembled segment of the surface has no more than one atom with unsaturated valence (a radical). At the same time, based on model quantum mechanical calculations, it was assumed that the increase in the binding energy of a cluster is small during formation of adjacent pentagons. Under these assumptions, the evolution of a cluster begins with the formation of a compact aggregation of six pentagons. In the course of the evolution, an extended cluster develops which is closed by an analogous aggregate of pentagons.

The scheme proposed in Ref. 2 is difficult to use directly for analyzing fullerene formation in arc discharges for two reasons:

1. The time of the first reaction, which is required to trigger isomerization, is on the order of 10^{-2} s for a structure with a central hexagon (Fig. 1b). Over these times, the gas dynamic stream carries a three-ring cluster into the low-temperature zone and it can no longer undergo isomerization.

2. The final structure of cluster evolution in Ref. 2 is extremely difficult to convert into a fullerene with the aid of a Stone–Wells transformation. Of any set of four adjacent polygons with such a structure, at least one tetra- or heptagon is formed, and to return to a structure of pentagons and hexagons requires new transformations involving large energy expenditures. It is difficult to create these energies in the low-temperature region at the periphery of an arc.

We assume that in an arc discharge, fullerenes are more efficiently formed when a growing cluster captures C_2 and, possibly, C_3 molecules from the surrounding space. This assumption creates two advantages. First, capture of a C_2 molecule by the central region of an initial three-ring system gives an effective push to the subsequent reactions. Second, the capture of even one molecule in the initial stage of assembly ensures more uniform filling of the surface by pentagons and hexagons. This greatly simplifies the annealing of the resulting surface into a fullerene.

The ''openness'' of the system during growth, however, means that it must be described statistically, since clusters with different numbers of polygons and different topologies

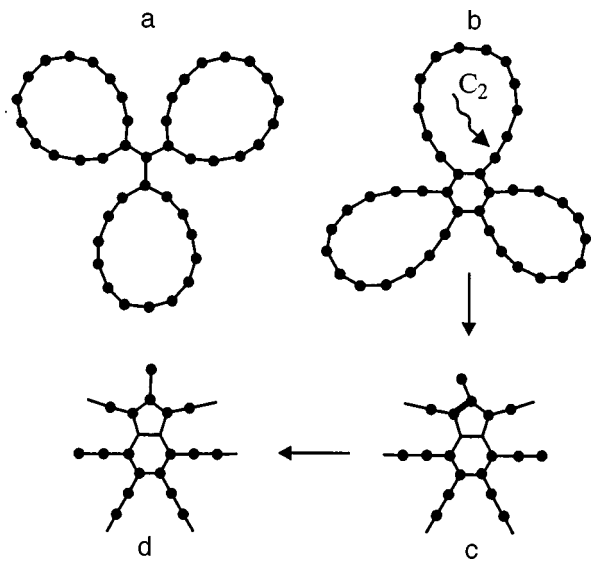

FIG. 1. Initial three-ring clusters in Ref. 2 (a,b) and the capture reaction of a C_2 molecule by the core of a structure (c,d).

coexist at any time. This naturally makes the calculations more complicated.

As in Ref. 2, we have taken a three-ring cluster, shown in Fig. 1b, as an initial structure. Although the scheme described here also applies to two-ring structures, a much more rapid transformation of the cluster is ensured in the three-ring variant and it requires substantially fewer combinations of C_2 molecules.

We have assumed that the most efficient reaction in the first stage of ''construction'' of a fullerene is the combination of a C_2 molecule with the atom of one of the rings nearest to the hexagon. The reaction cross section is $5\sigma_0$, where $\sigma_0 = 5.2\times10^{-16}\exp(-0.5/T)\ \mathrm{cm}^2$ corresponds to the combination of two carbon atoms into a C_2 molecule,[3] and the characteristic time is $\tau_0 = (5\sigma_0 v_T n_2)^{-1} \sim 10^{-4}$ s for a temperature $T = 0.3$ eV and density $n_2 = 10^{14}\ \mathrm{cm}^{-3}$. We estimate that concentrations of this order are entirely typical of the regions where three-ring systems are generated in arcs.

The most probable result of the combination of molecules is the formation of a bond between the first atom of a C_2 molecule and the first atom of a chain, which corresponds to the double line in Fig. 1 (for economy of space, in the figure the chains are indicated by fragments of rays), i.e., the formation of a pentagon and a carbon atom that ''sticks out.'' This is because a radial is not formed in such a reaction, but the reactivity of a chain of two atoms is very high.

It should be noted that, as the surface of the building cluster increases, the reaction rates rise rapidly and the protruding atoms vanish (Fig. 2a–c; the region occupied by the surface prior to initiation of the reaction is shaded). They are replaced by radicals, the number of which also decreases gradually to zero (Fig. 2d).

In the following we shall refer to the reactions among the carbon atoms that already exist in the system as spontaneous and to the reactions stimulated by C_2 molecules as induced. The characteristic times for the spontaneous reactions in the initial stage of isomerization (except for the first, ''slow'' reaction) can be estimated on the basis of Ref. 2. They are roughly $\tau_1 \approx 2\text{–}3\times10^{-5}$ s. Evidently, the time it takes to form a new polygon with the participation of a carbon atom of a connected C_2 molecule is much shorter than τ_1.

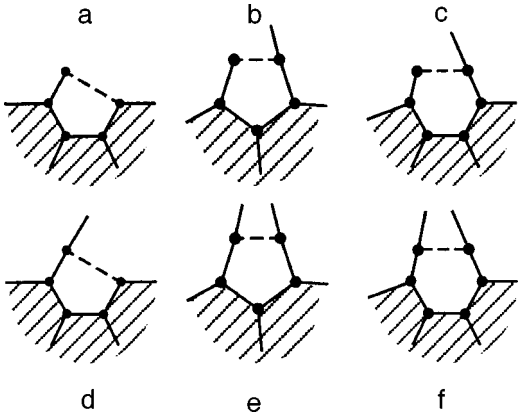

FIG. 2. Reactions leading to loss of protruding atoms and the formation of radicals (a,b,c), loss (d) and reaction without involvement of radicals and protruding atoms (e,f).

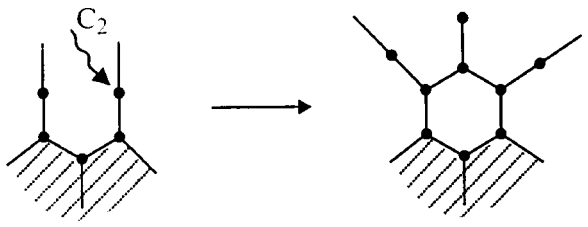

FIG. 3. An induced reaction with formation of a hexagon.

In order to estimate the contribution of induced reactions we must compare the hypothetical times over which a central core of pentagons is formed without the involvement of C_2 molecules[2] ($\approx 5\tau_1$) to the time τ_0.

For a concentration $n_2 = 10^{14}\ \mathrm{cm}^{-3}$ in our model, the core of a cluster can capture 1–2 C_2 molecules (not counting the first collision, which gives a push to isomerization). Here the chain of subsequent reactions keeps the maximum number of pentagons from exceeding three in the region of maximum aggregation.

The transformation of a three-ring cluster into a fullerene is described by a system of balance equations for the concentration Y of clusters with N_5 pentagons, N_6 hexagons, N atoms, μ protruding atoms, and ρ vacant radicals. It was assumed that if the assembled segment of the surface is large enough, then all possible positions of the radicals are occupied by the ends of chains with equal probability. In the initial stage of assembly, the possible reactions and configurations were examined ''manually.''

The equations included induced reactions, forward and reverse spontaneous reactions for creation and loss of polygons, and reactions for ''embedding'' C_2 molecules in chains. The cross sections for these reactions have been calculated[2] and for the typical ring sizes yield a characteristic time of roughly a third of τ_0, so that over the isomerization time, a total of 7–9 C_2 molecules are able to become attached to the cluster. Finally, the possibilities of an atom's breaking away in the final stage of isomerization if the number of atoms is odd and the closure of a cluster into a fullerene for an even number of atoms are taken into account.

Figure 2 shows the six spontaneous reactions that were considered. The reactions induced by the attachment of a C_2 molecule with subsequent formation of a pentagon or hexagon and a projecting atom are shown in Figs. 1 (the transition from Fig. 1b to 1c) and 3.

The system of equations had the following appearance:

$$
\begin{aligned}
dY_{N_5,N_6,\mu,\rho}/dt = {} & (W_{\Delta N_5=+1}Y)_{N_5-1,N_6} - (W_{\Delta N_5=+1}Y)_{N_5,N_6} \\
& + (W_{\Delta N_6=+1}Y)_{N_5,N_6-1} \\
& - (W_{\Delta N_6=+1}Y)_{N_5,N_6} \\
& + (W_{\Delta N_5=-1}Y)_{N_5+1,N_6} \\
& - (W_{\Delta N_5=-1}Y)_{N_5,N_6}
\end{aligned}
$$

$$
\begin{aligned}
&+(W_{\Delta N_6=-1}Y)_{N_5,N_6+1}\\
&-(W_{\Delta N_6=-1}Y)_{N_5,N_6}+n_2\nu_T\exp(-0.5/T)\\
&\times[(\sigma_{\Delta N_5=1,\Delta\mu=+1}Y)_{N_5-1,N_6,\mu-1,\rho}\\
&-(\sigma_{\Delta N_5=1,\Delta\mu=+1}Y)_{N_5,N_6,\mu,\rho}\\
&+(\sigma_{\Delta N_6=1,\Delta\mu=+1}Y)_{N_5,N_6-1,\mu-1,\rho}\\
&-(\sigma_{\Delta N_6=1,\Delta\mu=+1}Y)_{N_5,N_6,\mu,\rho}]\\
&+3n_2\nu_T[\sigma_{2,(N-2-N_s)/3}Y_{N_5,N_6,N-2}\\
&-\sigma_{2,(N-N_s)/3}Y_{N_5,N_6,N}]. \qquad (1)
\end{aligned}
$$

The subscript on the probabilities W of the spontaneous reactions indicates the change in the number of pentagons and hexagons. In fact, each term of the type WY is the sum over all reactions of a given type, during which the numbers μ and ρ change in different ways. For example,

$$
\begin{aligned}
W_{\Delta N_6=+1}Y_{N_5,N_6}\equiv{}&W_{\Delta N_6=+1,\Delta\mu=0,\Delta\rho=0}Y_{N_5,N_6,\mu,\rho}\\
&+W_{\Delta N_6=+1,\Delta\mu=-1,\Delta\rho=+1}Y_{N_5,N_6,\mu,\rho}. \qquad (2)
\end{aligned}
$$

Here the first term corresponds to reactions shown in Fig. 2f and the second, in Fig. 2c. In the system of Eqs. (1), $\sigma_{\Delta N_5=1,\Delta\mu=+1}$ and $\sigma_{\Delta N_6=1,\Delta\mu=+1}$ are the cross sections for formation of a pentagon or hexagon and protruding atom during capture of a C_2 molecule; $\sigma_{2,k}$ in the last term of Eq. (1) is the cross section for capture of a C_2 molecule by a ring of k atoms.

It was assumed that the probability W of spontaneous reactions can be represented as the product of three factors: the first characterizes the assembly state of the system; the second, a combinatorial term, characterizes the geometric probability of configurations of the active elements of the surface such that a given reaction takes place; and the third, the energetics of the reaction. For a reaction involving production of a radical, the third term equals $\exp(-E_r/T)$ (E_r is the ''cost of a radical'') and for a reaction with formation of a pair of adjacent pentagons, $\exp(-E^*/T)$ (E^* is a variable parameter). For example, the probability of the reaction shown in Fig. 2b was written in the form

$$
W=w\cdot D_2(C_{R-2}^{\mu+4}/C_R^{\mu+6})\exp(-2E^*\eta_5/T), \qquad (3)
$$

where $C_n^m=n!/(m!(n-m)!)$; $\eta_5=N_5/(N_5+N_6)$ is the probability of obtaining a pentagon among the polygons adjacent to a newly formed pentagon; $R=\mu+\rho+6$ is the number of possible positions of a radical or protruding atom; and, D_k is the number of segments of the surface between two possible positions of a radical (or protruding atom) separated by k carbon bonds. (Thus, in Fig. 2a, a segment with $k=3$ is indicated.)

The factor w was specified in the form

$$
w=A\exp(-E_a/T), \qquad (4)
$$

where $E_a=E_a(N_5,N_6)$.

The characteristic values of the set D_k with different k were determined by the general properties of a simply connected surface with the given number of pentagons and hexagons. In particular, it is easy to show that, if, for $k>3$, $D_k=0$, then $D_3-D_1=N_5-6$, while the sum $D_1+D_2+D_3$ obviously equals R. Thus, it was sufficient to specify independently only one of the numbers D_1,D_2,D_3, for example D_3. We have assumed that $D_3<3$ and the surface configurations with $D_3=1.2$ are equally probable.

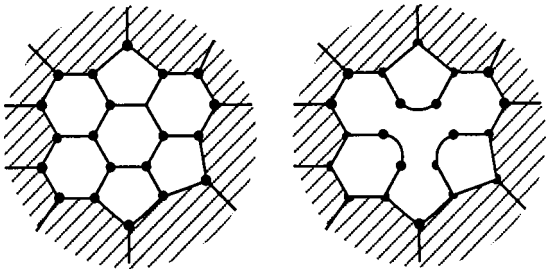

FIG. 4. State of a surface with $\rho=0$ vacant radicals and a minimum ''hole'' within which this number of radicals can be accommodated. The ''hole'' and chain ''residues'' are shown on the right, the ''hole'' in its assembled form, on the left.

The system of Eqs. (1) does not allow us to follow the evolution of the cluster to the end, since each of the reactions considered assumes the appearance of one new bond and one polygon. For the terminal reactions, this is not true. Thus, the calculation was carried out until the size of the unassembled piece of surface (or ''hole'') capable of accommodating a given number of radicals reached a minimum. If there are no radicals ($\rho=0$), this hole consists of a central hexagon and six polygons surrounding it (Fig. 4), i.e., the number of polygons lacking in the fullerene, $\Delta N_\Sigma=12-N_5+N_{6F}-N_6=7$. ($N_{6F}$ is the number of hexagons in a fullerene with the given number of atoms.) From this we find the number N_Σ of polygons in the assembled piece of surface of a cluster which could be referred to as a prefullerene: $N_5+N_6=2+N/2-\Delta N_\Sigma$ subject to the condition $N_5<12$. For $R=7$ ($\rho=1$), $\Delta N_\Sigma=9$. For $\rho=2$, $\Delta N_\Sigma=10$, while for $\rho=3$, $\Delta N_\Sigma=11$.

We shall represent the evolution of a cluster as its movement along a table N_5,N_6, which we shall call an activity table (Tables I and II). For each two dimensional table the numbers N, μ, and ρ are fixed. The entries in the table are to be filled with the values of the corresponding concentrations.

The above discussion implies that we can follow the fate of all clusters located in the activity table that do not lie below the diagonal $N_\Sigma=2+N/2-\Delta N_\Sigma$. To characterize the entries in this part of the table we shall use the following classification:

1. We shall denote the entries lying on the diagonal $N_\Sigma=2+N/2-\Delta N_\Sigma$ by the symbol F. It was assumed that the probability of a transition by clusters corresponding to these entries into a fullerene is determined by the number of pentagon–pentagon boundaries which must be formed,

$$
W_{F\to\mathrm{Full}}=w\exp(-E^*N_\nu\Delta N_5/\Delta N_\Sigma T), \qquad (5)
$$

where $\Delta N_5=12-N_5$, and N_ν is the number of edges in the unassembled segment of the surface.

2. Entries with $N_5+N_6=N_\Sigma$, $N_5<12$, and $N_6\leqslant N_{6F}$ correspond to a cluster in an intermediate stage of development and are denoted in the table by the symbol A (active).

TABLE I.

N_6	N_5										
	1	2	3	4	5	6	7	8	9	10	11
1	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
2	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
3	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
4	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
5	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
6	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
7	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
8	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*	
9	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*		
10	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*			
11	*A*	*A*	*A*	*A*	*A*	*A*	*F*				
12	*A*	*A*	*A*	*A*	*A*	*F*					
13	*S*	*S*	*S*	*S*	*S*						

Note: $N=44$, $\rho=0$.

3. The entries corresponding to a cluster with a deficient number of pentagons and an excess of hexagons are denoted by the symbol *S*. These clusters cannot be closed with the existing number of atoms and exists subsequently as a shell (S).

4. An entry with $N_5=11$ must correspond to an almost closed cluster, which is no longer capable of accommodating the given number of atoms in itself. Such a cluster cannot be realized and its predecessor in terms of the number of pentagons is denoted by *D* (for deadlock). Thus the right-hand boundary of the clusters of type *A* and *F* is arbitrarily taken to be $N_5=10$.

TABLE II.

N_6	N_5										
	1	2	3	4	5	6	7	8	9	10	11
1	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
2	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
3	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
4	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
5	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
6	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
7	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
8	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
9	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
10	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
11	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
12	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
13	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
14	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
15	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	
16	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*	
17	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*		
18	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*F*			
19	*A*	*A*	*A*	*A*	*A*	*A*	*F*				
20	*A*	*A*	*A*	*A*	*A*	*F*					
21	*S*	*S*	*S*	*S*	*S*						

Note: $N=60$, $\rho=0$.

The fate of type *S* and *D* clusters does not interest us, since they cannot be converted into fullerenes. These clusters are the end points of the evolution and determine the boundaries of the activity table. Clusters of type *F* transform to fullerenes with time. Thus, the entire set of objects within the boundaries

$$N_6=N_{6F}+1, \quad N_5=10, \quad N_5+N_6=2+N/2-\Delta N_\Sigma$$

and the fullerenes with a given *N* must ultimately come into equilibrium, when the bulk of the initial product is converted into fullerenes, *F*, *S*, and *D* clusters and there are almost no clusters of type *A* (if the reverse reactions are neglected, there are none of them left at all, or of *F* clusters).

In the calculations, it was assumed that three-ring clusters with a central hexagon and different numbers (from 26 to 74) of atoms in the chains exist at the initial time. The concentration of all the clusters was assumed to be $Y_N|_{t=0}=0.5$. This value was chosen for the linear problem so that, given the transformation of odd clusters into even, the concentration of fullerenes should be 1. The concentration of C_2 molecules and the temperature *T* were assumed to be fixed throughout the entire calculation.

The solution of the first order linear system of equations is the sum of an enormous number of exponential terms, which can be found analytically. Neglecting the reverse cluster decomposition reactions, the change in the concentration of all the clusters can be written in the form

$$\begin{aligned} Y_{N_5,N_6,\mu,\rho,N} = {} & a_{N_5,N_6,\mu,\rho,N}[\exp(-t/\tau_1) \\ & -\exp(-w^d_{N_5,N_6,\mu,\rho,N}t)] \\ & +\sum \alpha_{(N_5,N_6,\mu,\rho,N),(N_5',N_6',\mu',\rho',N')} \\ & \times[\exp(-w^d_{N_5',N_6',\mu',\rho',N'}t) \\ & -\exp(-w^d_{N_5,N_6,\mu,\rho,N}t)], \end{aligned} \tag{6}$$

where $w^d_{N_5,N_6,\mu,\rho,N}$ is the probability of decomposition of the clusters from the given set owing to all of the $N_5'<N_5$ and $N_6'<N_6$ and over all μ' and ρ'.

The coefficients a and α are determined by a chain of recurrence relations. In fact, however, the matrix α occupies an enormous amount of memory and there was not enough memory to include all pairs N_5 and N_6. The concentration of clusters that could not ''fit'' in the analytic solution was calculated by the Runge–Kutta method.

Figure 5 illustrates the formation dynamics of fullerenes with different numbers *N* of atoms. It is clear that the saturation level for these curves is close to 1. For fullerenes with small *N*, this happens because the diagonal $N_5+N_6=2+N/2-\Delta N_\Sigma$ almost completely overlaps the activity table. For the fullerenes with $N=56$–66, the direction of the most rapid development of the clusters also passes through this diagonal and only a small fraction of them settle in entries *S* or *D*. Thus, the saturation level is also high.

The characteristic time it takes to reach saturation is of order 10^{-3} s in the calculations. It has been shown experimentally[4] that the region in which fullerenes are

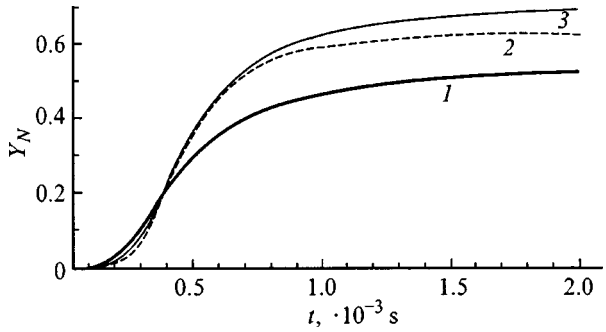

FIG. 5. The rise with time of the concentrations of fullerenes with different numbers of atoms: $N=46$ (*1*), 54 (*2*), 60 (*3*), $T=0.3$ eV; $n_2=10^{14}$ cm^{-3}.

formed in an arc is 2–3 cm long. When the gas dynamic flow velocity out of the discharge is $\sim 2\times 10^3$ cm/s,[5] the time of flight through this region is also $\sim 10^{-3}$ s, in accord with the calculations.

Figure 6 shows spectra of fullerenes in the initial stage of their growth (graph 2) and in saturation (graph 3). The calculated spectra are obviously shifted relative to the initial spectrum of three-ring clusters (a step) by roughly 14–18 atoms (which corresponds to attachment of 2–3 C_2 molecules to the growing core of a surface and 5–7 molecules to rings) and cut off at 68–70 atoms. The number of clusters with high N is limited because, as N increases, the diagonal $N_5+N_6=2+N/2-\Delta N_\Sigma$ in the activity table shifts downward, while the direction in which the clusters shift through the activity table through the most highly populated entries changes little.

If we assume that all the carbon atoms vaporized from the anode of the arc discharge are consumed in forming the initial three-ring clusters, while the end products of the cluster evolution with $N\neq 60$ are transformed to carbon soot, then the data of Fig. 6 imply that the ''percent content of fullerenes in the soot'' is $\sim 9\%$. This value is in good agreement with available experimental data,[5] although this agreement may be purely accidental.

The scheme we have developed can be included in a spatial calculation of the concentrations of the different clusters in a turbulent stream flowing from a discharge gap region. The concentration of the three-ring clusters, which was done ''manually'' in this paper, would be calculated self-consistently in that sort of problem from the initial discharge parameters and the known properties of the arc. The result of such a calculation and a discussion of the reactions specific to a plasma will be published in the later parts of this cycle of papers.

We thank A. A. Bogdanov for a constructive and generous critique of this work and D. V. Afanas'ev for help in preparing the figures.

This study was carried out under the aegis of the Russian Science and Engineering program ''Fullerenes and Atomic Clusters'' (Project No. 98-056) and was supported in part by the firm Fullerene Technologies, St. Petersburg, Russia.

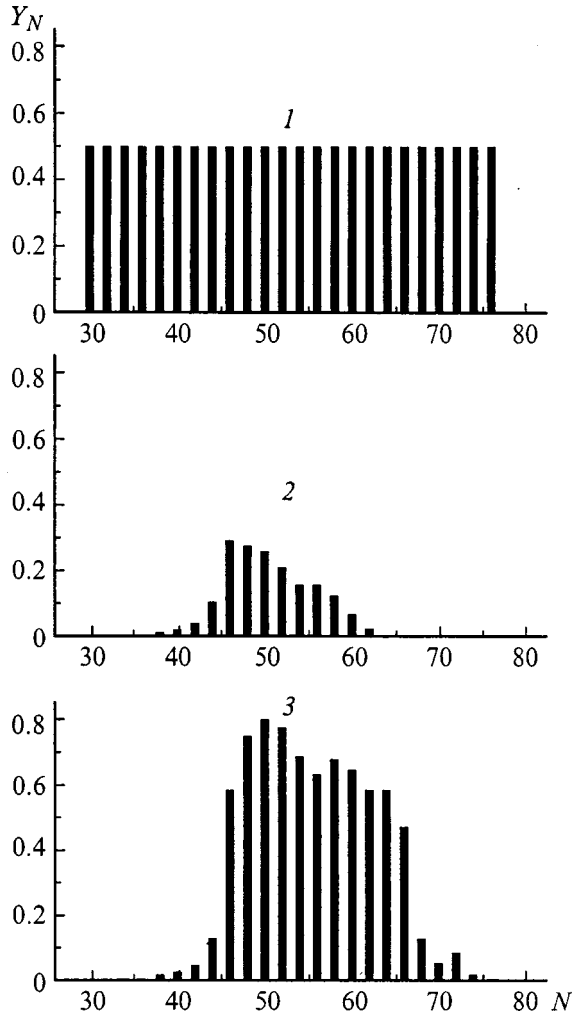

FIG. 6. Initial spectra of three-ring clusters (*1*) and of fullerenes in an early stage of development (*2*) (ten× magnification) and in saturation (*3*). $T=0.3$ eV, $n_2=10^{14}$ cm^{-3}.

[1] J. M. Hunter, J. L. Fye, and M. F. Jarrold, J. Chem. Phys. **3**, 1785 (1993); J. M. Hunter, J. L. Fye, E. J. Roskamp, and M. F. Jarrold, J. Chem. Phys. **98**, 1810 (1994).
[2] V. A. Schweigert, A. L. Alexandrov, Y. N. Morokov, and V. I. Bedanov, Chem. Phys. Lett. **235**, 221 (1995).
[3] V. A. Schweigert and A. L. Alexandrov, Chem. Phys. Lett. **263**, 551 (1996).
[4] G. A. Dyuzhev and V. I. Karataev, Fiz. Tverd. Tela (St. Petersburg) **36**, 2795 (1994) [Phys. Solid State **36**, 1528 (1994).
[5] D. Afanas'ev, I. Blinov, A. Bogdanov *et al.*, Zh. Tekh. Fiz. **64**(10), 76 (1994) [Tech. Phys. **39**, 1017 (1994)]; D. V. Afanas'ev, A. A. Bogdanov, G. A. Dyuzhev *et al.*, Zh. Tekh. Fiz. **67**(2), 125 (1997) [Tech. Phys. **42**, 234 (1994)].

Translated by D. H. McNeill

Crosstalk caused by incoherence of the readout light in wavelength-division multiplexing of reflection holograms

A. V. Chamrai, V. M. Petrov, and M. P. Petrov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted November 25, 1998)
Zh. Tekh. Fiz. **69**, 110–114 (September 1999)

The effect of incoherence of the readout light on the level of crosstalk in wavelength-division multiplexing of reflection holograms is studied. The level of crosstalk is determined theoretically as a function of the spectral width of the readout light, the total number of multiplexed holograms, and the frequency difference between neighboring holograms. Good agreement is obtained between the theory and experiment for the case of two sinusoidal gratings. © *1999 American Institute of Physics.* [S1063-7842(99)02409-5]

The high selectivity of volume holograms makes it possible to record a large amount of information using wavelength-division or angular multiplexing techniques.[1–3] Wavelength-division multiplexing (WDM) of reflection holograms has a number of advantages, primarily a low level of crosstalk (*XSR*). For this case,[4,5]

$$XSR=\left(\frac{K_{g\perp}}{K_g}\right)^2. \tag{1}$$

Here K_g is the wave vector of the recorded hologram and $K_{g\perp}$ characterizes the angular spectrum of the recorded signal. This result holds for the case when both recording and readout are done with perfectly coherent light. In practice, however, this is far from the case. In holographic memory systems[3,6] employing WDM, tunable lasers (for example, semiconductor lasers) with a relatively low level of coherence are used as a readout light source.

The prospects for the practical utilization of WDM for creating various holographic memory systems, as well as in channel switching systems in fiber-optic lines[7,8] make it necessary to analyze in detail the effect of the incoherence of the readout light on the crosstalk level.

In this paper we study the effect of the width of the temporal frequency spectrum of readout light on the level of crosstalk in WDM of reflection holograms. It is shown that incoherence of the readout light greatly increases the level of crosstalk and, therefore, limits the information capacity of systems employing WDM technology.

THEORETICAL ANALYSIS

As a model we consider holographic memory, where the information is recorded in a photorefractive crystal.[9] We shall assume that we have recorded M reflection holograms in ideally coherent light at different temporal frequencies $\nu^{(m)}$. The permittivity variation induced in the crystal as the holograms are recorded can be written in the form

$$\delta\varepsilon\approx\alpha\sum_{m=1}^{M}R_m^*(\mathbf{r})S_m(\mathbf{r})+\text{c.c.}, \tag{2}$$

where S_m and R_m are the complex amplitudes of the signal and reference beams, respectively, and α is a constant describing the properties of the recording medium.

It is assumed that the dynamic range of the optically sensitive medium is sufficient for recording all the holograms in a linear regime.

Suppose we want to reconstruct a certain hologram, number l, from the set of recorded holograms, using light which is not ideally coherent but has a temporal frequency bandwidth $\delta\nu$. Besides the lth hologram in which we are interested, the intensity of the diffracted light will contain contributions from all the other holograms recorded in the crystal. The crosstalk level is characterized by the ratio *XSR* of the intensities of the light diffracted on the ''undesired'' holograms to the intensity of the reconstructed lth hologram:[10]

$$XSR=\frac{I-I^{(l)}}{I^{(l)}}, \tag{3}$$

where I is the total intensity of the reconstructed light and $I^{(l)}$ is the intensity that the reconstructed light would have if the lth hologram alone had been recorded.

When incoherent light is used for readout, the reconstructed light will also be incoherent. Then the intensity distribution of the reconstructed light in the exit plane of the device is given by

$$I=\langle S(r_\perp,t)S^*(r_\perp,t)\rangle. \tag{4}$$

Here $S(r_\perp,t)$ is the amplitude of the reconstructed light wave and $\langle\ldots\rangle$ denotes averaging over all possible realizations, i.e., simultaneously in terms of both the time and the ensemble of recorded holograms. We write the amplitude of the reconstructed light wave in the form of the spectrum of the temporal and angular frequencies,

$$S(r_\perp,t)\sim\int\int\Xi(\sigma_\perp,\delta\nu)\exp(i(\sigma_\perp r_\perp-\delta\nu\cdot t))d\sigma_\perp d\delta\nu. \tag{5}$$

We express the Fourier spectrum of the amplitude of the reconstructed light ($\Xi(\sigma_\perp,\delta\nu)$) in terms of the frequency response function of the recorded holograms (F_{ml}),

$$\Xi(\rho_\perp,\delta\nu)=\sum_m F_{ml}(\rho_\perp,\delta\nu)P_l(\rho_\perp,\delta\nu). \tag{6}$$

Here $P(\rho_\perp,\delta\nu)$ is the spatial–temporal spectrum of the amplitude of the readout light; $\rho_\perp$ is that component of the wave vector of the readout light which describes the propagation direction of the light wave, and it is related to the wave vector of the reconstructed light and the wave vector of the mth hologram by

$$\boldsymbol{\sigma}=\boldsymbol{\rho}-\mathbf{K}_g^{(m)}. \tag{7}$$

The frequency response function of a volume hologram, $F_{ml}(\rho_\perp,\delta\nu)$, for an arbitrary spatial frequency spectrum at a fixed temporal frequency has been derived elsewhere.[11] We shall use the special case of simple sinusoidal gratings,

$$\begin{aligned}
F_{ml}(\rho_\perp,\delta\nu)&\sim A_m(\rho_\perp)\ \exp(i\xi_r^{(ml)}(\rho_\perp,\delta\nu))\\
&\quad\times\sin c(\xi_r^{(ml)}(\rho_\perp,\delta\nu)),\\
\xi_r^{(ml)}(\rho_\perp,\delta\nu)&=\frac{-K_g^{(m)2}+2(\boldsymbol{\rho}\cdot\mathbf{K}_g^{(m)})}{4\rho\sqrt{1-[(\rho_\perp-K_{g\perp}^{(m)})/\rho]^2}}T,\\
\rho&=2\pi n(\nu^{(l)}+\delta\nu)/c,\quad A_m(\rho_\perp)=\delta(\rho_\perp),
\end{aligned} \tag{8}$$

where $\nu^{(l)}$ is the center frequency of the readout light, n is the refractive index of the medium, c is the speed of light in vacuum, and T is the thickness of the crystal in which the recording is done.

$$\mathbf{K}_g^{(m)}=\boldsymbol{\rho}^{(m)}-\boldsymbol{\sigma}^{(m)},\quad \rho^{(m)}=\sigma^{(m)}=\frac{2\pi n\nu^{(m)}}{c}. \tag{9}$$

Here $\boldsymbol{\rho}^{(m)}$ and $\boldsymbol{\sigma}^{(m)}$ are the wave vectors of the coherent writing light. Since the level of crosstalk depends weakly on the angular spectrum of the recorded signal, the results obtained below for simple sinusoidal gratings with different periods can be carried over to the case of recording complicated holograms to within a factor of $(K_{g\perp}^{(m)}/K_g^{(m)})^2$.

Using the approximation of a quasimonochromatic extended source of readout light,[12] we obtain

$$\begin{aligned}
I\sim\int\int\sum_m F_{ml}(\rho_\perp,\delta\nu)P_l(\rho_\perp,\delta\nu)\\
\times\sum_m F_{ml}^*(\rho_\perp,\delta\nu)P_l^*(\rho_\perp,\delta\nu)d\rho_\perp d\delta\nu.
\end{aligned} \tag{10}$$

The integral is taken over the angular aperture and temporal frequency spectrum of the readout light source.

In order to minimize crosstalk the set of sinusoidal gratings is recorded in a crystal in accordance with the Rayleigh criterion,[5,10] that is with a temporal frequency step size such that the detuning parameter ($\xi_r^{(m,m+1)}$) in the Kogelnik formula[13] equals $s\pi$, where s is an integer, and the diffraction efficiency goes to zero; this ensures a minimal level of crosstalk during readout by ideally coherent light. We have defined the the spectral selectivity $\Delta\nu_0$ of the hologram as the frequency detuning of the readout light for which the diffraction efficiency reaches the first zero.

We shall assume that readout is done with a light source that has a rectangular temporal frequency spectrum with a center frequency ($\nu^{(l)}$) equal to the frequency at which the lth hologram was recorded. A very narrow, nearly delta-function angular spectrum of the readout light can be obtained using a collimator,

$$P(\rho_\perp,\delta\nu)=\begin{cases}\delta(\rho_\perp); & -\Delta\leqslant\delta\nu\leqslant\Delta,\\ 0; & \delta\nu\notin[-\Delta;\Delta].\end{cases} \tag{11}$$

We shall assume that the holograms have a random phase in the case of WDM, as is, in fact, true, if special measures are not taken. Then, on averaging over the large number of holograms, the crosstalk adds up incoherently. Substituting Eq. (8) for the frequency response function and the readout light spectrum (11) into Eq. (10), we obtain a final formula for the level of crosstalk owing to the incoherence of the readout light,

$$XSR=\frac{\displaystyle\int_{-\Delta}^{\Delta}\sum_{m\neq l}\sin c^2[2\pi n/c(\Delta\nu_0 s(l-m)+\delta\nu)T(1-\Delta\nu_0 s(l-m)-\delta\nu/\nu^{(l)})]d\delta\nu}{\displaystyle\int_{-\Delta}^{\Delta}\sin c^2[2\pi\delta\nu n/cT(1-\delta\nu/\nu^{(l)})]d\delta\nu}. \tag{12}$$

In deriving Eq. (12), the terms related to the angular spectrum of the recorded holograms ($(K_{g\perp}/K_g)^2$) were dropped because they are small. In fact, for the typical parameter values used in practice ($K_{g\perp}^{(m)}/K_g^{(m)}=5\times10^{-2}$) one can record on the order of 300 pixels/mm, and the correction introduced to *XSR* by the angular spectrum of the holograms is about 10^{-3}.

EXPERIMENT

In order to check these theoretical results, we conducted the following experiment. Two holograms with wave vectors $K_g^{(1)}$ and $K_g^{(2)}$ were recorded in a 0.05 mol. % Fe:$LiNbO_3$ crystal with a thickness $T=5$ mm (Fig. 1). Recording was done with two opposed He–Ne laser beams. The period of the first holographic grating was $\Lambda_1\approx137.61$ nm and that of

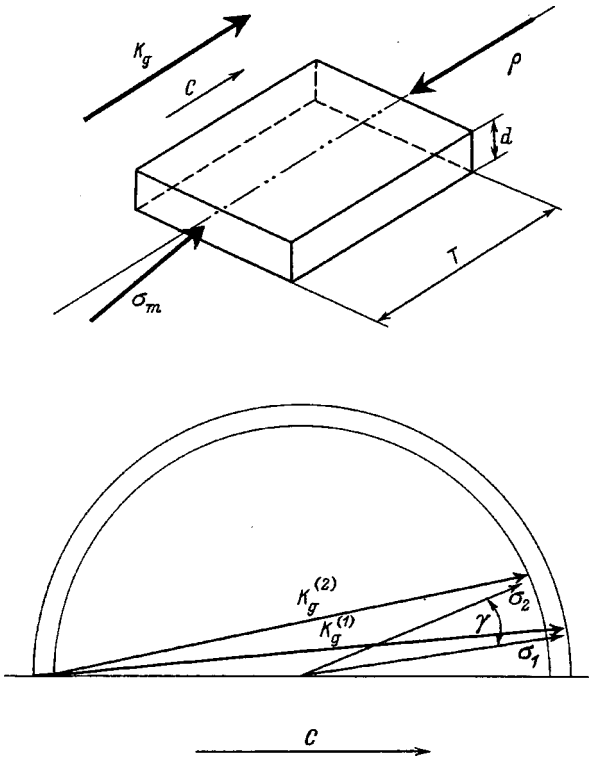

FIG. 1. Geometry for the wavelength-division multiplexing of reflection holograms in the $LiNbO_3$ crystal employed in the experiments.

the second, $\Lambda_2 \approx 138.04$ nm. While the Bragg condition was satisfied exactly ($\xi_r = 0$) for the first hologram with normal incidence of light with a wavelength of 633 nm, $\xi_r \approx 10\pi$ for the second hologram. In order for the Bragg condition to be met for the second hologram with the same incidence of the readout beam, the wavelength must be changed. Thus, these holograms are equivalent to two gratings recorded at different temporal frequencies of the writing light with a step $s \approx 10$.

The erasure time for holograms in doped $LiNBO_3$ crystals is much longer than the recording time. In our case, it was several tens of hours, so a fixing process was not required.

Light from a mercury lamp was used for reconstruction (Fig. 2). The degree of coherence was controlled using a monochromator. The center frequency was set equal to that of the He–Ne laser, $\nu^{(l)} = 4.74 \times 10^{14}$ Hz. When the width of the exit slit of the monochromator was much smaller than the width of the entrance slit, the temporal frequency spectrum of the readout light would be close to rectangular. We could change the temporal bandwidth of the readout light and, thereby, control the degree of coherence, by varying the width of the exit slit.

Another assumption in the theoretical analysis, namely that of a very narrow, nearly delta-function angular spectrum of the readout light, is also satisfied. The width of the angular spectrum can be estimated as $d/F \approx 5 \times 10^{-4}$, where d is the width of the exit slit and F is the focal distance of the collimating objective.

Besides the difference in the periods of the gratings, there was a small angle between the wave vectors of the recorded holograms. Because of this, we separated the light diffracted on the first and second holograms. The angle between the beams reconstructed on the first and second holograms was $\gamma = 12'$, which is smaller than the angular selectivity $\Delta\theta \approx 1°$ (the angle at which the Bragg conditions implies a diffraction efficiency of zero), but this was enough to resolve these two signals in the plane of the photodetector, given the width of the angular spectrum and the divergence induced by diffraction on the finite aperture.

The ratio of the intensity of the light reconstructed on the first hologram ($I^{(1)}$), for which the center frequency of the monochromator is the frequency of the writing beam, to that reconstructed on the second hologram ($I^{(2)}$) is also the level of crosstalk described by Eq. (9).

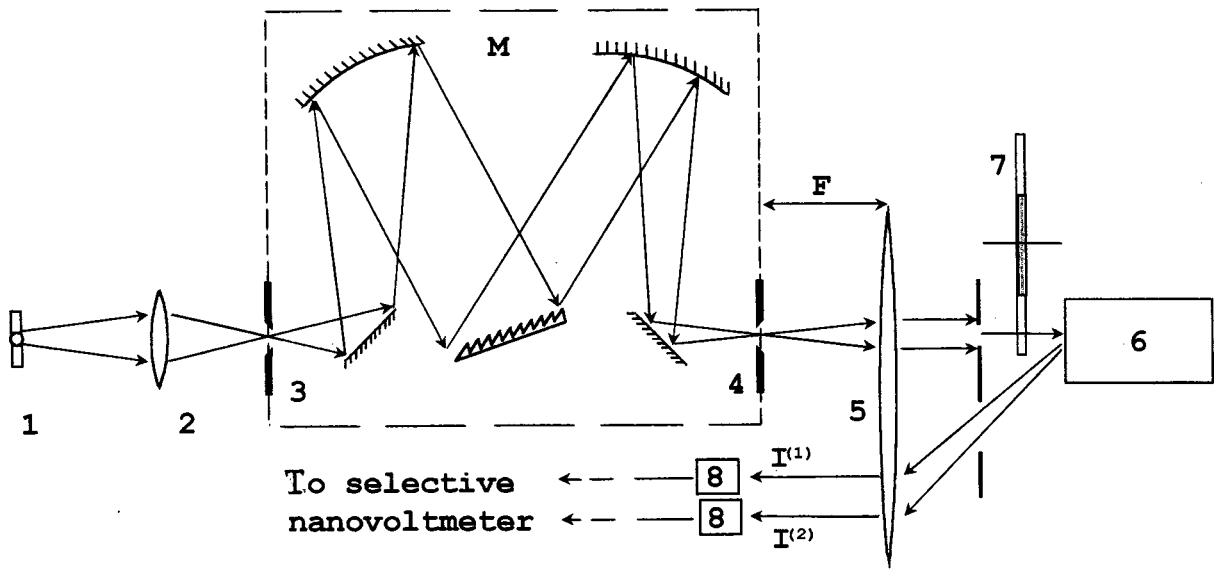

FIG. 2. Experimental apparatus for readout of holograms by partially coherent light: *1* — mercury lamp, *2* — focussing lens, *M* — monochromator, *3* — entrance slit, *4* — exit slit, *5* — collimating objective with a focal length of $F = 350$ mm, *6* — crystal of $Fe:LiNbO_3$, *7* — chopper, *8* — photomultiplier.

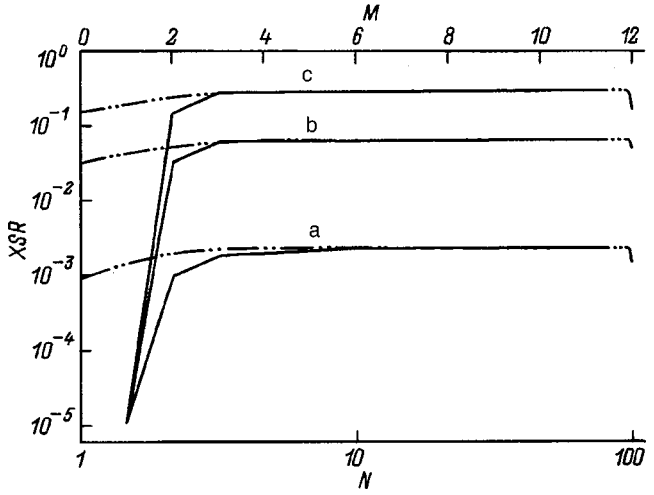

FIG. 3. Theoretical dependence of *XSR* on the total number *M* of holograms (solid curve, $l=M/2$) for $s=1$, $\nu^{(l)}=4.74\times10^{14}$ Hz, $\Delta\nu_0=1.4\times10^{10}$ Hz, as well as on the order number *l* (dot-and-dash curve, $M=100$) of a hologram for different relative widths of the temporal frequency spectrum of the readout light ($\Delta/\Delta\nu_0$): a — 0.3,b — 0.7, c — 1.

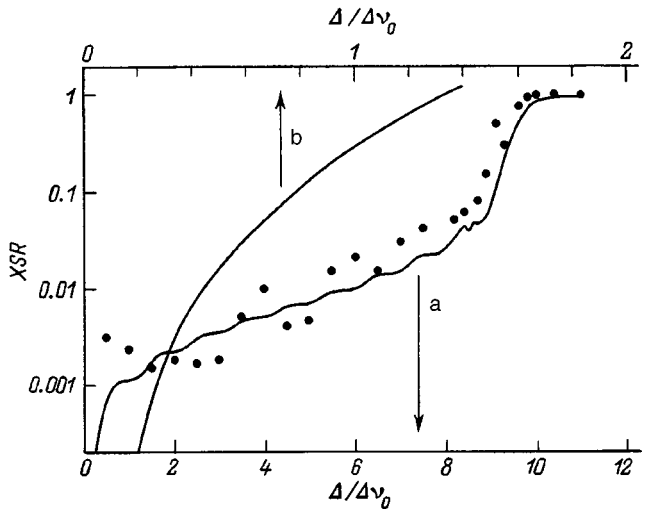

FIG. 4. *XSR* as a function of the relative width of the temporal frequency spectrum of the readout light: a —comparison of theory (solid curve) with experiment (●); b — theoretical dependence for $s=1$, $M=100$.

DISCUSSION OF RESULTS

Quantitative theoretical estimates of the level of crosstalk owing to the incoherence of the readout light have been made for the following system parameters. The holograms were recorded within the specified frequency range with a fixed adjustment step $s=1$. The *l*th hologram was recorded at the frequency of the He–Ne laser. The crystal thickness $T=5$ mm, so $\Delta\nu_0/\nu^{(l)}\approx3\times10^{-5}$. Figure 3 shows plots of *XSR* as a function of the total number of recorded holograms and the order number of the reconstructed hologram for several widths (Δ) of the temporal frequency spectrum of the readout light. It is evident from these graphs that the major contribution is from neighboring holograms and that after the total number of holograms exceeds five, the level of crosstalk no longer varies as the number of holograms is increased. Therefore experiments with just two holograms can provide an estimate of the order of magnitude of the crosstalk. Figure 4 shows a comparison of experiment with the theoretical *XSR* curve as a function of the bandwidth of the readout light calculated using Eq. (12) for the experimental conditions ($M=2$, $s=10$, $T=5$ mm). Also shown there is a theoretical plot of *XSR* as a function of the bandwidth of the readout light for a total number of holograms greater than five with $s=1$.

The level of crosstalk caused by the incoherence of the readout light greatly exceeds the noise associated with the complexity of the recorded signal. Thus, in creating practical systems with the maximum possible information capacity, high standards must be imposed on the coherence of the readout light source.

One possible way of reducing the crosstalk is multiplexing the holograms with a step size exceeding their spectral selectivity, i.e., with $s>1$. However, here the frequency tuning range of the light source must be increased substantially in order to preserve the capacity. Theoretical plots of the crosstalk as a function of the step size in WDM for different bandwidths of the readout light are shown in Fig. 5.

As the step size *s* is increased, the crosstalk decreases, but here it is worth mentioning that the minima in the curve with incoherent readout light are not the same as the conditions for minimum crosstalk with readout by absolutely coherent light (*s* is an integer[10]). In addition, as the temporal frequency spectrum is broadened, the requirements on the accuracy of positioning the center frequency of the readout light become less stringent. On specifying a certain crosstalk level for the system, we can use Eq. (12) to obtain a relationship among the bandwidth of the readout light, the spectral selectivity of the recorded holograms, and the step size in WDM. Thus, in order to obtain $XSR\leqslant0.1$, the following inequality must be satisfied:

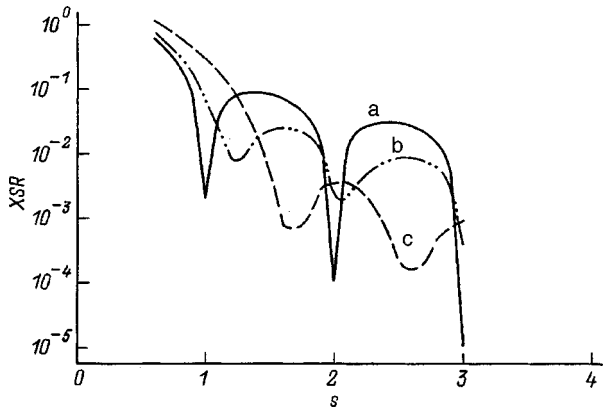

FIG. 5. Theoretical dependence of *XSR* on the step size between multiplexed holograms for different relative widths of the temporal frequency spectrum of the readout light ($\Delta/\Delta\nu_0$): a — 0.3, b — 0.7, c — 1.

$$\frac{s\Delta\nu_0-\Delta}{\Delta\nu_0}\geqslant 0.3. \tag{13}$$

CONCLUSION

We have shown that the level of crosstalk depends significantly on the degree of coherence of the readout light. The crosstalk owing to the nonmonochromaticity of the readout light substantially exceeds that owing to the complexity of the recorded images. Most of the noise is created by the nearest neighbors of a reconstructed hologram, so the level of crosstalk owing to the incoherence of the light depends only weakly on the total number of holograms. *XSR* depends on the relationship among the bandwidth of the readout light, the step size between the holograms, and the spectral selectivity of an individual hologram. A crosstalk level of less than 0.1 can be guaranteed when condition (13) is satisfied.

[1] F. H. Mok, Opt. Lett. **18**, 915 (1993).
[2] S. Yin, H. Zhou, F. Zhao *et al.*, Opt. Commun. **101**, 317 (1993).
[3] L. Hesselink and M. C. Bashaw, Opt. Quantum Electron. **25**, S611 (1993).
[4] C. A. Rakuljic, V. Leyva, and A. Yariv, Opt. Lett. **17**, 1471 (1992).
[5] K. Curtis, C. Gu, and D. Psaltis, Opt. Lett. **18**, 1001 (1993).
[6] J. H. Hong, I. McMichael, T. Y. Chang *et al.*, Opt. Eng. (Bellingham) **34**, 2193 (1995).
[7] R. T. B. James, C. Wah, K. Iizuka, and H. Shimotahira, Appl. Opt. **34**, 8230 (1995).
[8] R. W. Ziolkovski *et al.*, Opt. Lett. **22** 1033 (1997).
[9] M. P. Petrov, S. I. Stepanov, and A. V. Khomenko, *Photorefractive Crystals in Coherent Optics* [in Russian], Nauka, St. Petersburg (1992), 317 pp.
[10] M. C. Bashaw, J. F. Heanue, A. Aharoni *et al.*, J. Opt. Soc. Am. B **11**, 1820 (1994).
[11] Yu. L. Korzinin and V. I. Sukhanov, Opt. Spektrosk. **56**, 763 (1984) [Opt. Spectrosc. **56**, 467 (1984)].
[12] M. Born and E. Wolf, *Principles of Optics* [Pergamon, London (1969); Nauka, Moscow (1973), 719 pp.].
[13] W. H. Kogelnik, Bell Syst. Tech. J. **48**, 2909 (1969).

Translated by D. H. McNeill

New methods for determining the polarization state of vacuum ultraviolet radiation

V. Yu. Bakman, S. V. Bobashev, and O. S. Vasyutinskiĭ

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted March 18, 1998)
Zh. Tekh. Fiz. **69**, 115–122 (September 1999)

Two new methods are proposed for determining the polarization of vacuum ultraviolet radiation which permit the determination of an arbitrary polarization mode for photons with energies of 10–100 eV. The essence of these methods is to create and detect a nonequilibrium population of the magnetic sublevels of atoms and molecules excited by the original VUV radiation and then determine the polarization of this radiation based on these measurements in accordance with known formulas. © *1999 American Institute of Physics.* [S1063-7842(99)02509-X]

INTRODUCTION

Vacuum ultraviolet (VUV) radiation (radiation with wavelengths of 10 to 200 nm) is now used widely in various areas of science and technology.[1] Storage rings and undulators are extremely promising sources of intense VUV radiation[1,2] with a high degree of coherence and an elliptical polarization. Methods for studying the parameters of VUV radiation are of great importance when these sources are used. In addition, it is difficult to determine the polarization state of VUV radiation at wavelengths $\lambda < 105$ nm by standard methods.

The idea of using reflecting polarimeters was first proposed by Hamm and MacRae.[3] Schledermann and Skiblowski[4] made the first attempt to determine the degree of elliptical polarization of synchrotron radiation in the VUV. The proposed method was based on analyzing the intensity of the radiation reflected from a mirror surface with simultaneous measurements of its optical constants. Unfortunately, this attempt was not successful, to a great extent because of the technical complications. In the second half of the 1980's, increasing interest in using the polarization properties of VUV synchrotron radiation, in particular that with circularly polarization, led to the development of new methods for polarization analysis of VUV radiation and stimulated numerous experimental studies of this problem.[5–8]

The use of multilayer reflecting polarimeters has recently made it possible to increase the accuracy of the measurements substantially (to 10%).[9,10] The problem, however, has not yet been solved definitively, since, first of all, this accuracy is not adequate for many applications and, second, the known methods do not provide for nonperturbative monitoring of the polarization of the VUV radiation.

In this paper we examine a new approach to this problem based on some assumptions we have published earlier.[11] Two new methods of determining the polarization of VUV radiation, which presumably have a greater efficiency and accuracy than the standard methods, are studied theoretically. The essence of these methods is to create and detect a nonequilibrium population of the magnetic sublevels of atoms and molecules excited by the initial VUV radiation and then to determine its polarization from these measurements using some known formulas.[11–13]

Two methods for detecting a nonequilibrium population of the excited states are proposed: (1) determining the polarization of infrared, visible, or ultraviolet fluorescence radiation (the fluorescence method[11–13]) and (2) detecting the dichroism of probe laser light (probe beam method[12,13]). Measurements in either method can be used to determine the polarization of the VUV radiation to within about 3%. In addition, the proposed methods allow the measurements to be made more rapidly than before, and also offer the possibility of nonperturbative monitoring of the polarization of VUV radiation.

FLUORESCENCE METHOD FOR DETERMINING THE POLARIZATION OF VUV RADIATION

The fluorescence method for determining the polarization state of VUV radiation essentially involves a preliminary conversion of the frequency of that radiation to a longer wavelength and a determination of the polarization state of the converted radiation, which is then used to calculate the polarization state of the initial VUV radiation using standard formulas. An atomic gas can be used as the frequency converter. In this case, absorption of the VUV radiation causes excitation the gas atoms, whose fluorescence lies in a spectral region that is convenient for study,

$$A + h\nu_{\mathrm{vuv}} \rightarrow A^{**} \rightarrow A^{*} + h\nu_{\mathrm{fl}},$$

where A is an atom or ion, and A^{**} and A^{*} are the states of the atom or ion after absorption of a VUV photon and after reemission of a fluorescence photon, ν_{vuv} and ν_{fl} are the frequencies of the initial VUV and fluorescence radiation, respectively.

Here the polarization of the fluorescent emission from the excited atoms is uniquely determined by the polarization state of the absorbed VUV radiation and the angular momenta of the atomic (ionic) states involved in the photoprocess.[14]

The proposed technique for determining the polarization of the VUV radiation is illustrated in Fig. 1. Note that the absorption chamber containing the atomic or molecular gas

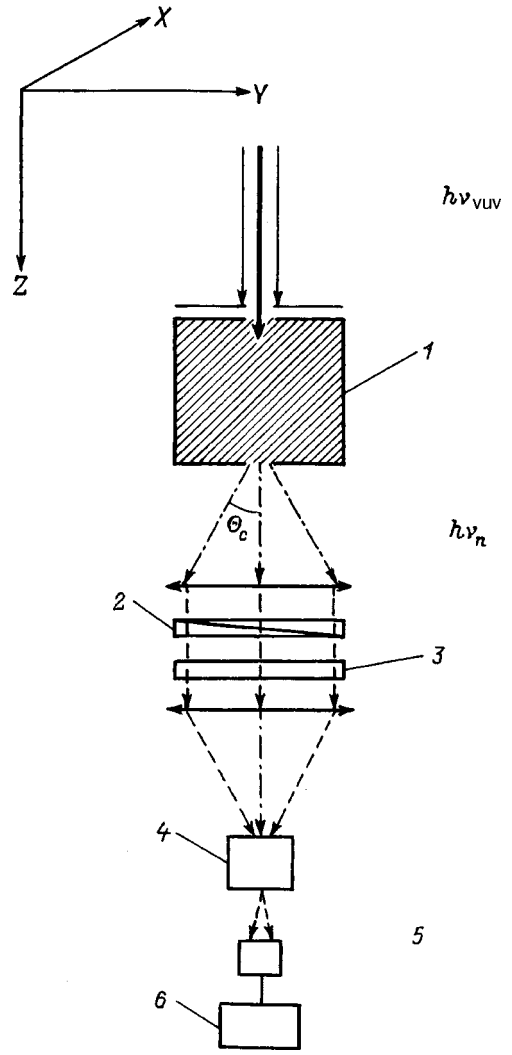

FIG. 1. A sketch of the experimental apparatus for determining the polarization of VUV radiation by the fluorescence method: *1* — absorption chamber with an atomic or molecular gas, *2* — λ/4 plate (only for analyzing circular polarization), *3* — linear polarizer, *4* — monochromator, *5* — photodetector, and *6* — recording system.

must be shielded from external magnetic fields in order to avoid depolarization of the fluorescence owing to the Hanle effect. The detected fluorescence emission is analyzed in the customary way using a λ/4 plate (only for analyzing circular polarization), linear polarizer, monochromator, and photodetector. It is assumed that, in order to enhance the signal-to-noise ratio, the fluorescence radiation is detected within some solid angle Ω with an opening half angle Θ_c. The experimentally determined quantities are the integral Stokes parameters $S_1(\Theta_c)$, $S_2(\Theta_c)$, and $S_3(\Theta_c)$, given by

$$S_1(\Theta_c)=\frac{S'_x(\Theta_c)-S'_y(\Theta_c)}{S'_x(\Theta_c)+S'_y(\Theta_c)}, \tag{1}$$

$$S_2(\Theta_c)=\frac{S'_{x_1}(\Theta_c)-S'_{y_1}(\Theta_c)}{S'_{x_1}(\Theta_c)+S'_{y_1}(\Theta_c)}, \tag{2}$$

$$S_3(\Theta_c)=\frac{S'_+(\Theta_c)-S'_-(\Theta_c)}{S'_+(\Theta_c)+S'_-(\Theta_c)}, \tag{3}$$

where S'_x, S'_y, S'_{x_1}, and S'_{y_1} are the intensities of the fluorescence radiation after transmission through an analyzer turned so that the X, Y, X_1, and Y_1 axes are the transmission axes, respectively, with the X_1 and Y_1 axes lying in one plane with X and Y and forming an angle of 45° with the latter.

As the solid angle of for detection of the radiation goes to zero, $S_1(\Theta_c)$, $S_2(\Theta_c)$, and $S_3(\Theta_c)$ approach the Stokes parameters for the fluorescence radiation propagating along the Z axis: S_1^0, S_2^0, S_3^0.

Let us determine the relationship between the normalized Stokes parameters of the VUV radiation, I_1, I_2, and I_3, and the integral Stokes parameters for the fluorescence radiation S_1, S_2, and S_3. Let the atomic gas used in the absorption chamber be excited by the VUV radiation from its ground state with angular momentum j_1 into state *2* with angular momentum j_2 and the nonequilibrium population of the magnetic sublevels of the excited state be registered by determining the state of polarization of the fluorescence radiation from level *2* to level *3* with angular momentum j_3. We choose the coordinate system so that the Z axis coincides with the propagation direction of the original radiation. We shall describe the propagation direction of the fluorescence light by the polar angles Θ and ϕ.

In a first approximation with respect to the intensity of the exciting radiation, the intensity of the fluorescence radiation with polarization vector $\mathbf{e}_f$ in the $2\rightarrow3$ transition can be written as[15]

$$S(\mathbf{e}_f,\mathbf{e}_i)=S'_0\sum_{mm'\mu} f_{mm'}(\mathbf{e}_i)(\mathbf{e}_f\cdot\mathbf{d}_{m'\mu})(\mathbf{e}_f\cdot\mathbf{d}_{m\mu})^*, \tag{4}$$

where S'_0 is a normalization constant; $\mathbf{e}_i$ is the polarization vector of the exciting radiation; $f_{mm'}$ is the density matrix of the excited atoms, which depends on the polarization of the exciting radiation; and, $d_{m\mu}$ is the matrix element of the dipole moment operator of the atom, where the subscript m enumerates the Zeeman sublevels of the excited state and μ, those for the ground state.

Note that this relation is valid when the exciting radiation is resonant with the $1\rightarrow2$ transition.

In the absence of an external magnetic field, the density matrix of the excited atoms can be written in the form[16]

$$f_{mm'}(\mathbf{e}_i)=F'_0\sum_{\mu}(\mathbf{d}_{m\mu}\cdot\mathbf{e}_i)(\mathbf{d}_{m'\mu}\cdot\mathbf{e}_i)^*, \tag{5}$$

where F'_0 is a normalization constant that is independent of the polarization of the radiation.

In order to find the intensity according to Eq. (4), it is convenient to use an expansion of the density matrix in terms of the irreducible tensor operators[17]

$$f_{mm'}(\mathbf{e}_i)=\sum_{\kappa=0}^{j_2}\sum_{q=-\kappa}^{\kappa}(-1)^q f_q^{\kappa}(\mathbf{e}_i)(T^{\kappa}_{-q})_{mm'}, \tag{6}$$

where

$$(T_q^\kappa)_{mm'}=(-1)^{j_2-m'}\frac{2\kappa+1}{(2j_1+1)^{1/2}}\begin{pmatrix} j_2 & \kappa & j_2 \\ -m & q & m' \end{pmatrix}, \quad (7)$$

and the expressions in brackets are the $3j$ symbols.[18]

The components of the density matrix f_q^κ and the fluorescence intensity S can be written in the form[15]

$$f_q^\kappa(\mathbf{e}_i)=F_0\begin{Bmatrix} 1 & 1 & \kappa \\ j_2 & j_2 & j_1 \end{Bmatrix}\Phi_q^\kappa(\mathbf{e}_i), \quad (8)$$

$$S(\mathbf{e}_f,\mathbf{e}_i)=S_0\sum_{\kappa=0}^{2}(2\kappa+1)\begin{Bmatrix} 1 & 1 & \kappa \\ j_2 & j_2 & j_3 \end{Bmatrix} \times\sum_{q=-\kappa}^{\kappa}(-1)^q f_q^\kappa(\mathbf{e}_i)\Phi_{-q}^\kappa(\mathbf{e}_f), \quad (9)$$

where F_0 and S_0 are normalization constants, the expressions in the curly brackets are the $6j$ symbols,[18] and the functions $\Phi_q^\kappa(\mathbf{e}_i)$ and $\Phi_q^\kappa(\mathbf{e}_f)$ characterize the polarization of the exciting radiation and the fluorescence radiation, respectively, and are given by

$$\Phi_q^\kappa(\mathbf{e})=\sum_{q_1,q_2=-1}^{1}(-1)^{q/2}e_{q1}(e_{q_2})^*\begin{pmatrix} 1 & 1 & \kappa \\ q_1 & -q_2 & -q \end{pmatrix}, \quad (10)$$

where e_q are the circular components of the polarization vector $\mathbf{e}$,

$$e_1=-(e_x+ie_y)\frac{1}{\sqrt{2}},\quad e_0=e_z,\quad e_{-1}=(e_x-ie_y)\frac{1}{\sqrt{2}}. \quad (11)$$

The $6j$ symbols[18]

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix}$$

are nonzero only if the triplets of numbers (j_1,j_2,j_3), (j_1,j_5,j_6), (j_4,j_2,j_6), and (j_4,j_5,j_6) satisfy the triangle condition. Thus, a $6j$ symbol in Eq. (8) or (9) is nonzero only for $\kappa=0,1,2$ and only the following components of the density matrix of the excited atomic state contribute to the intensity of the fluorescence radiation (9):

$$f_0^0,\quad f_q^1\ (q=0,\pm1)\quad\text{and}\quad f_q^2\ (q=0,\pm1,\pm2).$$

As can be seen from Eqs. (9)–(11), for a given polarization vector $\mathbf{e}_f$ of the fluorescence radiation, $S(\mathbf{e}_f,\mathbf{e}_i)/S_0'$ is uniquely determined by the polarization of the exciting radiation, $\mathbf{e}_i$, and by the angular momenta j_1,j_2,j_3 of the atomic states which participate in the photoprocess.

The functions $\Phi_q^\kappa(\mathbf{e})$ characterizing the polarization of radiation propagating in the direction $\mathbf{s}$ with polarization vector $\mathbf{e}$ can be expressed explicitly in terms of the Stokes parameters of this radiation. Let us define the normalized Stokes parameters I_1, I_2, and I_3 of the radiation as

$$I_1=\frac{I_{x'}-I_{y'}}{I_{x'}+I_{y'}},\quad I_2=\frac{I_{x''}-I_{y''}}{I_{x''}+I_{y''}},\quad I_3=\frac{I_+-I_-}{I_++I_-}, \quad (12)$$

where $I_{x'}$, $I_{y'}$, $I_{x''}$ and $I_{y''}$ are the intensities of the light after passing through the polarizer turned so that the X', Y', X'', and Y'' axes are the respective transmission axes.

The coordinate systems (x',y',z') and (x'',y'',z'') are oriented so that the $\mathbf{Z}'$ and $\mathbf{Z}''$ axes coincide with the propagation direction $\mathbf{s}$, the $\mathbf{X}'$ axis lies in a plane with the vectors $\mathbf{Z}$ and $\mathbf{s}$, and the $\mathbf{X}''$ axis is turned at an angle of 45° with respect to the $\mathbf{X}'$ axis; and, I_+ and I_- are the intensities of the light after passing through right- and left-circular analyzers, respectively.

The dependence of the functions Φ_q^κ on the normalized Stokes parameters of the radiation propagating in the direction specified by the polar angles Θ and ϕ can be written, using Eq. (10), in the form

$$\Phi_0^0=-\frac{1}{\sqrt{3}},\quad \Phi_0^1=\frac{I_3}{\sqrt{6}}\cos\Theta,\quad \Phi_1^1=\frac{I_3}{2\sqrt{6}}\sin\Theta e^{i\phi},$$

$$\Phi_0^2=\frac{1}{2\sqrt{30}}(1-3\cos^2\Theta)+\frac{3}{2\sqrt{30}}I_1\sin^2\Theta,$$

$$\Phi_1^2=-\frac{1}{2\sqrt{5}}\sin\Theta\cos\Theta e^{i\phi}-\frac{1}{2\sqrt{5}} \times\sin\Theta(I_1\cos\Theta+iI_2)e^{i\phi},$$

$$\Phi_2^2=-\frac{1}{4\sqrt{5}}\sin^2\Theta e^{2i\phi}+\frac{1}{4\sqrt{5}}(I_1(1+\cos^2\Theta) +2iI_2\cos\Theta)e^{2i\phi}. \quad (13)$$

The functions Φ_{-1}^1, Φ_{-1}^2, and Φ_{-2}^2 are determined from the Φ_1^1, Φ_1^2, and Φ_2^2 given above in accordance with an equation that follows directly from Eq. (10),

$$\Phi_{-q}^\kappa=(-1)^q(\Phi_q^\kappa)^*. \quad (14)$$

Since, according to Eq. (8), the components of the density matrix $f_q^\kappa(\mathbf{e}_i)$ are proportional to the functions $\Phi_q^\kappa(\mathbf{e}_i)$, these functions determine which components of the density matrix of the excited state of the atoms appear when the atoms interact with light incident in a given direction with a given polarization. In their turn, the $\Phi_q^\kappa(\mathbf{e}_f)$ determine just which components of the density matrix of the excited state contribute to the intensity of the fluorescence emission with polarization $\mathbf{e}_f$.[15]

Thus, if the exciting radiation is linearly polarized, then $\Phi_q^1(\mathbf{e}_i)=0$ for any of the allowed values $q=-1,0,1$; at the same time, among the $\Phi_q^2(\mathbf{e}_i)$, some are nonzero (alignment of the atoms). Excitation of the atoms by circularly polarized light leads both to alignment $(\Phi_q^2(\mathbf{e}_i)\neq0)$ and to orientation of the atoms $(\Phi_q^1(\mathbf{e}_i)\neq0)$.[19] In addition, $\Phi_0^0(\mathbf{e}_i)=1/\sqrt{2}$ and is independent of the polarization and propagation direction of the exciting radiation, with $f_0^0\sim\Phi_0^0(\mathbf{e}_i)$, while

$$f_0^0\sim\sum_m f_{mm}=\mathrm{Tr}\ f$$

is the population of level *2*, which, when the atoms are excited from an equilibrium ground state, depends only on the intensity of the exciting radiation and not on its polarization.

We now consider the scheme for determining the polarization of VUV radiation shown in Fig. 1. Let the initial radiation propagate along the $\mathbf{Z}$ axis and the fluorescence

radiation be detected within a solid angle Ω with an opening half angle Θ_c, whose axis is also directed along $\mathbf{Z}$. We shall assume that the surface of the linear or circular analyzer used to determine the polarization of the fluorescence radiation is parallel to the (x,y) plane. If the effects associated with oblique incidence of the fluorescence radiation on a linear analyzer are small,[20] then after it passes through this analyzer, the radiation can be regarded a linearly polarized in a plane parallel to (x,z), regardless of the direction of propagation (i. e., the polarization vector is determined by the intersection of the (x,z) plane and a plane perpendicular to the propagation direction of the light, s).

Substituting Eqs. (8) and (13) in Eq. (9) and taking the integral over the solid angle Ω, we obtain expressions relating the Stokes parameters of the exciting VUV radiation, I_k, to the integral Stokes parameters of the fluorescence radiation, S_k $(k=1,2,3)$,

$$S_1=\frac{3A_2M_2G(\Theta_c)}{2A_0M_0+\frac{1}{2}\cos\Theta_c(1+\cos\Theta_c)A_2M_2}\times(1-\Delta S_1(\Theta_c,\mathbf{e}_i))I_1,$$

$$S_2=\frac{3A_2M_2G(\Theta_c)}{2A_0M_0+\frac{1}{2}\cos\Theta_c(1+\cos\Theta_c)A_2M_2}\times(1-\Delta S_2(\Theta_c,\mathbf{e}_i))I_2,$$

$$S_3=\frac{\frac{3}{2}A_1M_1(1+\cos\Theta_c)}{2A_0M_0+\frac{1}{2}\cos\Theta_c(1+\cos\Theta_c)A_2M_2}\times(1-\Delta S_3(\Theta_c,\mathbf{e}_i))I_3, \tag{15}$$

where

$$G(\Theta_c)=\frac{1}{2\pi(1-\cos\Theta_c)}\int_0^{\Theta_c}\int_0^{2\pi}g(\Theta,\phi)\sin\Theta d\Theta d\phi,$$

$$g(\Theta,\phi)=(\cos^2\Theta\cos^2\gamma-\sin^2\gamma)\cos 2\phi+\cos\Theta\sin 2\gamma\sin 2\phi,$$

$$\gamma\equiv\gamma(\Theta,\phi)=\arctan(\cos\Theta\tan\phi),$$

$$A_\kappa=\begin{Bmatrix}1&1&\kappa\\ j_2&j_2&j_3\end{Bmatrix},\quad M_\kappa=\begin{Bmatrix}1&1&\kappa\\ j_2&j_2&j_1\end{Bmatrix},$$

while the functions $\Delta S_k(\Theta_c,e_i)$ $(k=1,2,3)$ account for errors in the polarization measurements when the analyzer operates in a nonparallel light beam.[20] It can be shown that the ΔS_k as functions of the Stokes parameters I_m have the following form $(m=1,2,3)$:

$$\Delta S_k(\Theta_c,I_1,I_2,I_3)=\frac{\sum_m\alpha_{km}I_m}{\sum_m\alpha'_{km}I_m}. \tag{16}$$

Thus, in general, to find the Stokes parameters I_m of the radiation being studied, it is necessary to solve a system of nonlinear equations; determining the coefficients α_{km} and α'_{km}, which depend on the optical properties of the analyzers being used, requires some rather cumbersome calculations. However, it has been shown[20] that for angles $\Theta_c\leqslant 40°$ the absolute magnitudes of the $\Delta S_{1,2,3}$ are small ($|\Delta S_{1,2}|\leqslant 0.01$, $|\Delta S_3|\leqslant 0.04$), so that errors owing to oblique incidence of the light on the analyzer can be neglected when determining the Stokes parameters I_m with the proposed accuracy.

If the fluorescence emission is detected in the direction defined by the polar angles Θ and ϕ, then, neglecting finite solid angle effects in the detection of this radiation, we obtain

$$S_1(\vartheta,\Omega)=\frac{A_2M_2\left(-\frac{3}{2}\sin^2\vartheta+\frac{3}{2}(1+\cos^2\vartheta)\cdot I_1\right)}{2A_0M_0-A_2M_2\left(\frac{1}{2}(1-3\cos^2\vartheta)+\frac{3}{2}I_1\sin^2\vartheta\right)},$$

$$S_2(\vartheta,\Omega)=\frac{3A_2M_2\cos\vartheta\cdot I_2}{2A_0M_0-A_2M_2\left(\frac{1}{2}(1-3\cos^2\vartheta)+\frac{3}{2}I_1\sin^2\vartheta\right)},$$

$$S_3(\vartheta,\Omega)=\frac{3A_1M_1\cos\vartheta\cdot I_3}{2A_0M_0-A_2M_2\left(\frac{1}{2}(1-3\cos^2\vartheta)+\frac{3}{2}I_1\sin^2\vartheta\right)}. \tag{17}$$

In this approximation, when the propagation direction of the initial VUV radiation and the detection direction for the fluorescence coincide, the dependence of S_k on I_k takes the simpler form[13]

$$S_1=\frac{3A_2M_2}{2A_0M_0+A_2M_2}I_1,$$

$$S_2=\frac{3A_2M_2}{2A_0M_0+A_2M_2}I_2,$$

$$S_3=\frac{3A_1M_1}{2A_0M_0+A_2M_2}I_3. \tag{18}$$

Let us consider an application of this method. The photoexcitation of helium atoms

$$\mathrm{He}(1^1S_0)+h\nu_{\mathrm{vuv}}\rightarrow\mathrm{He}(2^1S_1)\rightarrow\mathrm{He}(2^1S_0)+h\nu_{\mathrm{fl}} \tag{19}$$

can be used to determine the polarization state of VUV radiation at a wavelength of $\lambda_{\mathrm{vuv}}=58.4$ nm. Here the fluorescence should be detected at $\lambda_{\mathrm{fl}}=2058.1$ nm and in this case Eq. (18) reduces to $S_k=I_k$ $(k=1,2,3)$. Other reactions involving He atoms, as well as other atoms and molecules, can be used. In order to operate at shorter VUV wavelengths, photoexcitation reactions of atomic and molecular beams can be used.

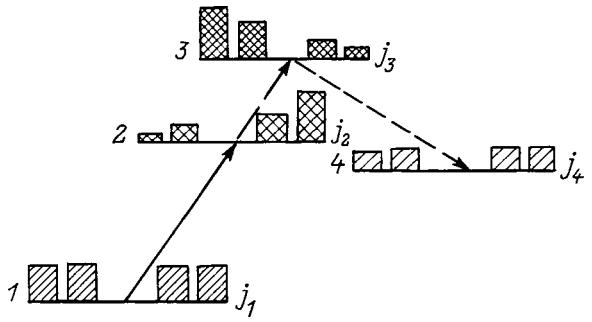

FIG. 2. An atomic level scheme illustrating the probe beam method: *1–2* — transition under the influence of VUV radiation with polarization $\mathbf{e}_i$, *2–3* — transition under the influence of the probe radiation with polarization $\mathbf{e}_p$, *3–4* — detected fluorescence emission.

PROBE BEAM METHOD

This method essentially involves detecting dichroism in the absorption of a probe light beam in an atomic gas that has been irradiated beforehand by the VUV radiation under study; the Stokes parameters of the VUV radiation are determined using standard formulas. The sketch of the energy levels of the atoms in Fig. 2 illustrates this photoprocess. It is assumed that the atoms are initially excited from their ground state with angular momentum j_1 by the VUV radiation into state *2* (the first excited state) with angular momentum j_2, and then the excited atomic gas is probed by a polarized probe beam (such as a laser) with a frequency corresponding to the transition *2→3*. Circular or linear dichroism in the absorption of the probe beam is detected by laser induced fluorescence through the intensity of the fluorescence from level *3* (a second excited state) to level *4*, as a function of the polarization of the probe radiation. The proposed arrangement of the experimental apparatus is shown in Fig. 3.

The advantage of this method compared to the fluorescence method discussed in the previous section is increased sensitivity and, for when a pulsed laser is used as an excitation source, the possibility of eliminating effects owing to the depolarization of state *2* as it is filled by cascading and by trapping of the fluorescence radiation in the volume of the gas.

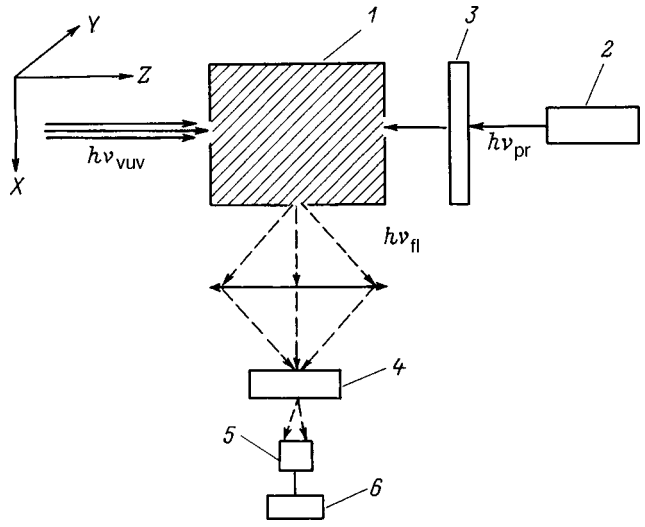

FIG. 3. Sketch of the experimental apparatus for determining the polarization of VUV radiation by the probe beam method: *1* — absorption chamber with atomic gas, *2* — probe radiation source (laser), *3* — polarizer, *4* — monochromator, *5* — photodetector, and *6* — recording system.

Let us consider the interaction of arbitrarily polarized VUV radiation with an atomic gas that leads to dichroism of the gas for probe light that is resonant with one of the atomic lines. The VUV absorption leads to a nonequilibrium population of the Zeeman sublevels in the first excited state of the atoms in the absorption chamber. As noted above, the population of the sublevels depends on the polarization of the exciting radiation,[15,16]

$$f_{m_1m_1'}(\mathbf{e}_i)=F_0'\sum_{\mu}(d_{m_1\mu}\cdot\mathbf{e}_i)(d_{m_1'\mu}\cdot\mathbf{e}_i)^*, \tag{20}$$

where $f_{m_1m_1'}$ is the density matrix of the atoms in the first excited state, F_0' is a constant of proportionality, $d_{m_1\mu}$ is the matrix element of the operator for the dipole moment of the atom, $\mathbf{e}_i$ is the polarization vector of the exciting radiation, and the subscripts m_1 and m_1' enumerate the sublevels of the first excited state and μ those of the ground state.

The population of the sublevels of the second excited state, formed during absorption of the probe light by atoms in state *2*, depends both on the population of the first excited level (i.e., on the polarization of the initial radiation) and on the polarization of the probe light. The matrix elements for the density of atoms in state *3*, $f_{m_2m_2'}$, where the indices m_2 and m_2' enumerate the sublevels of the given state, can be written in the form

$$f_{m_2m_2'}(\mathbf{e}_i,\mathbf{e}_p)=F_1'\sum_{m_1m_1'}f_{m_1m_1'}(\mathbf{e}_i)(\mathbf{d}_{m_2m_1}\cdot\mathbf{e}_p)(\mathbf{d}_{m_2'm_1'}\cdot\mathbf{e}_p)^*, \tag{21}$$

where F_1' is a constant of proportionality and $\mathbf{e}_p$ is the polarization vector of the probe light.

Transforming to a irreversible tensor operator representation, we can write

$$f_{m_1m_1'}(\mathbf{e}_i)=\sum_{\kappa_1=0}^{2j_2}\sum_{q_1=-\kappa_1}^{\kappa_1}(-1)^{q_1}f_{q_1}^{\kappa_1}(\mathbf{e}_i)(T_{-q_1}^{\kappa_1})_{m_1m_1'}, \tag{22}$$

$$f_{m_2m_2'}(\mathbf{e}_i,\mathbf{e}_p)=\sum_{\kappa_2=0}^{2j_3}\sum_{q_2=-\kappa_2}^{\kappa_2}(-1)^{q_2}f_{q_2}^{\kappa_2}(\mathbf{e}_i,\mathbf{e}_p)(T_{-q_2}^{\kappa_2})_{m_2m_2'}, \tag{23}$$

and

$$f_{q_2}^{\kappa_2}(\mathbf{e}_i,\mathbf{e}_p)=(-1)^{q_2}\frac{2j_3+1}{2\kappa+1}\sum_{m_2m_2'}f_{m_2m_2'}(\mathbf{e}_i,\mathbf{e}_p)(T_{-q_2}^{\kappa_2})_{m_2m_2'}, \tag{24}$$

where $f_{q_1}^{\kappa_1}$ are the irreversible components of the density matrix $f_{m_1m_1'}$ and $f_{q_2}^{\kappa_2}$, of the density matrix $f_{m_2m_2'}$.

According to the Wigner–Eckhart theorem,[19] the matrix elements of the circular components q_2 $(q_2=0,\pm1)$ of the

dipole moment operator can be expressed in terms of the product of $3j$ symbols and these matrix elements $(j_3\|d\|j_2)$,

$$(d_{q_2})_{m_2m_1}=(-1)^{j_3-m_2}\begin{pmatrix} j_3 & 1 & j_2 \\ -m_2 & q_2 & m_1 \end{pmatrix}(j_3\|d\|j_2). \qquad (25)$$

Substituting Eqs. (21), (22), and (25) in Eq. (24), after a number of algebraic transformations, we obtain

$$f^{\kappa_2}_{q_2}(\mathbf{e}_i,\mathbf{e}_p)=F_2(-1)^{1+\kappa_2}\sum_{\kappa_1 q_1}(-1)^{q_1}\Pi^2_{\kappa_1}f^{\kappa_1}_{q_1}(\mathbf{e}_i) \times\sum_{x'q'}\Pi^2_{\kappa'}\begin{pmatrix} \kappa_1 & \kappa_2 & \kappa' \\ -q_1 & q_2 & -q' \end{pmatrix} \times B_{\kappa_2\kappa_1\kappa'}\Phi^{\kappa'}_{q'}(\mathbf{e}_p), \qquad (26)$$

where for simplicity, we have used the following notation

$$\Pi_\kappa=\sqrt{2\kappa+1},$$

$$B_{\kappa_2\kappa_1\kappa'}\equiv B_{\kappa_2\kappa_1\kappa'}(j_3,j_2)=\begin{Bmatrix} j_3 & j_2 & 1 \\ j_3 & j_2 & 1 \\ \kappa_2 & \kappa_1 & \kappa' \end{Bmatrix}. \qquad (27)$$

Here the expression in the curly brackets is a $9j$ symbol.[18] Equation (26) relates the density matrix of the atoms in the first and second excited states to the polarization of the probe light. Note that the elements of the density matrix for any state of the atoms subjected to cascade excitation can be calculated by successive use of Eq. (26).

An expression for the intensity of the fluorescence emission at the *3→4* transition can be obtained by inserting the $f^{\kappa_2}_{q_2}$ specified by Eq. (26), where the $f^{\kappa_1}_{q_1}$ are given by Eq. (8), as the components of the density matrix of the upper fluorescence level in the expression for the intensity (4). Using the definition of the irreducible vector product of tensors,[18]

$$\{M_{j_1}\otimes N_{j_2}\}_{jm}=\sum_{m_1m_2}C^{jm}_{j_1m_1j_2m_2}M_{j_1m_1}N_{j_2m_2}, \qquad (28)$$

where M and N are arbitrary irreducible second rank tensors of order j_1 and j_2, respectively, and using the definition of the scalar product,

$$[M'_jN'_j]=\sum_m(-1)^{-m}M'_{jm}N'_{j-m}, \qquad (29)$$

where M' and N' are the irreducible tensors of order j, the expression for the intensity S of the fluorescence emission with polarization specified by the vector $\mathbf{e}_f$ can be written in the form

$$S(\mathbf{e}_i,\mathbf{e}_p,\mathbf{e}_f)=S''_0\sum_{\kappa,\kappa_1,\kappa_2=0}^{2}G_{\kappa\kappa_1\kappa_2}(j_1,j_2,j_3,j_4) \times[\{\Phi_{\kappa_1}(\mathbf{e}_i)\otimes\Phi^*_{\kappa_2}(\mathbf{e}_p)\}_\kappa\Phi_\kappa(\mathbf{e}_f)], \qquad (30)$$

with

$$G_{\kappa\kappa_1\kappa_2}(j_1,j_2,j_3,j_4)=\Pi_{\kappa_1\kappa_2}\begin{Bmatrix} 1 & 1 & \kappa_1 \\ j_2 & j_2 & j_1 \end{Bmatrix} \times\begin{Bmatrix} 1 & 1 & \kappa \\ j_3 & j_3 & j_4 \end{Bmatrix} \times\begin{Bmatrix} j_3 & j_2 & 1 \\ j_3 & j_2 & 1 \\ \kappa & \kappa_1 & \kappa_2 \end{Bmatrix},$$

where S''_0 is a normalization constant.

Using the representation (13) for the functions $\Phi^{\kappa_1}_{q_1}(\mathbf{e}_i)$, $\Phi^{\kappa_2}_{q_2}(\mathbf{e}_p)$, and $\Phi^{\kappa}_{q}(\mathbf{e}_f)$, we can express the intensity of the fluorescence radiation given by Eq. (30) in terms of the Stokes parameters of the original VUV radiation, I_k, and of the probe radiation, P_k $(k=1,2,3)$. As an example, let us consider one of the possible arrangements for this experiment (Fig. 3). Let the VUV radiation propagate along the **Z** axis and the probe light, in the opposite direction, while the fluorescence radiation is detected in a direction corresponding to the **X** axis. After a number of simple, but cumbersome algebraic transformations, we obtain an explicit dependence of the fluorescence intensity S on the Stokes parameters I_k and P_k,

$$S=K\{R_1+R_2P_1+(R_3+R_4P_1)I_1-R_4P_2I_2+R_5P_3I_3\}, \qquad (31)$$

where K is a constant and the coefficients R_1-R_5 are determined by the angular moments j_1-j_4 as follows:

$$R_1=-\frac{1}{3\sqrt{3}}M_0\left(A_0B_{000}+\frac{1}{2\sqrt{5}}A_2B_{022}\right) +\frac{\sqrt{5}}{6\sqrt{6}}M_2\left(\frac{1}{\sqrt{2}}A_2B_{202}+\frac{5}{\sqrt{10}} \times\left(\frac{1}{\sqrt{5}}A_2B_{220}-\frac{1}{\sqrt{7}}B_{222}\right)\right),$$

$$R_2=\frac{\sqrt{5}}{4\sqrt{3}}M_2\left(\frac{\sqrt{5}}{\sqrt{7}}A_2B_{222}+A_0B_{202}\right),$$

$$R_3=\frac{\sqrt{5}}{4\sqrt{3}}M_2\left(\frac{\sqrt{5}}{\sqrt{7}}A_2B_{222}+A_0B_{220}\right),$$

$$R_4=\frac{1}{2\sqrt{3}}A_2\left(\frac{1}{\sqrt{5}}M_0B_{022}+\frac{5}{\sqrt{14}}M_2B_{222}\right),$$

$$R_5=-\frac{1}{2}A_1\left(\frac{1}{3}M_0B_{011}+\frac{\sqrt{3}}{\sqrt{2}}M_2B_{211}\right),$$

$$A_\kappa=\begin{Bmatrix} 1 & 1 & \kappa \\ j_2 & j_2 & j_1 \end{Bmatrix},\quad M_\kappa=\begin{Bmatrix} 1 & 1 & \kappa \\ j_3 & j_3 & j_4 \end{Bmatrix},$$

$$B_{\kappa\kappa_1\kappa_2}=\begin{Bmatrix} j_3 & j_2 & 1 \\ j_3 & j_2 & 1 \\ \kappa & \kappa_1 & \kappa_2 \end{Bmatrix}.$$

As Eqs. (31) imply, four measurements of the intensities S for different combinations of the Stokes parameters P_k of the probe light are required to determine the required Stokes parameters I_k. Then the four unknown quantities K and I_k ($k=1,2,3$) are to be sought as the solution of a system of four equations based on Eq. (31).

Let us consider a possible sequence of measurements. In the first measurement it is necessary to choose probe light that is linearly polarized along the **X** axis. This corresponds to the following normalized Stokes parameters: $P_1=1$, $P_2=0$, and $P_3=0$. In the second measurement, probe light that is linearly polarized along the **Y** axis, with Stokes parameters $P_1=-1$, $P_2=0$, and $P_3=0$, is to be used. In the third measurement, probe light that is linearly polarized at an angle of 45° to the **X** axis with $P_1=0$, $P_2=1$, $P_3=0$ is to be used, and in the fourth, circularly polarized light with Stokes parameters $P_1=0$, $P_2=0$, and $P_3=1$. In each experiment, one measures the fluorescence intensity S_n, where n is the number of the measurement. Equation (31) yields a relationship between the normalized Stokes parameters I_k of the VUV radiation and the experimentally determined intensities S_n ($n=1,2,3,4$; $k=1,2,3$),

$$I_1=\frac{(S_1-S_2)(R_1+R_2)-2R_2}{(S_2-S_1)(R_3+R_4)+2R_4},$$

$$I_2=\left(2\frac{S_1-S_3}{S_1-S_2}-1\right)\left(\frac{R_2}{R_4}+I_1\right),$$

$$I_3=\left(2\frac{S_4-S_1}{S_1-S_2}+1\right)\left(\frac{R_2}{R_4}+I_1\right)\frac{R_4}{R_5}. \quad (32)$$

As an example of determining the polarization state of VUV radiation with a wavelength of $\lambda=58.4$ nm, let us consider the following photoexcitation reactions for helium atoms:

$$\text{He}\ (1^1S_0)+h\nu_{\text{vuv}}\rightarrow\text{He}(2^1P_1),$$

$$\text{He}\ (2^1P_1)+h\nu_{\text{pr}}\rightarrow\text{He}(4^1S_0)\rightarrow\text{He}(3^1P_1)+h\nu_{f1}. \quad (33)$$

Here the wavelength of the probe light is $\lambda_{\text{pr}}=504.8$ nm and that of the fluorescence light is $\lambda_{f1}=2074$ nm. In this case,

$$I_1=\frac{11}{3}\frac{S_1-S_2}{S_1+S_2},$$

$$I_2=\frac{11}{3}\frac{S_1+S_2-2S_3}{S_1+S_2},$$

$$I_3=\frac{11}{5}\frac{S_1+S_2-2S_4}{S_1+S_2}. \quad (34)$$

CONCLUSIONS

Before proceeding to our conclusions, it should be said that the fluorescence method has recently been demonstrated successfully in an experiment[21] to measure the circular polarization of synchrotron radiation in channel (3.3).

In this paper we have made a theoretical study of two new methods for determining the polarization state of VUV radiation. The proposed methods can be used to make polarization measurements of radiation at wavelengths of 10–100 nm and have a number of advantages compared to the existing methods: greater accuracy and speed, and the possibility of nonperturbative monitoring of the polarization of the VUV radiation. These methods essentially involve creating and detecting a nonequilibrium population of the magnetic sublevels of atoms, ions, or molecules that have been excited by the initial VUV radiation and then determining the polarization of the VUV radiation from these measurements using standard formulas.

The fluorescence method assumes detection of a nonequilibrium population in the excited state of the atoms used in the experiment by determining the polarization of fluorescence emission which may lie in the infrared, visible, or ultraviolet. In this paper we have derived equations relating the normalized Stokes parameters of the initial VUV radiation and of the detected fluorescence light for the general case in which the directions of propagation of the initial and detected light differ and the fluorescence emission is collected in a solid angle with an opening half angle Θ_c.

The major shortcomings of the fluorescence method are: 1) it is not possible to measure the polarization of VUV radiation over a wide frequency range, since the frequency of the light under investigation must be resonant with one of the transitions of the atoms, ions, or molecules that are used; thus, in general, different absorbing media must be used to determine the polarization of light at different wavelengths; and, 2) collisional depolarization of the excited state of the atoms that are used, as well as depolarization owing to cascade population of the given state and owing to trapping of the fluorescence radiation in the volume occupied by the gas.

The probe beam method, the second method considered here, provides a radical way of avoiding depolarization of the excited state through collisions, cascading, or radiation trapping of the fluorescence light. This method relies on detecting a nonequilibrium population of the excited state by detecting dichroism in the absorption of the probe laser light from a measurement of the intensity of the fluorescence induced by the probe light. Equations have obtained here for determining the Stokes parameters of the initial VUV radiation from the known Stokes parameters of the probe light and measured intensities of the fluorescence light. Some examples of the proposed techniques were considered.

We thank the authors of Ref. 21 for providing us with a preprint of their work before publication.

This research was supported by a grant from the Russian Fund for Fundamental Research (No. 96-15-96368). V. Yu. Bakman thanks the Soros Foundation for a Student Grant.

[1] A. N. Zaĭdel' and E. Ya. Shreĭder, *Vacuum Ultraviolet Spectroscopy and its Applications* [in Russian], Nauka, Moscow (1976), 456 pp.
[2] H. Winick and S. Doniach, *Synchrotron Radiation Research*, Plenum Press, New York (1980).
[3] R. N. Hamm and R. A. MacRae, J. Opt. Soc. Am. **55**, 1460 (1955).
[4] M. Schledermann and M. Skiblowsk, Appl. Opt. **10**, 321 (1971).
[5] W. B. Westerveld, K. Becker, P. W. Zetner *et al.*, Appl. Opt. **24**, 2256 (1985).
[6] N. Saito, Appl. Opt. **29**, 4538 (1990).
[7] T. Koide, T. Shidara, M. Yuri *et al.*, Phys. Rev. B **44**, 4697 (1991).
[8] R. L. Johnson, J. Barth, M. Cordona *et al.*, Phys. Rev. B **41**, 3291 (1990).
[9] E. S. Gluskin, Rev. Sci. Instrum. **63**, 1523 (1986).
[10] T. Koide, T. Shidara, M. Yuri *et al.*, Rev. Sci. Instrum. **63**, 1458 (1992).

[11] S. V. Bobashev and O. S. Vasyutinskiĭ, Pis'ma Zh. Tekh. Fiz. **11**, 1453 (1985) [Sov. Tech. Phys. Lett. **11**, 599 (1985)].
[12] V. Yu. Bakman, S. V. Bobashev, and O. S. Vasyutinskiĭ, Pis'ma Zh. Tekh. Fiz. **20**(19), 14 (1994) [Tech. Phys. Lett. **20**(10), 771 (1994)].
[13] V. Yu. Backman, S. V. Bobashev, and O. S. Vasyutinskii, Solar Phys. **164**, 397 (1996).
[14] P. P. Feofilov, *Polarized Luminescence of Atoms, Molecules, and Crystals* [in Russian], Fizmatgiz, Moscow (1959).
[15] M. I. D'yakonov, Zh. Éksp. Teor. Fiz. **47**, 2213 (1964) [Sov. Phys. JETP **20**, 1484 (1964)].
[16] O. V. Konstantinov and V. I. Perel', Zh. Éksp. Teor. Fiz. **45**, 279 (1963) [Sov. Phys. JETP **18**, 195 (1964)].
[17] A. R. Edmonds, in *Deformation of Atomic Nuclei* [collected Russian translations, edited by L. A. Sliv], Nauka, Moscow (1958), p. 305.
[18] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* [in Russian], Nauka, Moscow (1975).
[19] A. C. Mitchell and M. W. Zemansky, *Resonance Radiation and Exited Atoms*, Macmillan, New York (1934).
[20] V. Yu. Bakman and O. S. Vasyutinskiĭ, ''A study of the transmission of polarized radiation through an analyzer'' [in Russian], Preprint No. 1670, Physicotechnical Institute, St. Petersburg (1996), pp. 1–32.
[21] C. J. Latimer, M. A. MacDonald, and P. Finetti, in *Proceedings of the 12th International Conference on Vacuum Ultraviolet Radiation Physics*, USA, (1998), p. Tu114.

Translated by D. H. McNeill

First experiments on field photodesorption IR spectroscopy

N. M. Blashenkov and G. Ya. Lavrent'ev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted July 16, 1998)
Zh. Tekh. Fiz. **69**, 123–127 (September 1999)

The ion current of adsorbed water molecules is studied experimentally as a function of the frequency of near-IR radiation incident on a surface at frequencies in the intrinsic vibrational bands of the water molecule. The ions are produced by nonequilibrium field surface ionization. The observed band (near one of the combination frequencies) has a width of 100 cm^{-1} and is shifted relative to the free molecular band by 130 cm^{-1}. Estimates show that the cross section for absorption of the radiation by the adsorbed molecules is 3–4 orders of magnitude larger than for free molecules, as is typical of surface processes. © *1999 American Institute of Physics.* [S1063-7842(99)02609-4]

INTRODUCTION

All the known methods of surface spectroscopy have been classified by R. F. Willis in terms of the physical nature of the exciting (probe) radiation: electrons, photons, neutral particles.[1] A further classification in terms of the physical nature of the system response can be introduced: photons, electrons, neutrals, ions. The method of vibrational surface spectroscopy proposed in this paper uses IR photons as a probe and the response is desorbed ions.

This class (photons, ions) of surface (but not vibrational) spectroscopy includes photoionization microscopy[2] and photodissociative surface ionization.[3] Note that here we are only considering methods with resonant absorption of light. In the first of the above methods, laser light excites electronic transitions in the adsorbed molecules and this leads (as a rule, with multiphoton absorption) to the ionization of molecules or their fragments. In the second method, visible light increases the heterogeneous dissociation of alkali halide salts with the release of alkali metals which ionize at the surface.

FEATURES AND EXPERIMENTAL IMPLEMENTATION OF THE METHOD

In this method, adsorbed molecules are excited by monochromatic IR radiation within their intrinsic vibrational bands. The excited molecules are ionized by nonequilibrium surface ionization, i.e., ionization from vibrationally excited states of the adsorbate.[4] In order to keep these states from relaxing before the ions are desorbed, the lifetime of the adsorbed ions is simultaneously reduced by lowering the Schottky barrier by applying an electric field to the surface.[5] The phenomenon of nonequilibrium surface ionization has been observed[6] and described[4,7] in studies of the ionization of the heterogeneous exothermic decomposition products of acetone peroxide. The measured ion currents of the radicals formed in the reaction exceeded the equilibrium currents calculated according to the theory of surface ionization by 6–7 orders of magnitude. Studies of the energy distribution of these ions have shown that the effective ion temperature T_N of this distribution exceeds the surface temperature T by $\Delta T = T_N - T \approx 500–800$. It has been shown[4,7] that such a large increase in the ion current can be explained in terms of the vibrational excitation of desorbed ions by the energy released during the exothermic chemical reaction.

Thus, the surface ionization efficiency of excited particles should also increase during IR excitation of vibrational states of adsorbed molecules. However, at the low emitter temperatures necessary to form significant coatings, the lifetime of the adsorbate (τ^0 for neutrals and τ^+ for ions) can greatly exceed the relaxation time for vibrational excitation, τ^*. In order for the particles to be desorbed into excited states, it is necessary to reduce their lifetimes on the surface. The lifetimes of neutral particles cannot be changed under these circumstances. In contrast, the lifetime of ions formed from excited molecules by nonequilibrium surface ionization can be shortened, by lowering the Schottky barrier by means of an electric field applied to the surface to make $\tau^+ \leqslant \tau^*$ (Ref. 7).

The novelty of the proposed method, therefore, lies in establishing a relationship between the current of ions desorbed from a surface (in the presence of an strong electric field) and the varying frequency of IR radiation incident on that surface within bands corresponding to eigenmodes of the adsorbed molecules. Since this method is mass spectrometric, observations are made of the current of ions of a certain mass which increases only when the incident radiation excites those vibrations in the adsorbed molecules or their fragments which lead to desorption of ions of these particles. Here the excitation process must involve single photons, in order for the vibrational modes to be resolved. Thus, the recorded spectrum (i.e., the dependence of the ion current on the frequency of the radiation) is the desorption vibrational spectrum of the adsorbed molecules and their fragments.

Figure 1 is a sketch of the apparatus, which consists of a spectroscopic computational system (KSVU-23) and a static sector mass spectrometer with a field surface-ionization ion source. The ion emitter is an oxidized tungsten filament with a diameter of 5 (or 10) μm. A voltage $U = -(0.5–20)$ kV is applied to an electrode that is coaxial with it to produce an

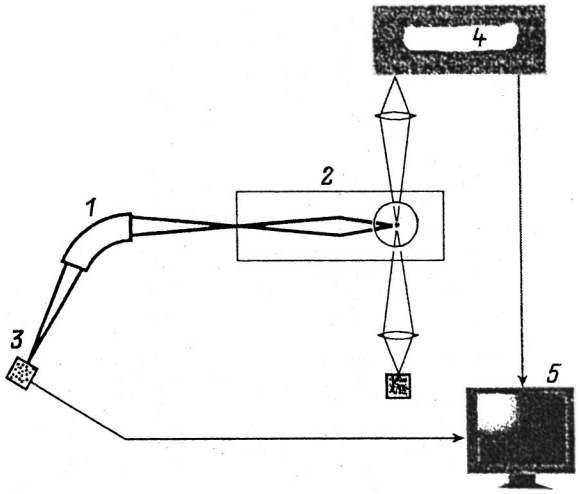

FIG. 1. Block diagram of the apparatus: *1* — mass spectrometer, *2*— field surface-ionization ion source, *3* — secondary electron multiplier, *4* — monochromator, *5* — multipurpose spectroscopic and computational system.

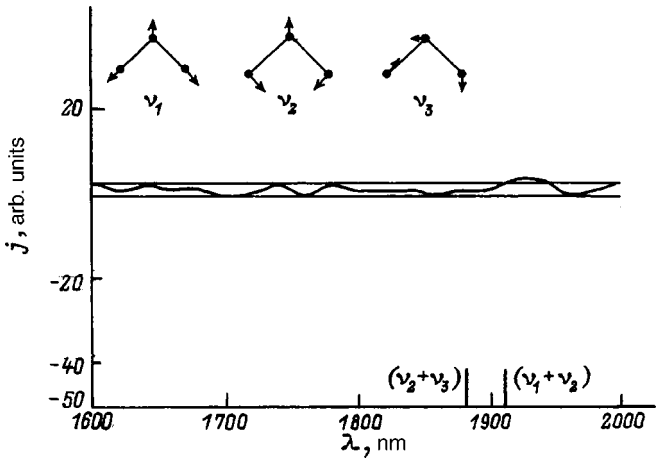

FIG. 2. The current j of desorbed H_2O^+ ions as a function of the wavelength λ of the light incident on a surface: the horizontal lines denote the root-mean-square error; ν_1, ν_2, and ν_3 are the displacements of the atoms in the normal vibrational modes of the water molecule; $(\nu_2+\nu_3)$ and $(\nu_1+\nu_2)$ are the positions of the combination frequencies on the abscissa.

electric field strength at the emitter of from 10^5 to 8×10^6 V/cm with a smooth filament and up to 10^8 V/cm after scales have grown out of the tungsten oxides.[8] The desorbed ions pass through a system of electron lenses which ensure a constant geometry for the ion beam over the entire range of variation in the electric fields and make it possible to analyze the ions in energy by the retarding potential method.[5] The ions are then focussed onto the entrance slit of a mass spectrometer with a resolution of $m/\Delta m=200$. The emitter is illuminated by monochromatic light with a wavelength of from 1.6 to 3 μm (obtained using an MDR-23 monochromator) perpendicular to the axis of the emitter and to the electron-optical axis. The collecting angle for the ion beam is $\pm3°$. The angle at which the IR radiation is absorbed most efficiently by the surface is $\sim88°$,[1] which corresponds to a central angle of $\sim2°$ when the light is incident on a cylindrical surface.

The residual gas pressure in the chamber was found to be 10^{-9} Torr using an additional electron-impact ionization ion source. The pressure of water, whose desorption spectrum we were studying, was found to be $(1-2\times10^{-6}$ Torr. The emitter temperature in these experiments was $\sim$350–400 K. Water ions ($m/e=18$ amu) were detected in a pulse counting mode. An SIRSh-200 lamp was used as a light source. The spectrum was recorded by scanning on the wavelength abscissa over 1600–2000 nm and 2000–3000 nm with a step size of 20 nm (300 lines/mm grating). The ordinate was the signal from the secondary electron multiplier of the mass spectrometer fed the KSVU-23 computer.

A background scan (without illuminating the emitter) was made in each of these spectral ranges and then a scan with the emitter illuminated. The working scan was the difference between the two. 110 and 80 working scans, respectively, were made in these spectral intervals.

EXPERIMENTAL RESULTS AND DISCUSSION

Figure 2 shows the spectrum of the ion current for H_2O^+ ($m/e=18$ u) during a scan over 1.6 to 2 μm. The current exceeded the root-mean-square deviation in the background for $\lambda=1.91$–1.95 μm. The eigenfrequencies of the water molecules nearest these frequencies are the combination frequencies $(\nu_1+\nu_2)^{-1}=1.91$ μm and $(\nu_2+\nu_3)^{-1}=1.88$ μm. (Besides these frequencies there is a fourth harmonic $(4\nu_2)^{-1}=1.63$ μm within this interval, but its excitation probability is very low.) Here we use the following notation for the parallel normal vibrations: skeletal $(\nu_1)^{-1}=2.73$ μm, deformation $(\nu_2)^{-1}=6.27$ μm; and for the perpendicular skeletal vibrations $(\nu_3)^{-1}=2.66$ μm.[9]

Vibrations at the combination frequency $(\nu_2+\nu_3)$, like the fundamental ν_3, are not excited on the surface if the symmetry axis of an adsorbed molecule water lies perpendicular to the surface. (This position of the molecules is most probable, especially in an electric field.) Dipoles positioned parallel to the surface are not excited, since the appropriate component in the radiation incident on the metallic surface is zero.[1] Thus, the observed band can be attributed to the combination vibration frequency $(\nu_1+\nu_2)$. The width of the band in the desorption spectrum is ~100 cm^{-1}, which corresponds to the band width for an isolated molecule. This means that, in our case, the adsorbed molecules are not bound by hydrogen bonds, i. e., they do not form a water film on the surface.[10] The shift in the absorption band by ~130 cm^{-1} toward longer wavelengths indicates the existence of a chemical shift that overlaps with the dipole interaction of the adsorbed molecules.[1]

Some words are in order about the possible appearance of a band of skeletal vibrations ν_1. No statistically reliable spikes in the ion current were observed in the course of an analysis of the spectra over 2–3 μm. Nevertheless, a change in the form of the noise pattern was observed, in particular, a drop by a factor of 2–3 in the number of times the noise curve passes through zero near $\lambda=2.7$ μm (Fig. 3). Given that the band corresponding to the frequency ν_1 is less intense than that at ν_2,[11] we can hope that when the irradiation

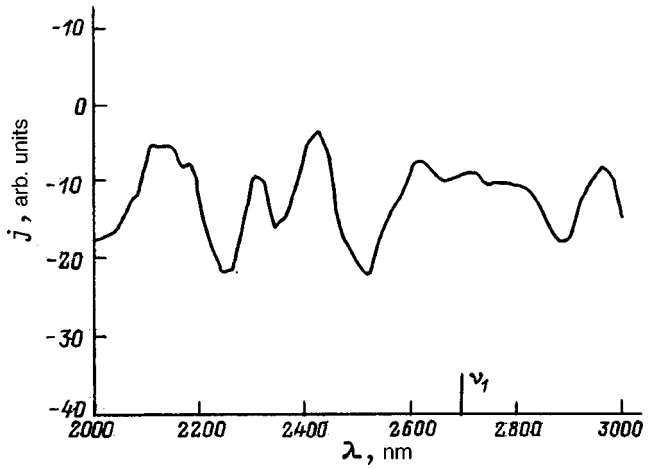

FIG. 3. The current j of desorbed H_2O^+ ions as a function of wavelength λ: the horizontal is the average of the noise path line and ν_1 denotes the location of the skeletal vibration on the abscissa.

intensity is raised, the band at ν_1 will appear in a statistical significant form.

We now derive a formula for the change in the nonequilibrium surface-ionization ion current during absorption of IR radiation by adsorbed molecules for the purpose of estimating the cross section σ for absorption of radiation by adsorbed molecules.

Let us consider the case of hard-to-ionize particles, for which the degree of surface ionization, α, is given by

$$\alpha=\frac{g_+}{g_0}\ll 1.$$

Then the equilibrium surface-ionization current j (the current in scans without illumination of the emitter) is given by[12]

$$j\approx eg\alpha=egA\exp\frac{e(\varphi+\sqrt{eF}-V)}{kT},$$

where k is the Boltzmann constant, g is the flux of particles onto the surface, which for $\alpha\ll 1$ is essentially equal to g_0; g_+ and g_0 are the fluxes of ions and neutrals from the surface (in the steady state, $g=g_0+g_+$), e is the electronic charge, A is the ratio of the statistical weights of the ionic and neutral states of the particles, $e\varphi$ is the work function of the surface, F is the electric field strength at the surface, and V is the ionization potential of the particles.

The flux g determines the concentration N of particles at the surface and is related to N through the lifetimes of the neutral and ionic components,[12]

$$g=N\left(\frac{1}{\tau_+}+\frac{1}{\tau_0}\right).$$

When a strong electric field is present, $\tau_+\ll\tau_0$, so that

$$g=\frac{N}{\tau_+}. \tag{1}$$

When the surface is illuminated, some part of the adsorbed particles enter a vibrationally excited state and are desorbed in the form of ions, whose current will be determined by the extent of nonequilibrium surface ionization (α_N),[7]

$$\Delta j\approx e\Delta g\alpha_N=e\Delta gA_N\exp\frac{e(\varphi+\sqrt{eF}-V)}{kT_N},$$

where A_N is the ratio of the statistical weights of the ionic and neutral states of an excited particle, whose populations are determined by the nonequilibrium temperature T_N.

We note, once again, that T_N is an effective temperature that characterizes the quasi-Maxwellian distribution of the ions with respect to their initial translational velocities and, when the adsorbed particles are vibrationally excited, it exceeds the surface temperature T by an amount ΔT.

The flux of excited particles Δg is proportional to the concentration of particles, N, the radiation intensity I, and the cross section σ for the process, i.e.,

$$\Delta g=NI\sigma,$$

where $I=P/Sh\nu$, $h\nu$ is the photon energy, P is the radiation power, S is the area of the surface illuminated by the light and falling within the aperture of the mass spectrometer.

Then, using Eq. (1), we write the ratio of Δj to j as

$$\frac{\Delta j}{j}=\frac{\Delta g}{g}\times\frac{\alpha_H}{\alpha}=I\sigma\tau^+\exp\left[\frac{e(\varphi+\sqrt{eF}-V)}{kT_N}\left(-\frac{\Delta T}{T}\right)\right]. \tag{2}$$

We leave out the ratio A_N/A, which should enter in Eq. (2), because it is close to unity. In fact, since vibrational levels of the molecules are excited, we only consider the factors in the vibrational statistical weights,

$$\frac{A_N}{A}\approx\frac{\theta_N^+}{\theta_N^0}\times\frac{\theta^0}{\theta^+},$$

where θ^0, θ_N^0 and θ^+, θ_N^+ are the vibrational statistical weights in the ground and excited states for the neutrals and ions, respectively.

At low temperatures, the ratio θ^0/θ_N^0 can be written in the form

$$\frac{\theta^0}{\theta_N^0}=\frac{\prod_{i=1}^{3n-6}\left(1-e^{-\frac{h\nu_j}{kT_N}}\right)}{\prod_{i=1}^{3n-6}\left(1-e^{-\frac{h\nu_j}{kT}}\right)}\approx\frac{\Pi_{i=1}^{3n-6}\frac{h\nu_i}{kT_N}}{\prod_{i=1}^{3n-6}\frac{h\nu_i}{kT}}=\left(\frac{T}{T_N}\right)^{3n-6},$$

where n is the number of atoms in a molecule and h is Planck's constant.

For θ_N^+/θ^+ we obtain the reciprocal of this; thus, $A_N/A\approx 1$. Note also that τ^+ appears in Eq. (2) because, as noted above, the condition $\tau^+\leqslant\tau^*$ applies. Unfortunately, not all the experimentally measured quantities in Eq. (2) were measured in our experiments. Thus, we could not measure ΔT by the lag method because of the small currents Δj. The ionization potential of the water molecules was determined from tables[13] as $eV=12.6$ eV.

The equilibrium current j was used to estimate F ($\sim 10^8$ V/cm); φ was determined from the temperature curves of the bismuth ion current[14] and was found to be

given by $e\varphi = 6.7 \pm 0.05$ eV; the intensity of the radiation ($I \approx 10^{18}$ ph/cm^{2}s) was estimated from the electrical power of the radiator, its temperature with a correction for the black body emission spectrum, and the relative aperture of the monochromator; τ^+ can be estimated based on the following circumstance: the experimental fact that the translational velocity distribution of the excited ions is Maxwellian[6] indicates that during ionization of the excited molecules we are dealing, not with vertical transitions, but with adiabatic ones. In other words, after an IR photon is absorbed with subsequent ionization, the energy absorbed by the particle is randomized over its degrees of freedom, including the translational degrees of freedom. For this, the ions must complete at least a few vibrations. Then τ^+ cannot be less than 10^{-13}–10^{-12} s.

The value of Δj was taken to be the amount by which the current exceeds the root-mean-square deviation from the average current in the working scan. Thus, it was assumed that the spread in the values of the current j corresponds, on the whole, to a normal distribution and that deviations in the current from this distribution at a confidence level of 0.67 are determined by a systematic factor, in our case the vibrational excitation of the ions. The exponent in Eq. (2) was estimated within the limits of variation of $\Delta T \approx 10-50$. Under these assumptions and given the experimental value of $\Delta j/j \approx 10^{-3}$, the cross section for this process calculated using Eq. (2) lies within the range 10^{-12}–10^{-11} cm^2. Such high values of the cross sections σ for various processes involving adsorbed molecules compared to those for gaseous phase processes have been noted in several experimental papers,[3,15] but have not yet been explained theoretically.

CONCLUSIONS

The data obtained using the new method of field photodesorption IR spectroscopy yield the following conclusions: 1) when parallel vibrations are excited in water molecules, they are desorbed with the highest probability from solid surfaces and undergo deformation vibrations; 2) water molecules are adsorbed on a surface by the oxygen atoms; this follows from the closeness of the deformation vibration frequencies of the adsorbed and free molecules; these frequencies are determined by the motion of the hydrogen atoms toward one another, and they remain free when the molecule is adsorbed by its oxygen atom; and, 3) the symmetry axis of the molecule lies perpendicular to the surface; this is implied by the absence of absorption bands with perpendicular vibrations ν_3.

In conclusion, we sincerely thank M. V. Knat'ko, V. I. Paleev, and N. D. Potekhina for helpful discussions.

[1] *Vibrational Spectroscopy of Adsorbates* [in Russian], Mir, Moscow (1984), 246 pp.
[2] V. S. Letokhov, *Laser Photoionization Spectroscopy* [in Russian], Nauka, Moscow (1987), 318 pp.
[3] É. Ya. Zandberg, M. V. Knat'ko, V. I. Paleev *et al.*, Izv. RAN. Ser. Fiz. **56**, No. 8, 21 (1992).
[4] N. M. Blashenkov and G. Ya. Lavrent'ev, Pis'ma Zh. Tekh. Fiz. **14**, 1359 (1988) [Sov. Tech. Phys. Lett. **14**, 593 (1988)].
[5] N. M. Blashenkov and G. Ya. Lavrent'ev, Zh. Tekh. Fiz. **60**(2), 154 (1990) [Sov. Phys. Tech. Phys. **35**, 227 (1990)].
[6] I. N. Bakulina, N. M. Blashenkov, G. Ya. Lavrent'ev *et al.*, Pis'ma Zh. Tekh. Fiz. **1**, 170 (1975) [Sov. Tech. Phys. Lett. **1**, 81 (1975)].
[7] N. M. Blashenkov and G. Ya. Lavrent'ev, Pis'ma Zh. Tekh. Fiz. **21**(24), 15 (1995) [Tech. Phys. Lett. **21**(12), 1000 (1995)].
[8] H. D. Beckey, *Principles of Field Ionization and Field Desorption Mass Spectrometry* (Vol. 61 of International Series in Analytical Chemistry), Pergamon Press, (1977), 335 pp.
[9] L. A. Gribov, *Introduction of Molecular Spectroscopy* [in Russian], Nauka, Moscow (1976), 399 pp.
[10] Ya. S. Bobovich, Usp. Fiz. Nauk **162**(6), 81 (1992) [Sov. Phys. Usp. **35**, 481 (1992)].
[11] G. Herzberg, *Molecular Spectra and Molecular Structure. Vol. 2. Infrared and Raman Spectra of Polyatomic Molecules* [Van Nostrand, New York (1945); Nauka, Moscow (1949), 647 pp.].
[12] É. Ya. Zandberg and N. I. Ionov, *Surface Ionization*, Nauka, Moscow (1969), 432 pp.
[13] *Dissociation Energies of Chemical Bonds. Ionization Potentials and Electron Affinities*, Nauka, Moscow (1974), 351pp.
[14] N. M. Blashenkov and G. Ya. Lavrent'ev, Pis'ma Zh. Tekh. Fiz. **16**(4), 72 (1990) [Sov. Tech. Phys. Lett. **16**(2), 154 (1990)].
[15] St. Y. Dixon-Warren, I. Harrison, K. Reggett *et al.*, J. Chem. Phys. **88**, 4092 (1988).

Translated by D. H. McNeill

Inverse problems in the theory of electrostatic energy analyzers. I

Yu. K. Golikov, K. G. Utkin, and D. V. Grigor'ev

St. Petersburg State Technical University, 195257 St. Petersburg, Russia
(Submitted June 26, 1998)
Zh. Tekh. Fiz. **69**, 128–131 (September 1999)

General formulas are derived for reconstructing two-dimensional Laplace fields from specified focusing and dispersion characteristics in the plane of symmetry. The electrode configuration is determined for an energy analyzer with ideal focusing of a fan beam. The parameters of the plane trajectories, their shape, and their energy dispersion are determined.
 [S1063-7842(99)02709-9]

STATEMENT OF THE PROBLEM

Many papers have been devoted to energy analysis of electron flows in electric fields.[1] As a rule, the choice of electrode configurations is dictated by simplicity of fabrication, symmetry considerations, and engineering estimates of the proposed electron-optical properties based on analogies taken from the literature. In systems found by such a direct approach, it is not usually possible to attain high focal quality and high dispersion simultaneously. For example, a spherical reflector has substantial dispersion, but its focal quality is inadequate. A cylindrical mirror with second order focusing does not provide the required dispersion, while prism spectrometers have a large reserve of dispersion but low throughput, are complicated in design, etc.

For fields with a plane of symmetry, it is possible to propose a more flexible and efficient approach to choosing the field structure, taking as a basis the theory of inverse problems for particle motion, in which the characteristics of the controlled flow are specified and the potential $\varphi(x,y,z)$ of the field is sought. This program for synthesizing energy analyzer systems with previously specified parameters is most easily realized for two-dimensional Laplace potentials with a plane of symmetry along which a fan beam spreads out. These fields include some very convenient variants which combine small size with high dispersion and sharp focusing. In particular, on this path it is easy to find systems with a previously specified order of focusing, and among them, a field with ideal focusing in one plane. An energy analyzer based on this approach, the ''Tutankhamon'' analyzer,[2] has been studied in several papers,[3–5] but a complete theory is given here for the first time. For generality and mathematical clarity, all these investigations are done in dimensionless variables.[6]

DIMENSIONLESS MODEL

We write the physical potential of the field in the form

$$\Phi(X,Y,Z)=\Phi_0\varphi(x,y,z), \tag{1}$$

where Φ_0 is a characteristic value of the potential in ordinary measurement units, such as volts and the factor $\varphi(x,y,z)$ is the dimensionless potential, which depends on the dimensionless coordinates x,y,z, which are related to the real cartesian coordinates X,Y,Z by equations of the form

$$X=lx, \quad Y=ly, \quad Z=lz. \tag{2}$$

The scale length l is usually chosen in accordance with the dimensions of the system. The real time t is conveniently expressed in terms of the dimensionless variable τ by

$$t=T\tau. \tag{3}$$

In the following the derivatives with respect to τ will be denoted by an overhead dot. If the choice of units for the time, T, is subject to the condition

$$T=l\sqrt{\frac{m}{q\Phi_0}}, \tag{4}$$

then the nonrelativistic equations of motion of a particle with mass m and charge q will take the form

$$\ddot{x}=-\partial\varphi(x,y,z)/\partial x, \quad \ddot{y}=-\partial\varphi(x,y,z)/\partial y,$$
$$\ddot{z}=-\partial\varphi(x,y,z)/\partial z. \tag{5}$$

When condition (4) is satisfied, the initial kinetic energy of a particle, $E_0=mv_0^2/2$, is replaced by a dimensionless energy parameter w with the clear physical significance

$$w=(\dot{x}_0^2+\dot{y}_0^2+\dot{z}_0^2)/2=E_0/q\Phi_0. \tag{6}$$

The parameter w expresses the interrelationship between the initial particle energy E_0 and the characteristic potential energy $q\Phi_0$ of the field. Evidently, when $w\gg 1$, the particle moves with a weakly curved trajectory in the field region and for $w\approx 0$ the motion takes place near a line of force of the field that passes through the initial starting point x_0,y_0,z_0, while all the basic transformations of the particles with a small spatial density must take place for $w\approx 1$. The function φ in Eq. (1) is most often chosen subject to the condition $0<|\varphi|<1$.

INVERSION FORMULAS

Let us consider the motion of a particle in the plane of symmetry of a two-dimensional field with a potential $\varphi(x,y)$ distributed along it in accordance with

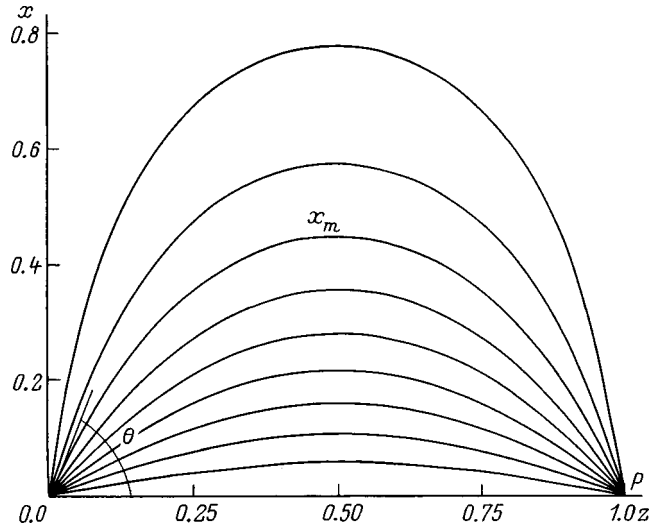

FIG. 1. Particle trajectories in an ideally focusing field.

$$\varphi\Big|_{y=0}=f(x), \quad f(0)=0, \quad \frac{\partial f}{\partial x}\Big|_{x>0}\geqslant 0. \tag{7}$$

The particles start from the point $x_0=y_0=z_0=0$ at an angle θ to the z axis with energy $w=(\dot{x}_0^2+\dot{y}_0^2)/2$, form a symmetric arc under the influence of the field (Fig. 1), and return to the z axis at the point P. In this case, the equations of motion have the form

$$\ddot{x}=-\partial f/\partial x, \quad \ddot{z}=0. \tag{8}$$

For $y_0=0$ and $\dot{y}_0=0$, all the trajectories lie in the plane of symmetry. The energy integral for motion along x is given by

$$\dot{x}^2/2+f(x)=A, \tag{9}$$

where

$$A=\dot{x}_0^2/2=w\sin^2\theta. \tag{10}$$

The time of flight of a particle from the start to point P equals twice the time of flight to the peak, where $\dot{x}=\sqrt{2(A-f)}=0$. We find this time from Eq. (9) by quadrature,[7]

$$\tau_P=\sqrt{2}\int_0^{x_m}\frac{dx}{\sqrt{A-f(x)}}. \tag{11}$$

Along the z axis, there is a drift

$$z=\dot{z}_0\tau=\sqrt{2w}\cos\theta\cdot\tau. \tag{12}$$

In the class of monotonic potentials $f(x)$ the integral (11) is profitably transformed to the variable of integration f by specifying the relation between f and x through the implicit dependence

$$x=F(f), \quad F(0)=0. \tag{13}$$

With the aid of Eqs. (7) and (11)–(13) we calculate the arrival function P of the particle,

$$P=2\sqrt{w}\cos\theta\int_0^A\frac{F'(f)df}{\sqrt{A-f}}. \tag{14}$$

Equation (14) can be used to study the focal quality and the magnitude of the dispersion in any specified field $f(x)$ in a very much more constructive physical sense in another variant. Let the beam be monoenergetic, with $w=w_0=\text{const}$. We express $\cos\theta$ in terms of A using Eq. (10) and rewrite Eq. (14) in the form

$$\frac{P(A)}{2\sqrt{w_0-A}}=\int_0^A\frac{F'(f)df}{\sqrt{A-f}}. \tag{15}$$

This can be regarded as an Abel integral equation in the unknown function $F'(f)$, if the function $P(A)$ is considered given for $w=w_0$. Solving Eq. (15) by a standard method[7] and integrating with respect to f, we arrive at the following key formula in the theory of analyzers:

$$x=F(f)=\frac{1}{2\pi}\int_0^f\frac{P(A)dA}{\sqrt{(w_0-A)(f-A)}}. \tag{16}$$

It can be used to solve various problems concerning the focusing of fan beams in the plane of symmetry of a field, uniquely determining the structure of the potential in the plane of symmetry from a specified angular dependence of the arrival function $P(A)$, where $A=w_0\sin^2\theta$ serves as the angular variable in this case.

Let us assume, further, that the angle θ_0 stands out in the flow and we consider the variation of the arrival function P with the energy w and, thereby, study the dispersion properties of the field. It is convenient, as before, to retain the parameter $A=w\sin^2\theta_0$ in the equations as a basic parameter, assuming that it varies as a result of the energy w. Then Eq. (14) should bee transformed to a new form, on expressing w in terms of A for $\theta=\theta_0=\text{const}$,

$$\frac{\tan\theta_0 P(A)}{2\sqrt{A}}=\int_0^A\frac{F'(f)df}{\sqrt{A-f}}. \tag{17}$$

Solving this equation for F' and integrating with respect to f subject to the condition $F(0)=0$, we obtain the second key inversion formula in the theory of analyzers,

$$x=F(f)=\frac{\tan\theta_0}{2\pi}\int_0^f\frac{P(A)dA}{\sqrt{A(f-A)}}. \tag{18}$$

Now we have the possibility of uniquely determining a two-dimensional field from a specified dependence of the arrival function P on the energy w through the parameter $A=w\sin^2\theta_0$, which is now an energy parameter, in other words, actually from the magnitude of the dispersion:

$$D=w\frac{\partial P}{\partial w}=A\frac{\partial P}{\partial A}. \tag{19}$$

If we intend to specify the variation in the dispersion $D(A)$ for a fixed entrance angle θ_0, then it is first necessary to recover the function $P(A)$ from Eq. (19) by integrating,

$$P(A)=\int\frac{D(A)dA}{A}. \tag{20}$$

In both cases, the spatial distribution of the potential is easy to find explicitly using the complex potential $\Omega=\psi(x,y)+i\varphi(x,y)$ in accordance with the obvious formula

$$\Omega=if(x+iy) \tag{21}$$

or in the implicit notation

$$x+iy=F(-i\Omega)=F(\varphi-i\psi). \tag{22}$$

From Eq. (22) we obtain a convenient real representation of the equipotentials in the parametric form

$$x=F_1(\psi,\varphi), \quad y=F_2(\psi,\varphi), \tag{23}$$

where for fixed $\varphi=\text{const}$ it is necessary to change ψ as a parameter.

FIELD WITH IDEAL FOCUSING

Let us set $w_0=1$ in Eq. (16), which in no way reduces the generality, and choose P as a constant, say $P=1$. Then., using Eq. (16), we quickly find the variation in the potential which ensures complete independence of the arrival function P from the angle θ, i.e., ideal focusing of a monoenergetic flow $w=1$. Integrating Eq. (16) yields the expression

$$x=F(f)=\frac{1}{2\pi}\ln\frac{1+\sqrt{f}}{1-\sqrt{f}} \tag{24}$$

and, therefore,

$$f=\tanh^2\pi x. \tag{25}$$

Equation (25) is contained, in somewhat hidden form, as a special case in the most interesting papers of Geronimus,[8,9] which have not had a proper influence on electron optics, apparently because of an excessively abstract initial premise. Here we associate Eq. (25) with the variation of the potential in a real two-dimensional Laplace field, in its midplane, and the entire construction acquires a clear and distinct physical significance.

Separating the imaginary part from the complex potential Ω,

$$\Omega=\psi+i\varphi=i\tanh^2\pi(x+iy), \tag{26}$$

we obtain the ordinary scalar potential

$$\varphi(x,y)=\frac{\sinh^2 2\pi x-\sin^2 2\pi y}{(\cos 2\pi x+\cos 2\pi y)^2}. \tag{27}$$

Its equipotential portrait is extremely complicated (Fig. 2) and can be used extensively in beam optics, in particular, where there is a convenient closed fragment (Figs. 2 and 3), as the basis for an energy analyzer.

PARTICLE DYNAMICS IN A PLANE OF SYMMETRY

With the aid of the integral (9),for the potential (25) it is easy to find a trajectory shape and arrival function $P(w,\theta)$ with arbitrary energy w, but subject to the condition

$$A=w\sin^2\theta<1. \tag{28}$$

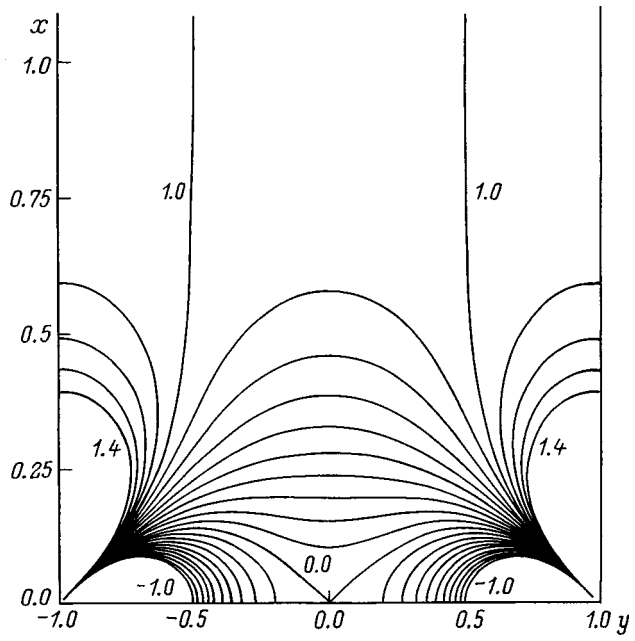

FIG. 2. Equipotentials of a field with ideal focusing.

Then

$$x=\frac{1}{\pi}\ln\frac{\sqrt{2-A(1+\cos kz)}+\sqrt{A(1-\cos kz)}}{\sqrt{2(1-A)}}, \tag{29}$$

where

$$k=2\pi\sqrt{\frac{1-A}{w-A}}.$$

The structure of a focussed flux ($w=1$) is shown in Fig. 1. Equation (14) yields the arrival function $P(w,\theta)$,

$$P(w,\theta)=\cos\theta\sqrt{\frac{w}{1-w\sin^2\theta}}. \tag{30}$$

The corresponding linear energy dispersion along the z axis is

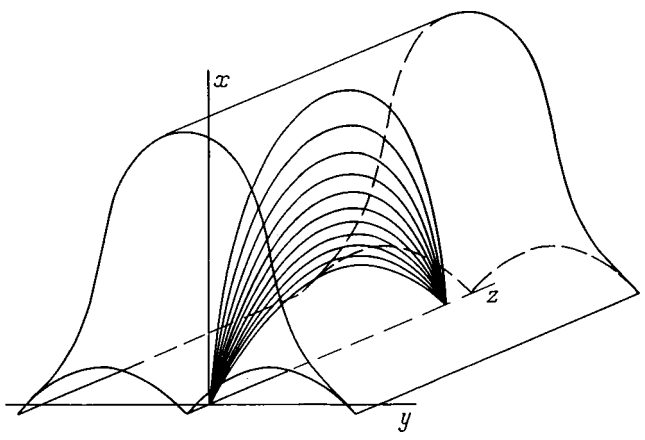

FIG. 3. Possible design of an electrostatic energy analyzer with ideal focusing in one plane (xz).

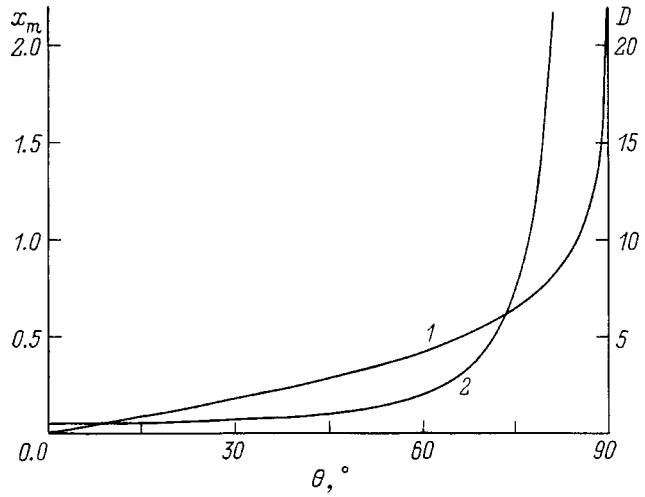

FIG. 4. The height of rise of the trajectories (*1*) and the energy dispersion (*2*) as functions of the entrance angle in the field, θ.

$$D = w\frac{\partial P}{\partial w} = \frac{\cos\theta}{2}\sqrt{\frac{w}{(1-w\sin^2\theta)^3}}. \tag{31}$$

Under ideal focusing conditions with $w=1$,

$$D = \frac{1}{2\cos^2\theta}. \tag{32}$$

The maximum swing of the trajectory along the x axis is given by a formula for the coordinate of the peak x_m (Fig. 1),

$$x_m = \frac{1}{2\pi}\ln\frac{1+\sin\theta}{1-\sin\theta}. \tag{33}$$

The curves $x_m(\theta)$ and $D(\theta)$ (Fig. 4) manifest a remarkable property of this field, which is optimal for an energy analyzer, specifically a combination of small trajectory dimensions with a high level of dispersion, and all this occurs under ideal focusing conditions in at least one plane. For example, for angles $78<\theta<88°$, the dispersion varies within $11.56<D<410$, so that the minimum dispersion is already an order of magnitude higher than for plane and cylindrical mirrors of similar size. The height of the peak, $x_m<1.29$. A real device with dimensions of $100\times100\times130$ mm with a 0.1 mm exit slit can have a resolution $\Delta E/E=0.01\%$.

THE CONFIGURATION OF THE ''TUTANKHAMON'' ENERGY ANALYZER

From the overall equipotential portrait (Fig. 2) it is possible to cut out a fragment bounded by the equipotentials $\varphi=0$ and, for example, $\varphi=0.95$. In space this forms an elongated box, whose elegant outline resembles an Egyptian sarcophagus, which led to the idea of calling this analyzer ''Tutankhamon'' (Fig. 3).

The great convenience of this configuration is that the entrance and exit slits lie on the z axis in the recess of the grounded electrode, where there is almost no field and the apertures need not be covered with a grid.

The other electron-optical characteristics will be described in part II of this paper.

[1] V. P. Afanas'ev and S. Ya. Yavor, *Electrostatic Energy Analyzers for Charged Particle Beams*, Nauka, Moscow (1978), p. 224.
[2] Yu. K. Golikov, *Abstracts of the Second All-Union Conference on Mass Spectrometry* [in Russian], Moscow (1974), p. 324.
[3] L. N. Gall', R. N. Gall', Yu. K. Golikov, and K. G. Utkin, Inventor's Certificate No. 544307 (1976), BI, No. 3 (1977).
[4] Yu. K. Golikov, S. N. Davydov, V. V. Korablev, and K. G. Utkin, Inventor's Certificate (USSR) No. 1365987 (1987); pub. Byull. Izobret., No. 1 (1988).
[5] S. N. Davydov, M. Sh. Sugaitov, V. V. Korablev, and Yu. A. Kudinov, *Abstracts of the Seventh Symposium on Secondary Electron and Photoelectron Emission and Spectroscopy of Solid Surfaces* [in Russian], Tashkent (1990), pp. 14–15.
[6] Yu. K. Golikov, K. G. Utkin, and V. V. Cheparukhin, ''Design calculations for elements of electrostatic electron-optical systems'' [in Russian], LPI im. M. I. Kalinina, Leningrad (1984), 80 pp.
[7] L. D. Landau and E. M. Lifshitz, *Mechanics* [in Russian], IFML, Moscow (1958), 206 pp.
[8] Ya. L. Geronimus, Zh. Tekh. Fiz. **32**, 3 (1962) [Sov. Phys. Tech. Phys. **7**, 1 (1962)].
[9] Ya. L. Geronimus, Zh. Tekh. Fiz. **32**, 848 (1962) [Sov. Phys. Tech. Phys. **7**, 620 (1963)].

Translated by D. H. McNeill

Production of thick $YBa_2Cu_3O_{7-\delta}$ films on sapphire with a cerium oxide sublayer

E. K. Hollmann, D. A. Plotkin, S. V. Razumov, and A. V. Tumarkin

St. Petersburg State Electrotechnical University, 197376 St. Petersburg, Russia
(Submitted April 28, 1998)
Zh. Tekh. Fiz. **69**, 132–136 (September 1999)

The feasibility of producing high-quality films of the high-temperature superconducting material $YBa_2Cu_3O_{7-\delta}$ with thicknesses up to 2.6 μm by dc magnetron sputtering is demonstrated. It is found that inclusions consisting of CuO and $YBa_2Cu_3O_8$ coexist with the growing film and are''sinks'' for defects, nonstoichiometric atoms, and mechanical stresses. X-ray diffraction and Rutherford backscattering studies show that the structural perfection of the films increases as the thickness is increased in the proposed production regime. © *1999 American Institute of Physics.* [S1063-7842(99)02809-3]

INTRODUCTION

YBCO films with thicknesses of 0.6–2.6 μm offer great promise as a material for creating microwave and high-power electronic devices. Microwave applications require that the film thickness be greater than a few times the London penetration depth λ_L, otherwise the microwave signal losses increase. λ_L is roughly 140–180 nm for a perfect crystal along the c axis of the unit cell[1] and is considerably greater in real films because of their structural imperfection, i.e., point defects, intergrain walls, etc. This means that films with satisfactory quality must be obtained with thicknesses over 0.5 μm. However, attempts to grow thick films have encountered major difficulties associated with a deterioration in quality with thickness and changes in the orientation of the growing film. YBCO films with their c axes oriented perpendicular to the substrate film (c orientation) are optimal from the standpoint of their electrical and physical properties and for use in microwave devices. Beyond some critical thickness t_C, the growth of grains with their c axes parallel to the surface (a orientation) begins to predominate. In a-oriented films the critical current density is an order of magnitude lower and λ_L is three times greater. Thus, we face the problem of avoiding changes in the orientation of the growing film during deposition of films with thicknesses of 0.5–1.5 μm.

Orientation during growth of YBCO films has been the subject of many papers.[1–4] It has been established that the formation of a-oriented grains takes place at lower temperatures than that of c-oriented grains, regardless of the substrate material. (Studies have been made of growth on substrates with good matching of the lattice parameters of the substrate and film (STO, $LaAlO_3$) and with poor matching (MgO), as well as on a previously deposited thin c-oriented YBCO film.) In films grown at the optimum temperature for formation of c-oriented grains, a-oriented grains have also been observed, either at the film–substrate interface[4,5] or at some film thickness.[1,3] In any case, beginning with the critical thickness t_C, growth of the a-oriented phase predominated. It has been suggested[4] that growth along the a axis is faster because the unit cell ''aligns itself'' more easily to the $a-c$ and $b-c$ facets than to $a-b$, while the orientation of the grains observed at the substrate in the initial stages of growth depends on temperature. This hypothesis explains the fact that the relative volume of the a phase rose with thickness in all the experiments. Computer calculations[4] have shown that for a 10:1 ratio of the growth rages, the amount of the a-oriented phase at the surface increases from 2 to 100% as the film grew to a thickness of 300 nm. These results are consistent with experiment.[4] On the other hand, it has been reported[6] that a-oriented growth is thermodynamically preferred, while c-oriented growth is kinetically preferred and can only take place with high supersaturation. Another hypothesis is that cooling of the growing film surface leads to a change in its orientation. According to the data of Refs. 1 and 2, in order for formation of a-phase grains to predominate, the temperature of the growing surface must fall by 120 °C. Several possible mechanisms for this amount of cooling have been proposed, but direct observations of such a large temperature drop have not been confirmed. However, it has been possible[2] to increase t_C to several microns using a linear or stepwise increase in the temperature during film growth. Nevertheless, the structural perfection of these films deteriorated with thickness and the volume of the a-oriented phase increased, although it remained rather small.

It may be assumed that both a- and c-oriented grains develop at the substrate surface, as well as at the surface of existing grains. The rate at which they are formed depends on temperature, supersaturation, the orientation (c and a) of the previous layer, and the presence of defects on the surface.

We have recently studied the growth of secondary dielectric phases (inclusions) during the initial stages of growth of YBCO films.[7] It was found that these inclusions consist of CuO and $YBa_2Cu_3O_8$ phases and that their concentration varies with time as the first layers of the film grow and then reaches a saturation level of about 10^9 cm^{-2}. These inclusions can coexist with the growing film and are ''sinks'' for defects, nonstoichiometric atoms, and mechanical stresses. As they are dielectrics, they do not detract from the superconducting and microwave properties of the YBCO

films.[8] In this paper we assume that the secondary phases can affect the relative rates of growth of the *a*- and *c*-oriented YBCO (123) phases and study $YBa_2Cu_3O_{7-\delta}$ films with thicknesses of 0.3–2.6 μm grown under conditions such as to ensure the formation of dielectric inclusions.

EXPERIMENT

The films were deposited in a planar dc magnetron system in an atmosphere of pure oxygen at a pressure of 1 Torr. An oxygen enriched atmosphere ensures complete oxidation of the inclusions, ensuring that they are dielectric and permitting their use in microwave devices. $[1\underline{1}02]$ (*r*-cut) sapphire was used as a substrate with a 200-Å -thick sublayer of cerium oxide with a mixed (001)/(111) orientation.[9] Sputtering was done at two discharge currents, 400 and 200 mA, which ensured two different deposition rates. During the first 30 min of deposition the discharge current was raised, and then it was kept constant. The substrate holder temperature was maintained at a level of 650 °C and not raised during deposition. The deposition time varied from 5 to 45 h.

The thickness of the resulting films was measured with a DEKTAK 3030 profilometer and ranged from 0.1 to 2.6 μm. The rate of deposition determined from the thicknesses was 10.5 Å /min ($\pm$5%) for a current of 400 mA and and 3 Å /min ($\pm$5%) for a current of 200 mA. It should be noted that halving the discharge current (and, therefore, the deposition rate) reduces the growth rate by a factor of three. This is because the the growth rate is quite low and is close to the kinetic minimum at which all the atoms arriving at the substrate are reevaporated and the film does not grow.

The structural perfection of the grown films was studied using a Rigaku–Dmax Geigerflex x-ray diffractometer. The amounts of the *c*- and *a*-oriented phases were calculated from the relative intensities of the YBCO (005) (*c* orientation) and (200) (*a* orientation) peaks. To do this we compared the integrated intensities of the peaks obtained from the rocking curve and normalized to the intensities for powdered samples (we followed the method described in Ref. 1). The half-widths of the (005) peak from the rocking curve were used to estimate the degree of disorientation of the film along the *c* axis.

Rutherford backscattering (RBS) was also used to analyze the film structure. RBS with intermediate energy ions and an electrostatic analyzer provides a resolution of about 5 Å and yields information on the structural perfection of the films along the *c* axis, so that the quality of the film structures can be compared at a fixed depth.

The morphology of the film surfaces was studied using a scanning electron microscope. The concentration and average sizes of the objects on the surface were determined with the aid of a specially developed computer program.

The parameters that are most sensitive to the structural perfection and orientation, the critical current density j_C and the microwave surface resistance R_S, were chosen for comparing the electrical properties of these films. The critical current density was measured by a contactless method at 77 K and the surface resistance, in a copper cavity at 77 K at 60 GHz.

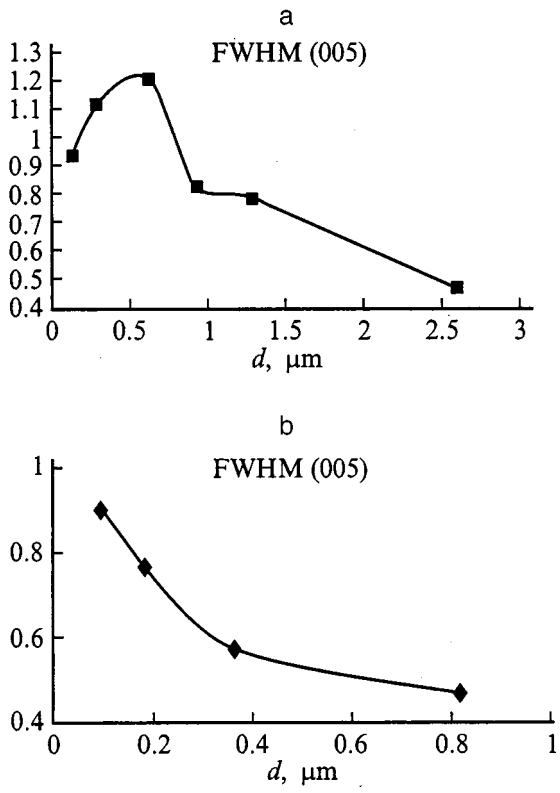

FIG. 1. The half-width of the (005) YBCO peak in the rocking curve as function of film thickness. Current (mA): a — 400, b — 200.

RESULTS

Figure 1a and 1b shows the dependence of the full width at half maximum of the (005) YBCO peak as a function of film thickness obtained by φ scanning. There is a clear tendency toward a reduction in the degree of disorientation of the film along the *c* axis for both deposition rates. The broadening caused by the finite thickness of the film is negligible: it can be estimated[10] as 0.1° for the thinnest film (thickness 0.1 μm). The instrument error is the same for all samples and is less than 0.1°, i.e., the curves in Fig. 1 reflect a genuine improvement in the structural perfection of the films with thickness. At the same time, the half-width of the (005) peak for relatively thin films with thicknesses of 0.1–0.3 μm is comparatively large, at 1–1.1°, and the 0.7–0.8° obtained for films with a thickness of 0.5 μm is typical for our films

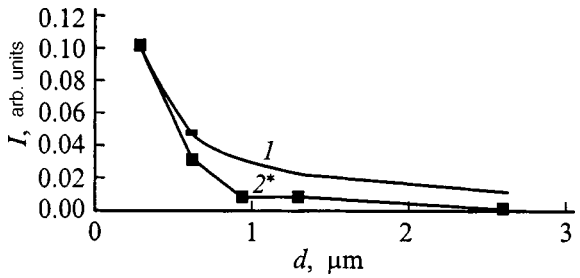

FIG. 2. The number of *a*-oriented grains: *1* — without annealing, *2* — with normalizing; current 400 mA; * — normalizing with a weight factor: $R = I(200)/T(200)/(I(005)/T(005) + I(200)/T(200))$, with $T(200) = 28.7$, $T(005) = 11.2$.

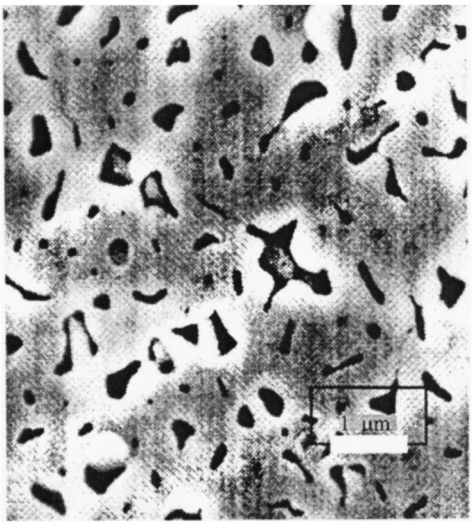

FIG. 3. The surface (scanning electron microscopy).

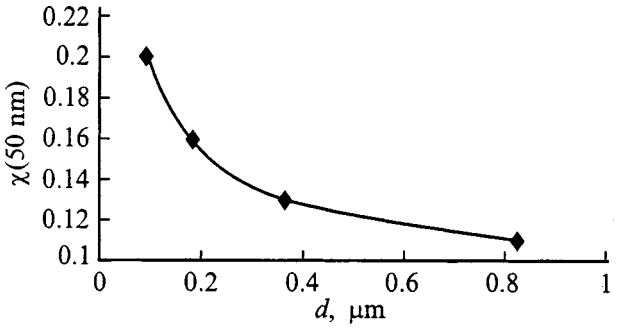

FIG. 4. Ratio of the channeled and random signals in the Rutherford backscattering spectrum as a function of film thickness.

used in the microwave devices, while the 0.45–0.5° observed in the films with a thickness of $1-2\ \mu$m is close to the lowest widths reported in the literature for much thinner films.

As the same time the half-width of the (005) peak is decreasing, the amount of the *a*-oriented phase decreases with increasing thickness. This dependence is shown in Fig. 2 for a discharge current of 400 mA. Note that the relative volume of the *a* phases decreases more rapidly with increasing film thickness than if the initial amount of *a*-oriented grains was kept constant (Fig. 2). This means that some sort of recrystallization is taking place: annealing in a pure oxygen atmosphere and with a constant flux of the components of YBCO onto the substrate. For all the films deposited at a lower deposition rate (current 200 mA), and for the films with a thickness of 2.6 μm deposited with a discharge current of 400 mA, no *a*-phase grains were observed by x-ray analysis, i.e., the intensity of the (200) peak was below the resolution threshold for our apparatus. This suggests an upper limit for the relative volume of the *a* phase of $\sim 5\%$.

Scanning electron microscopic studies of the film surfaces revealed two types of macroscopic defects: holes and inclusions of secondary phases (Fig. 3). Here there were more holes on the surface of the thicker samples, and more inclusions on the surface of the thinner samples. There are probably also slowly growing inclusions at the bottom of the holes. The overall concentration of macroscopic defects remains at a level of about $5\times 10^8\ \mathrm{cm}^{-2}$, which is close to the saturation level for the concentration of islands of secondary phases as a function of time in the initial stages of YBCO film growth.[11]

RBS analysis, which was used to compare the magnitude of the dechannelling $\chi(50)$, confirmed that as the film thickness increases, the orientation of the structure improves. $\chi(50)$ is proportional to the ratio of the channelled and unchannelled signals at a depth of 50 mm from the film surface.

The dechannelled χ signal is higher for films with a greater concentration of defects, intermediate layers, and damage in the lattice. It also rises with increasing disorientation of the grains along the channels perpendicular to the surface. Figure 4 is a comparison for films deposited at 200 mA. $\chi(t)$, where t is the distance from the film surface also yields information on the types of defects. In our case, this function can be approximated linearly, as $\chi=\chi(0)+Dt$. The values of $\chi(0)$, $\chi(50)$, and D are listed in Table I. $\chi(0)$ and D, measured for a single crystal and calculated by a Monte Carlo method for a perfect crystal, are also listed there. Note that, in order to evaluate the structural perfection we have chosen $\chi(50)$, which is observed directly, while $\chi(0)$ and D are approximations. D reflects the presence of defects in the volume of the film, such as point defects and dislocations. The concentration of these defects can be estimated if their type is known. Assuming that the defects are dislocations, we estimated their concentration as $5\times 10^{10}\ \mathrm{cm}^{-2}$. In turn, the difference between the measured and theoretical values of χ develops because of disorientation of the grains with

TABLE I.

Sample	Current, A	Thickness, 10^{15}at./cm^2	Thickness, μm	$\chi(50\ \mathrm{nm})$	D, 10^{-19} cm^2/at.	$\chi(0)$
507	0.4	770	0.11	0.2	1.4	0.14
508	0.2	620	0.093	0.2	2	0.1
509	0.2	1250	0.183	0.16	2	0.07
513	0.2	2000	0.366	0.13	1.4	0.06
516	0.6	3100	0.4	0.17	1.4	0.1
512	0.2	>4000	0.8235	0.11	1.2	0.06
515	0.4	>8000	2.6	0.11	1.3	0.065
Single crystal (theoretical min.)					0.7 0.2	0.03 0.015

Note: The thickness was measured by RBS and only for films with thicknesses less than 0.5 μm.

respect to the *c* axis or the presence of nonchannelling phases at the surface. In the first case, our films should have had a disorientation on the order of 0.12–0.45°, which is smaller than that observed in the x-ray analysis. The second variant seems more probable, since the macroscopic defects that are visible in the microphotographs are most likely not channelling defects. The contribution of inclusions from secondary phases at the surface to the nonchannelling signal is proportional to their concentration and average size, while the contribution of the holes also depends on their depth. Taking both types of macroscopic defects into account, we should reduce $\chi(0)$ for the *c*-oriented phase to 0.11 in the best films. This value is close to that measured for single crystals.

The measurements of the critical current density yielded less unambiguous results. The dependence of the critical current density on the film thickness was irregular; j_C varied over 0.9×10^6 to 2×10^6 A/cm^2. It should be noted that both values are typical of good films. The best results were obtained for films of intermediate thickness. One possible explanation for this effect may be the presence of dielectric macroscopic defects. It is not known how they affect the critical current recorded by the contactless inductive method. Microwave measurements showed that the surface resistance of all the films thicker than 0.2 μm was less than 50 mΩ for 60 GHz at 77 K. A study of the characteristics of microwave devices (filters) fabricated using our films allowed us to compare the properties of films with different thicknesses in more detail. The results of this study will be reported later.

We conclude that the structural perfection of our films, which was fairly low for films with thicknesses of 0.1–0.3 μm, improved with the thickness of the films, and for films with thicknesses of 1–2 μm, approached the best results from processes optimized for growing thin single-phase YBCO films. The amount of the *a* phase also decreased with increasing thickness and the grains of the *a* phase which grew at the start underwent recrystallization. The production of YBCO films with thicknesses of 1–2 μm that have high structural perfection and a low content of the *a*-oriented phase has also been reported previously, but the tendency toward a reduction in these parameters with increasing thickness is a new fact. Given that the growth of the *a* phase predominates at lower temperatures,[1,2] we might conclude that in our case the surface of the growing film is not cooled. Since heating of the surface is unlikely (there is no mechanism for such heating in thick films with a constant heater temperature), the surface temperature does not change enough to affect the film orientation. Thus, some other factors must predominate; of these, only macroscopic defects are on hand. The physics of their interaction with growing films has not yet been studied. They certainly act as sinks for defects and nonstoichiometric atoms. In addition, they can act as preferred sites for nucleation of the *a* phase, limiting its growth and changing the relationship between the volumes of the *c*- and *a*-oriented grains in favor of the *c* orientation. The above results suggest that the formation of secondary phases and macroscopic defects may be used in growing high-quality YBCO films of the required thickness.

In sum, a study of these thick (up to 2.6 μm) YBCO films has shown that the amount of the *a*-oriented phase decreases as the thickness increases (contrary to previously published data) and is less than 0.5% for the thickest films (2.6 μm) grown with a discharge current of 400 mA and for all the films obtained at 200 mA. The orientation of the *c* phase also improves with thickness, as confirmed by RBS and x-ray diffraction studies. Macroscopic defects have been mentioned as a possible cause of such a phenomenon. It is assumed that these defects, which have were observed by scanning electron microscopy, serve to stabilize the growth of the *c*-oriented phase and inhibit the growth of the *a* phase. In order to clarify the mechanism for the structural improvements and the role of macroscopic defects in this process, further studies of the growth of YBCO films with dielectric inclusions will be required. In particular, the question of what thickness of high-quality *c*-oriented films can be obtained is of practical interest.

[1] F. Vassenden, G. Linker, and J. Geerk, Physica C **175**, 566 (1991).
[2] S. Sievers, F. Mattheis, and H. Krebs, Appl. Phys. **78**, 5545 (1995).
[3] A. H. Carim, S. N. Basu, and R. E. Muenchausen, Appl. Phys. Lett. **58**, 871 (1991).
[4] C. W. Nieh, L. Anthony, J. Y. Josefowicz *et al.*, Appl. Phys. Lett. **56**, 2138 (1990).
[5] E. Garcia-Gonzalez, G. Wagner, and M. Reedyk, Appl. Phys. **78**, 353 (1995).
[6] S. J. Pennycook, M. F. Chisholm, D. E. Jesson *et al.*, Physica C **202**, 1 (1992).
[7] E. K. Hollmann, V. I. Goldrin, D. A. Plotkin *et al.*, Tech. Phys. Lett. **22**, 942 (1996).
[8] Y. J. Tian, L. P. Guo, L. Li *et al.*, Appl. Phys. Lett. **65**, 234 (1994).
[9] E. K. Hollmann, A. G. Zaitsev, V. E. Loginov *et al.*, J. Phys. D **26**, 504 (1993).
[10] J. P. Gong, M. Kawasaki, and K. Fujito, Phys. Rev. B **50**, 3280 (1994).
[11] E. K. Hollmann, V. I. Goldrin, D. A. Plotkin *et al.*, Fiz. Tverd. Tela (St. Petersburg) **39**, 216 (1997) [Phys. Solid State **39**, 189 (1997)].

Translated by D. H. McNeill

Sputtering of tungsten, tungsten oxide, and tungsten–carbon mixed layers by deuterium ions in the threshold energy region

M. I. Gusev and S. N. Korshunov

Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

A. L. Suvorov and N. E. Lazarev

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

(Submitted June 22, 1998)
Zh. Tekh. Fiz. **69**, 137–142 (September 1999)

An original experimental method is developed for determining the sputtering coefficients of electrically conducting materials during bombardment by light gas ions at threshold energies. This information is very valuable in both purely scientific and practical terms. The basis of the method is a special field-ion-microscopic analysis regime. The procedure for measuring the sputtering coefficients includes cleaning the surface by field desorption and evaporation, with the subsequent work on an atomically clean and atomically smooth surface. The method permits identification of single vacancies on the irradiated surface, i.e., it is possible to count individual sputtered atoms. The method is tested on commercially pure tungsten, tungsten oxide, and a W–C mixed layer on tungsten under deuterium ion bombardment. The energy dependences of the sputtering coefficients of these materials for sputtering by deuterium ions at energies of 10–500 eV are obtained and analyzed. An important relationship between the energy threshold for sputtering and the conditions for oxidation of tungsten is found. The energy threshold for sputtering of an oxidized tungsten surface is 65 eV. The energy threshold for sputtering of the W–C mixed layer is almost equal to the corresponding value for pure tungsten. [S1063-7842(99)02909-8]

INTRODUCTION

Carbon composites (CFC) and tungsten are currently under examination as candidate materials for controlled thermonuclear reactor designs. CFC and tungsten are to be used for the most energy stressed parts of the diverter system, which are subjected to high fluxes of particles and electromagnetic radiation from the plasma. The presence of different materials in the components of the divertor and first wall will inevitably lead to the formation of pulverized, mixed layers on the surface of the facing materials in the divertor and first wall.

Tungsten is characterized by a high mass number and, because of this, a high self-sputtering coefficient. In addition, because of the high chemical affinity of tungsten for oxygen, its surface is always oxidized. It has been established[1] that the presence on a tungsten surface of tungsten oxide, formed during bombardment of the tungsten surface by light ions when a high partial pressure of oxygen is present (8×10^{-5} Torr) in a vacuum vessel, causes a substantial reduction in the threshold for sputtering of tungsten owing to a drop in the binding energy of the atoms to the surface.

Under conditions such that the materials in the divertor and first wall of a thermonuclear reactor are pulverized, the surface layer of tungsten will consist of a mixture of tungsten and oxygen atoms. In this regard, a study of the effect of oxide and mixed W–C layers on the threshold sputtering energy of tungsten is of fundamental importance.

It should be noted that the experimental determination of sputtering thresholds requires especially sensitive, high-resolution methods capable of observing each vacancy formed on a sputtered surface by a bombarding ion (i.e., of counting each atom removed from the surface). Of the modern experimental techniques for surface diagnostics, these requirements are best met by field-ion microscopy, which is widely used for examining the structure of surface atomic layers, studies of the formation, behavior, and evolution of defects in crystal structure, and analyzing various radiation defects on the surface and in the volume of electrically conducting materials.[2,3]

In this paper field-ion microscopy is used for the first time to determine the energy thresholds for sputtering of tungsten, tungsten oxide, and mixed W–C layers on tungsten by D^+ ions, as well as to measure the energy dependences of the sputtering coefficients of these materials for sputtering by D^+ ions in the threshold energy region.

EXPERIMENTAL TECHNIQUE

We have used standard needle tips for field-ion microscopy with average radii of curvature at the tip on the order of $\overline{R_0}\leqslant 100$ nm as test samples in this work. In order to obtain a layer of oxide on the tungsten surface, the samples were heated in the atmosphere to a temperature of $\sim 750\,^{\circ}$C. Films of a W–C mixture were prepared by collecting the products of simultaneous sputtering of tungsten and graphite by 20 keV Ar^+ ions on needle and bulk samples of tungsten. The area of the ion beam was 2×10^{-2} m^2. The collectors

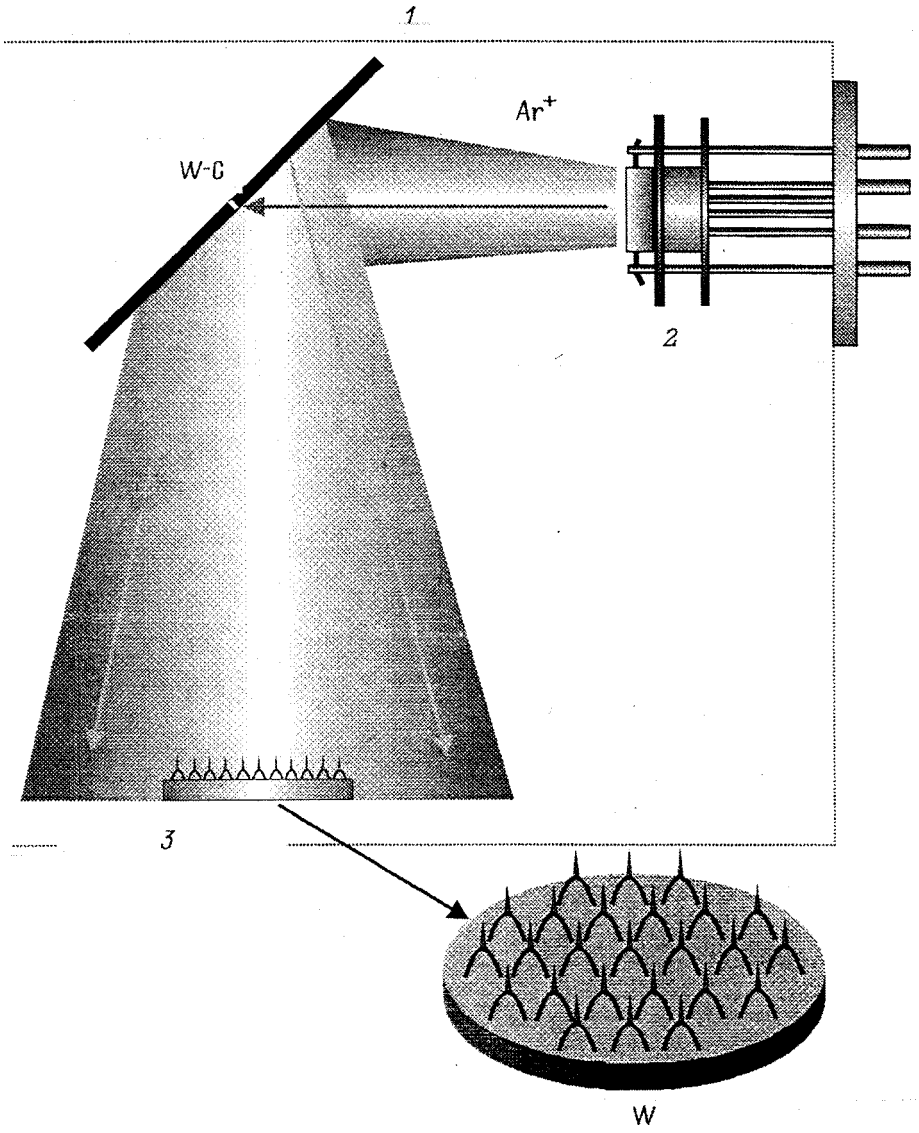

FIG. 1. Experimental apparatus for creating mixed W–C layers on the surface of tungsten samples– tips for field-ion-microscopy: *1* — vacuum chamber, *2* — ion source, *3* — collector.

were placed a distance of 25 mm from the sputtering targets. Figure 1 is a sketch of the apparatus used to deposit the mixed layer.

The thicknesses of the oxide and mixed layers, and their roughness and chemical composition were determined using a Sloan Instruments profilometer and Auger electron analysis in combination with layer-by-layer etching on the bulk samples. The phase composition of the samples was determined using x-ray structural analysis in a grazing incidence beam geometry. Figure 2 shows a typical depth distribution of oxygen on the oxidized tungsten surface, and Fig. 3 shows the depth distribution of carbon in a W–C mixed layer.

In this paper the sputtering coefficient of tungsten and its compounds was determined by directly counting the number of vacancies that were formed. (See, for example, the field-ion microscope image of a pure tungsten surface in Fig. 4, where the individual vacancies formed during a single beam pulse are indicated by arrows.)

A field-ion microscope of original design[4] at the Institute of Theoretical and Experimental Physics was used in our experiments. In order to form a system of pulses between the needle sample (cathode) and the microchannel plate (anode), a metal-ceramic cathode assembly grid was used. This assembly was placed in the spacce between the sample and fluorescent screen only during single beam irradiation pulses.

An original method for a pulsed two-step change in the polarity of the high voltage was developed and used for irradiating the samples. Here the gas, which could also serve for imaging, was ionized by electron impact at the moment a short high-voltage pulse of the opposite (negative) polarity was delivered and the sample generated electrons by field emission. These ''ion imaging'' high-voltage pulses were shorter than 0.1 μs and their amplitude was 5.0 kV. At the same time as the high-voltage pulse, a low-voltage negative pulse was applied to the sample; its duration was $\geqslant 10\,\mu$s and its amplitude, which was varied between 10 and 500 V with a step of at least 10 V, corresponded to the specified energy of the bombarding ions. Various control analyses were carried out in parallel, with a different sequence of pulses and their parameters (in diode and triode configurations; in the latter case the above mentioned cathode assembly grid was used).

The procedure for measuring the sputtering coefficients involved the following basic operations: a) fabrication of the

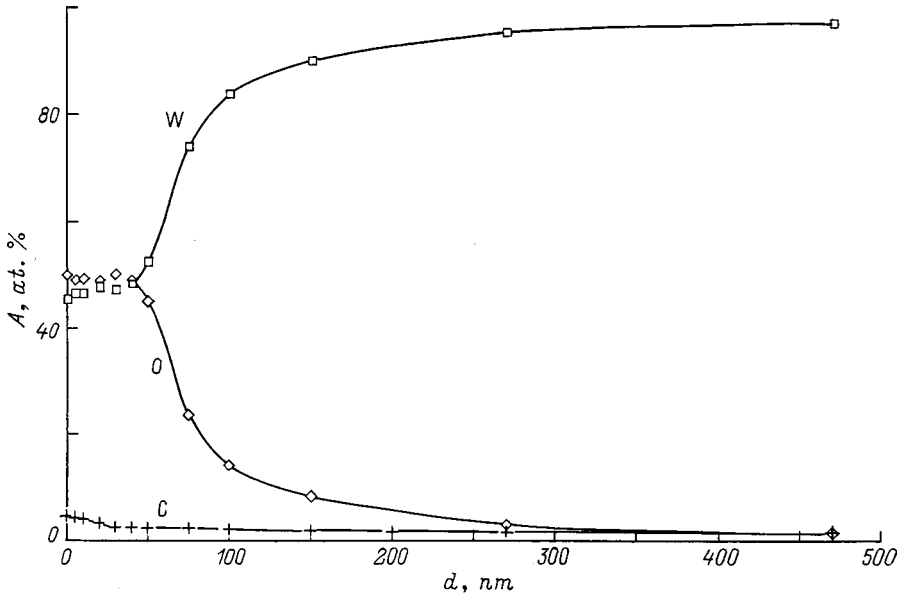

FIG. 2. Depth distributions of elements in the surface layer of WO after oxidation of tungsten samples: A is the concentration and d the depth.

sample tips and their installation in a field-ion microscope capable of operating in three modes: a field-ion microscope *per se*, a desorption ion microscope, and a field-emission microscope;[5] b) vacuum pumping of the microscope and admission of deuterium; the initial vacuum was 3×10^{-9} Torr, and the pressure after the deuterium was admitted was 10^{-6}–10^{-5} Torr; c) a preliminary field-ion or desorption microscopy analysis of the initial surface, cleaning and evaporation of the surface by the field to ensure maximum smoothness; d) pulsed irradiation of the sample surface by D^+ ions; e) a second microscopic analysis of the samples with the object of identifying the formation of isolated vacancies on their surface; and, f) calculation of the sputtering coefficients corresponding to the given bombarding ion energy. For a given sample, operations c–f were repeated many times. The number of bombarding pulses ranged from 10 to several hundred for low sputtering coefficients.

An important stage in the procedure for determining the sputtering coefficients Y was calculating the irradiation fluences. We emphasize that this is by no means a trivial problem. This problem, together with the estimate of the bombarding ion energy and the calculation of the number of individual vacancies formed during bombardment, imply a certain arbitrariness in the resulting values of Y. In this paper, the irradiation fluence was calculated using the formulas of Ref. 6. Here the required values of the average work function $\overline{\varphi}$ for W, WO, and WC (4.54, 5.0, and 3.6 eV, respectively) were taken from Ref. 7. For the given irradiation parameters, the surface of a tip with an average radius at its end of $\overline{R_0}$ $\approx$100 nm experienced $\sim2.5\times10^7$ D^+ ion impacts per sec-

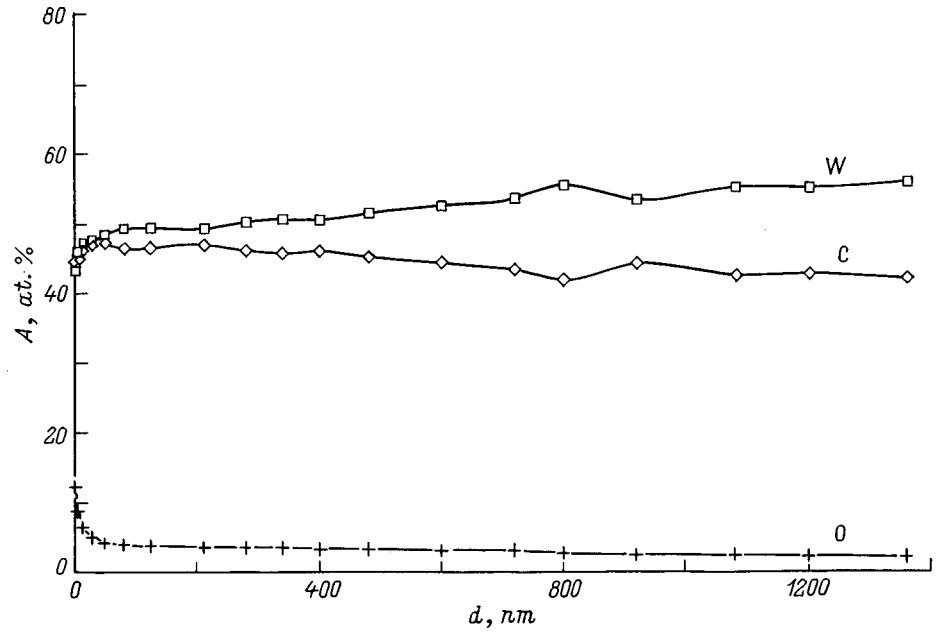

FIG. 3. Depth distribution of elements in a mixed surface layer of W–C on a tungsten sample.

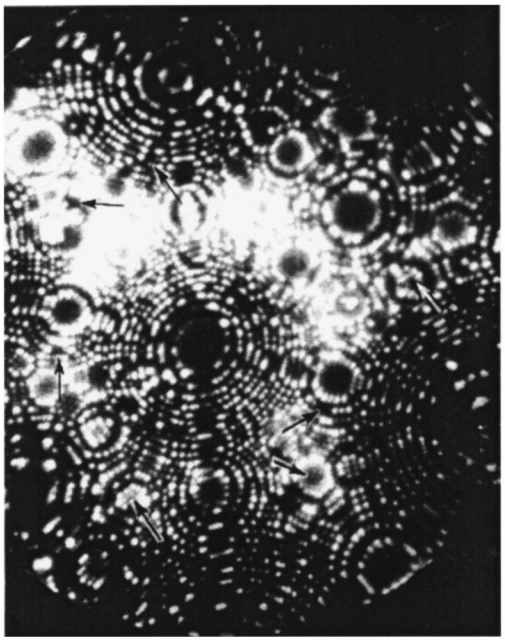

FIG. 4. Field-ion microscope image of the surface of a sample of pure tungsten: the arrows indicate individual vacancies formed as a result of a single pulse of D^+ ions.

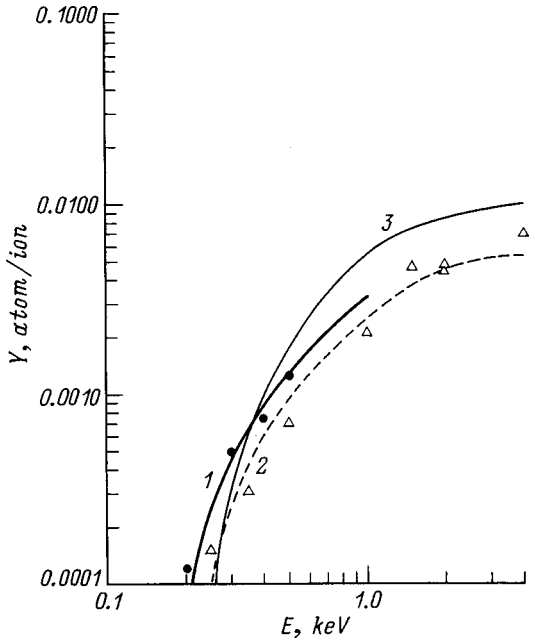

FIG. 5. Energy dependence of the sputtering coefficient Y of tungsten for sputtering by D^+ ions: *1* — results of the present work, *2* — experimental data of Ref. 9, *3* — theoretical calculation[9].

ond, which corresponds to 250 ions striking the sample surface during a single pulse or a flux of 2.5×10^{17} ions/cm^{2}s. In turn, $Y\approx10^{-1}$ corresponds to the observation of 25 individual vacancies over the entire sample surface observed under the microscope after a single beam pulse (the field-ion image of the end surface of a sample tip with $\overline{R}_0\sim100$ nm contains $\geqslant10^5$ atoms), $Y\approx10^{-2}$ corresponds to an observation of 5 vacancies in two beam pulses, $Y\approx10^{-3}$ corresponds to observation of 1 vacancy in 4 pulses, etc.

EXPERIMENTAL RESULTS

Table I is a comparison of the experimental values of the threshold sputtering energies of pure tungsten, tungsten carbide, a mixture of W–C, and tungsten oxide by deuterium ions, together with corresponding published experimental and theoretical data.[1,8–10]

The table shows that during D^+ ion bombardment there are no significant differences between E_{thr} for W and for a W–C mixed layer. The results for W and the W–C mixture are in good agreement with previous measurements of the threshold sputtering energy of W and W–C.[1,8,9]

TABLE I. Threshold sputtering energies.

Material No.	Material	T, K	E_{thr}, eV (experiment, present paper)	E_{thr}, eV (experiment)	E_{thr}, eV (theory)
1	W	293	160	175 (Ref. 8) 178 (Ref. 9)	201 (Ref. 9) 160 (Ref. 10) 209, 37 (Ref. 9)
2	W + C	78–293	150		
3	WC			171 (Ref. 9) 150 (Ref. 8)	
4	W oxide	293	65	$<$18 (Ref. 1)	

The threshold energy for sputtering of tungsten oxide by deuterium ions measured by field-ion microscopy is 65 eV, as compared to $E_{thr}\approx18$ eV for tungsten oxide sputtered by D^+ ions in a vessel with a residual oxygen pressure of 8×10^{-5} Torr measured by weighing. As noted in the reference, the low value of E_{thr} is caused by a low binding energy ($E_c=0.3$ eV) of the tungsten oxide molecules to the surface.[1] In our experiments, an oxide film with a thickness of ~50 nm developed on the tungsten surface. An estimate of the binding energy analogous to that in Ref. 1 gives ~1.1 eV.

The substantial increase in the threshold energy for sputtering of a tungsten oxide layer by deuterium ions observed by the new method employed here may also be related to differences between these two methods (weighing and field-ion microscopy). The new method identifies only the vacancies in the tungsten, that is, it counts only tungsten atoms,while the weighing method records all sputtered surface species, including adsorbed atoms and molecules.

Auger electron analysis of a mixed W–C layer showed that in a layer with a thickness of $\sim10^4$ nm, the W and C are distributed essentially uniformly, while the oxygen impurity on the surface is no more than 9.5 at. %. The thickness of the layer of WO on the tungsten resulting from heating it in air is ~50 nm.

The energy dependence of the sputtering coefficient for tungsten by deuterium ions obtained in the present experiments is shown in Fig. 5 (curve *1*). Besides these data, Fig. 5 also shows the experimental (curve *2*) and theoretical (curve *3*) energy dependences obtained in Ref. 9. As can be seen from Fig. 5, the tungsten sputtering coefficients measured by

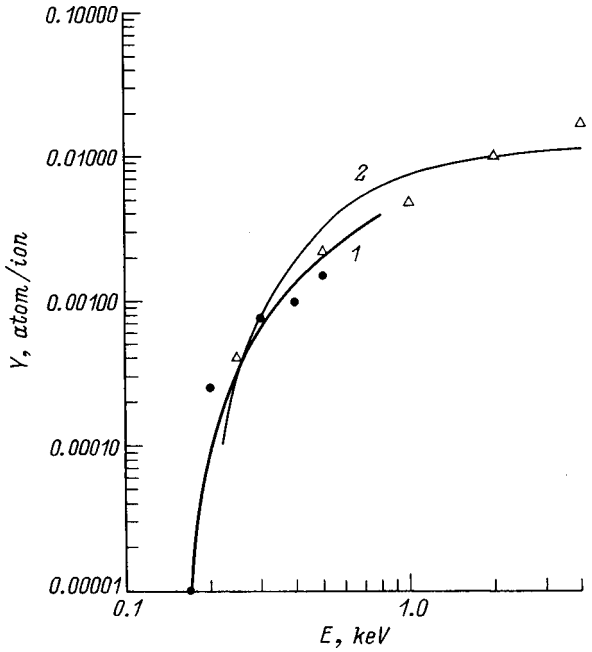

FIG. 6. Energy dependence of the sputtering coefficient of a mixed W– C layer on a tungsten surface for sputtering by D^+ ions: *1* — results of the present work; *2* — experimental data from Refs. 9 and 10 for a WC system.

the different methods are similar and agree fairly well with the theoretical curve.

The energy dependences of the sputtering coefficients for a pulverized layer of W–C and WC[9] are also essentially the same as those obtained here (Fig. 6).

Figure 7 shows the energy dependence of the sputtering coefficient for tungsten oxide with sputtering by deuterium ions (curve *1*). Experimental data from Ref. 1 for tungsten

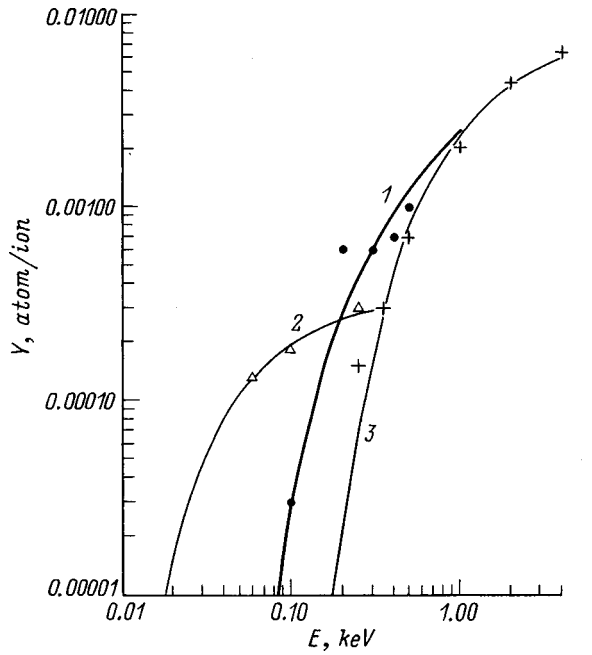

FIG. 7. Energy dependence of the sputtering coefficient of tungsten oxide on the surface of a tungsten sample for sputtering by D^+ ions: *1* — results of the present work, *2* — data of Ref. 9 for tungsten for an oxygen pressure of 8×10^{-5} Torr, *3* — data of Ref. 1 for pure tungsten.

sputtered at an oxygen pressure of 8×10^{-5} Torr (curve *2*) and for pure tungsten (curve *3*) are shown there for comparison. At energies <100 eV, the sputtering coefficients for tungsten irradiated in a high background pressure of oxygen substantially exceed the field-ion microscopy measurements. For $E\geqslant350$ eV the sputtering coefficients for tungsten oxide (curve *2* of Fig. 7) measured by the weighing method approach those for pure tungsten (curve *3*). This means that when the sputtering coefficient increased to 3×10^{-4} atom/ion under the experimental conditions of Ref. 1, the oxide film on the tungsten surface had already been sputtered away by the D^+ ions. In our experiments, a thick film of tungsten oxide was present on the tungsten surface, so that a higher sputtering coefficient was required to sputter it away.

CONCLUSION

1. It has been shown that measuring the sputtering coefficients of materials by field-ion microscopy has the following advantages: the ability to clean the sample surfaces of adsorbed atoms and molecules, of oxide films, etc., directly in the field-ion microscope and the possibility of observing and identifying each vacancy produced by the bombarding ions on the sample surfaces, i.e., the possibility of counting each sputtered atom.

2. A sharp increase has been found in the threshold energy for sputtering of tungsten oxide by deuterium ions. E_{thr} is 65 eV.

3. The threshold energy for sputtering of pure tungsten by deuterium ions at room temperature is in satisfactory agreement with the calculations of Ref. 1, but is considerably lower than the corresponding values determined by the weighing method at room temperature.

4. The energy dependences of the sputtering coefficients of pure tungsten and of a W–C mixed layer on tungsten for sputtering by deuterium ions are in agreement with previously published data for W and WC.

[1] J. Roth, J. Bohdansky, and W. Ottenberger, ''Data on low energy light ion sputtering,'' Max-Planck-Institut für Plasmaphysik. Report IPP 9/26 (1979).
[2] E. W. Muller and T. T. Tsong, *Field Ion Microscopy. Principles and Application*, Elsevier, New York (1969).
[3] A. L. Suvorov, *Field-Ion Microscopy of Radiation Defects in Metals* [in Russian], Énergoizdat, Moscow (1982), 167 pp.
[4] A. L. Suvorov *et al.*, PTÉ, No. 6, 188 (1985).
[5] A. L. Suvorov, *Structure and Properties of Surface Atomic Layers on Metals* [in Russian], Énergoatomizdat, Moscow (1990), 296 pp.
[6] P. A. Bereznyak and V. V. Slezov, Radiotekh. Elektron. **17**, 354 (1972).
[7] V. S. Fomenko, *Emission Properties of Materials* [in Russian], Naukova Dumka, Kiev (1981), 340 pp.
[8] J. Bohdansky, *Research Coordination Meeting on Plasma-Interaction Induced Erosion of Fusion Reactor Materials*, IAEA, Vienna (1989).
[9] W. Ekstein, C. Garcia-Rosales, J. Roth, and W Ottenberger, ''Sputtering data,'' Max-Planck-Institut für Plasmaphysik, Report IPP 9/82 (1993).
[10] J. Roth, J. Bohdansky, and A. P. Martinelli, *Proceedings of the First Conference on Ion Beam Modification of Materials*, Budapest (1978), p. 1541.

Translated by D. H. McNeill