Pseudoultrarelativistic behavior and specification of spinlike effects in the two-dimensional electron gas in Kane semiconductors with direct and inverted band structure

V. F. Radantsev, T. I. Deryabina, G. I. Kulaev, and E. L. Rumyantsev
Institute of Physics & Applied Mathematics, Ural State University, Ekaterinburg 620083, Russia
(Received 11 August 1995)

The comparative study of the spin splitting of two-dimensional subbands in surface layers on small gap $Hg_xCd_{1-x}Te$, with direct ($E_g>0$) and inverted ($E_g<0$) bands, is carried out experimentally and theoretically. The occupations of spin-splitted subbands n_i^+ are determined from Fourier transforms of capacitance magneto-oscillations of $Hg_xCd_{1-x}Te$ metal-oxide-semiconductor structures. The values $(n_i^- - n_i^+)/(n_i^- + n_i^+)$ are found to be independent of surface densities n_s, subband index i, and $|E_g|$ at high enough n_s when the band bending sufficiently exceeds the gap E_g and the conditions corresponding to pseudoultrarelativistic behavior of surface electrons are fulfilled. However, they are different for $E_g>0$ and $E_g<0$ cases. To analyze spinlike effects in narrow-gap and gapless semiconductors, we employed for Kane Hamiltonian the conception, offered by Zel'dovich and Migdal for the description of vacuum condensate of Dirac's electrons near supercritical nuclei. In this way, we obtain ''usual'' Schrodinger-like subband equations with some effective potential. The terms responsible for nonparabolicity, spin-orbit splitting, and ''resonant'' shift, due to interband mixing by surface electric fields, are easily singled out. Such equations admit the simple physical interpretation, and difference in values of ''spin'' effects for $E_g>0$ and $E_g<0$ cases is easily seen. The dependencies of total subband occupations $n_i=n_i^- + n_i^+$ and average cyclotron masses on n_s nevertheless are close for both cases in agreement with previous experimental observations. In a pseudoultrarelativistic limit $E_g=0$ the simple analytical expressions for subband parameters of experimental interest are obtained with an allowance for spin effects. The calculations agree with experiment for $Hg_xCd_{1-x}Te$ with both direct and inverted bands, excluding the region of low n_s in heavily doped samples. Possible reasons for disagreement are discussed. [S0163-1829(96)00420-1]

I. INTRODUCTION

The interest in both theoretical[1–11] and experimental[12–21] investigations of two-dimensional electron gas in the surface layers on the narrow-gap semiconductors, essentially extensive since the mid 1980s, is caused by a various number of the specific peculiarities inherent to these systems. Part of them directly follows from the smallness of effective mass: (a) the large depth and width of surface quantum well and as a result the multisubband character of spectrum; (b) the high values of Fermi energy and, consequently, the weak impact of many-body effects and fluctuations of potential; (c) the degeneration of electron gas in substrate for accumulation layers leading to the effects caused by the involvement of continuum electrons in the screening of surface field.[22–24] The most important peculiarities of these systems follow from the multiband nature of the Hamiltonian describing the bulk spectrum in narrow-gap semiconductors. Neglecting spinlike effects (but keeping the Fermi statistic), the effective-mass equation for electrons in surface potential is reduced to Klein-Gordon (KG)-like equation, i.e., such systems are the relativistic analog, with respect to two-dimensional (2D) systems based on wide-gap semiconductors.[17,25] This leads to such relativisticlike effects as strong nonparabolicity and kinetic confinement (motional binding).[24,26,27]

Even for small surface densities n_s the band bending in narrow-gap semiconductors with the energy gap $|E_g|<100$ meV $[E_g=E(\Gamma_6)-E(\Gamma_8)]$ exceeds E_g and nonparabolicity not only cannot be ignored, but also cannot be consider as the correction to the parabolic approach. What is more, the ''rest energy'' $ms^2=|E_g/2|$ (m and s are Kane mass and Kane velocity, respectively) can be neglected in the KG subband equation for the wide range of n_s of experimental interest. The direct consequences of such a pseudoultrarelativistic character of electron motion in the surface layers of narrow-gap semiconductors (at moderate doping level when the contribution to the screening from the depletion field in inversion layers or from the continuum electrons in accumulation layers is small) are the scale invariance of subband spectrum, with respect to surface potential μ_s and the universality of subband parameters.[17] The latter implies the independence of ratios of subband occupations n_i/n_j and cyclotron masses m_{ci}/m_{cj} for different 2D subbands of surface density n_s (and, consequently, μ_s), band parameters of materials and subband index i (at $i,j>1$). The experiment shows that n_i and m_{ci}, as a function of n_s, are indeed well described by the universal (the same for all narrow-gap semiconductors) dependencies coinciding with those obtained within the frame of KG approach.[17,21]

Moreover the experimental dependencies n_i (n_s) and m_{ci} (n_s) are the same for narrow-gap semiconductors with both direct ($E_g>0$) and inverted ($E_g<0$) band structure.[21] This result is unexpected from theoretical point of view because, as we shall see, the similarity of the subband parameters in the cases $E_g>0$ and $E_g<0$ is to break down (even at $E_g=\pm 0$) if we take spinlike effects into account. At the same time, the large magnitude of spin effects (the spin-orbit splitting and the ''resonant'' shift of the subbands, due to the

mixing between states of Γ_6 and Γ_8 bands by the surface electric field) is yet another relativisticlike peculiarity of 2D systems on narrow-gap semiconductors. Unlike in wide-gap semiconductors, where the spin-orbit interaction is introduced by including in Hamiltonian additional Rashba term,[28] arising from inversion asymmetry on interface, in narrow-gap semiconductors, the spatial inversion is absent in the bulk (symmetry T_d) and the spin-orbit interaction is already present in the Kane Hamiltonian. Why a large magnitude of spin-orbit interaction simply does not affect n_i (n_s) and m_{ci} (n_s) dependencies and does not violate their equality for the case of $E_g>0$ and $E_g<0$ is a question requiring an explanation.

Experimental manifestations of spin-orbit splitting in the surface layers on $Hg_xCd_{1-x}Te$ have been observed in Refs. 29 and 30, but the results were interpreted within a simple empirical Rashba model. Such treatment is justified from a point of view of the comparison of Rashba parameters α_i in the wide-gap and narrow-gap semiconductors, where they differ by a factor of $\sim 10^2$. However, the error in the theoretical estimation of magnitude of spin-orbit splitting is of the order of this magnitude, because of used approximations. This does not allow the accurate comparison of parameters of spin-orbit splitting in narrow-gap semiconductors, with $E_g>0$ and $E_g<0$. On the other hand, because of the adequacy of ultrarelativistic (UR) limit $E_g=0$ and its scale invariance, one can expect also the universality of the parameters characterizing spin effects. In order to prove such universality and define these parameters, a more rigorous theoretical treatment is required. At the same time, the subband calculations with an allowance for spin effects undertaken in a number of works (used several models and various boundary conditions) were done only for narrow-gap semiconductors with direct but not inverted band structure.

The purpose of this paper is the comparative experimental and theoretical investigations of spin effects in Kane semiconductors with direct and inverted bands (experimentally on $Hg_xCd_{1-x}Te$) and also in semiconductors with a Dirac-like spectrum realized in IV–VI semiconductors. In Sec. II, the experimental results concerning the spin-orbit splitting of a subband spectrum in inversion and accumulation layers on narrow-gap semiconductor $Hg_xCd_{1-x}Te$ are presented and their specific character for the cases of $E_g>0$ and $E_g<0$ is shown. In Sec. III, we present the theoretical consideration based on the conception developed previously for the related problem of the description of vacuum condensate of Dirac electrons near nuclei with supercritical charge. In the framework of such treatment, which is more transparent and easy to interpret in contrast to the method based on a direct numerical solution, the question reduces to Schrodinger-like equation with an effective potential in which the terms, responding for nonparabolicity, spin-orbit splitting and ''resonant'' shift are easy singled out. The qualitative similarity and the quantitative difference between cases $E_g>0$ and $E_g<0$ are clearly seen also. The boundary conditions in the bulk for envelopes of both surface and continuum states in this method are dictated by the form of derived effective potential. The more detailed results of subband calculations in an UR limit $E_g=0$, in which spin effects are more pronounced and which often correspond to typical experimental situation, are given in Sec. IV. We show that in this limit the simple analytical expressions for experimentally measured subband parameters, characterizing the spin-orbit splitting, may be obtained using the WKB approach. These results are presented in Sec. V. In Sec. VI, the spin-orbit splitting and resonant shift as functions of gap $|E_g|$ are presented. The results of the comparison of experiment and theory for the samples with $E_g>0$ and $E_g<0$ are discussed.

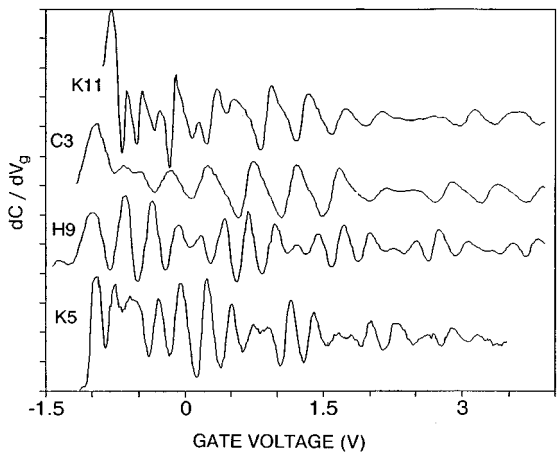

FIG. 1. Oscillations of the derivative dC/dV_g of differential capacitance C of $Hg_xCd_{1-x}Te$ MOS structures versus gate voltage V_g in magnetic field $B=4.5$ T, for the samples $K11$ ($E_g=+85$ meV, $N_A-N_D=2\times10^{16}$ cm^{-3}), $C3$ ($E_g=+50$ meV, $N_A-N_D=7\times10^{17}$ cm^{-3}), $K5$ ($E_g=-50$ meV, $N_D-N_A=3.5\times10^{15}$ cm^{-3}), and $H9$ ($E_g=-85$ meV, $N_A-N_D=7.5\times10^{17}$ cm^{-3}).

II. EXPERIMENTAL OBSERVATIONS

We investigated the two-dimensional electron gas in metal-oxide-semiconductor (MOS) structures based on $Hg_{1-x}Cd_xTe$ with $x=0.1$–0.22. An anodic oxide $\sim$1000 Å thick serves as a dielectric for MOS capacitors. The gate electrodes of the typical area $\sim5\times10^{-4}$ cm^2 were formed by the evaporation of Pb. In gapless semiconductors, because of the shunting of the surface conductivity by the volume, the traditional galvanometric techniques cannot be used. We employed the capacitance spectroscopy method in quantizing magnetic fields,[14] which is applicable for any E_g and high doping level also. Figure 1 shows the dependencies of derivative dC/dV_g of the differential capacitance C of MOS structures (the amplitude of test signal was equal to 10 meV, the frequency $f=1$ MHz) on the gate voltage V_g at $T=4.2$ K in the normal to the surface magnetic field $B=4.5$ T for the samples $K11$ ($E_g=+85$ meV, $N_A-N_D=2\times10^{16}$ cm^{-3}), $C3$ ($E_g=+50$ meV, $N_A-N_D=7\times10^{17}$ cm^{-3}), $K5$ ($E_g=-50$ meV, $N_D-N_A=3.5\times10^{15}$ cm^{-3}), and $H9$ ($E_g=-85$ meV, $N_A-N_D=7.5\times10^{17}$ cm^{-3}). The specific capacitances of the oxide for all MOS structures are approximately the same and correspond to the change of the induced charge surface density N_s on $\Delta N_s=(0.7$–$1)\times10^{12}$ cm^{-2} at $\Delta V_g=1$ V.

The specific feature of presented in Fig. 1 oscillations related to the ground subband is the oscillation beats, which indicate the existence of two ladders of Landau levels with the different periods in V_g (and in the reciprocal magnetic field B^{-1}). This is a typical manifestation of the spin-orbit splitting of the subband spectra.[30,33] In $Hg_xCd_{1-x}Te$ with

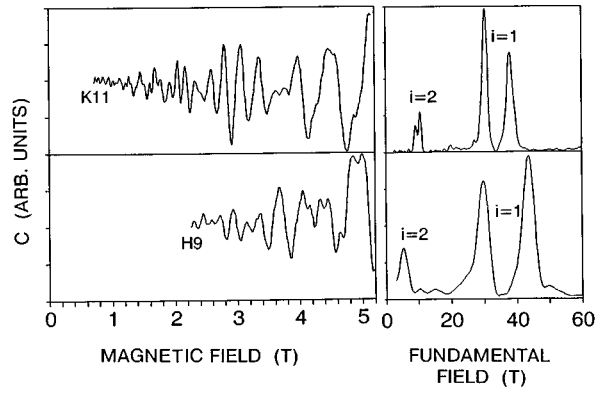

FIG. 2. Magneto-oscillations of capacitance at $V_g=9$ V and their Fourier transforms for $K11$ and $H9$ samples.

$E_g<0$, such behaviour is observed for excited subbands also. Due to the large magnitude of the beat periods and strong magnetoresistance of the surface channel in $Hg_xCd_{1-x}Te$ with $E_g>0$, which leads to the increase in series resistance of the MOS equivalent circuit and as a result of the sudden decreasing of the oscillation amplitudes in the region of large V_g, the second node of the beats has not been observed for these materials at $B=4.5$ T. The periodicity of oscillations on n_s in the range of small n_s shows that the influence of the depletion layer charge on the parameters of the subband is insignificant even in the materials with the greatest doping level $N_A-N_D\sim7.5\times10^{17}$ cm^{-3}.

The oscillations dC/dV_g (V_g) clearly indicate the universality of the beat structure: the ratio of the beat period Δ_{bi} to fundamental period Δ_i is almost independent of the surface density and magnitude of $|E_g|$, but essentially differs for $E_g>0$ ($\Delta_{bi}/\Delta_i\sim5.5$) and $E_g<0$ ($\Delta_{bi}/\Delta_i\sim3$). Such a large difference in values of spin splitting and consequently in beat periods explains the following experimental fact: while in lightly doped HgCdTe with $E_g>0$, the separate spin components are distinctly observed for low Landau levels ($N<3$) of the ground subband (see the oscillations for sample $K11$ in Fig. 1 in the range $V_g<0.5$ V, where their period is half as much as at $V_g>0.5$ V), in materials with inverted bands, in which the frequency of the beats is twice as large, the individual spin components have not been observed at any doping level.

Although the oscillations dC/dV_g (V_g) illustrate well the universality of the beat structure, the comparison with the theory is difficult and does not bear the clear physical interpretation, because of the changing of the subband spectrum under the varying of applied gate bias voltage V_g. From this viewpoint, the measurements of the capacitance (or dC/dV_g) oscillations as a function of magnetic field at the fixed gate voltage are of the most interest. Such investigations give, for the same band bending, the information about parameters of the several subbands, essentially extending the amount of extracted information and reliability of the used theoretical models. In Fig. 2, such oscillations are shown for the first and second excited subbands. The Fourier transforms of these oscillations distinctly demonstrate for every electric subband the presence of two frequencies connected with spin splitting of the subband spectra, because of the spin-orbit interaction. The surface densities of the carriers in the spin-

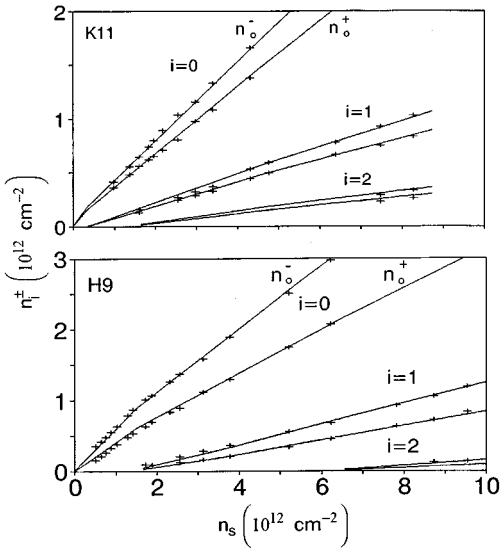

FIG. 3. Experimental (+) and theoretical (—) occupations of spin-orbit split subbands $n_i^{\pm}$ versus the total surface density of electrons n_s in inversion layers of $Hg_xCd_{1-x}Te$, with direct (sample $K11$) and inverted (sample $H9$) bands.

splitted subbands n_i^+ and n_i^- determined from Fourier transform of the magneto-oscillations are plotted in Fig. 3 as the functions of total surface density for samples $K11$ and $H9$. The relative differences of the occupations of spin-splitted subbands $\Delta n_i/n_i=(n_i^- - n_i^+)/(n_i^- + n_i^+)$ at high enough surface densities n_s practically do not depend on n_s and subband index i and are equal to ≈0.09 for $Hg_xCd_{1-x}Te$ with $E_g>0$ ($K11$ and $C3$ samples) and ≈0.2 for $Hg_xCd_{1-x}Te$ with $E_g<0$ ($H9$ and $K5$ samples). The corresponding ratios $\Delta_{bi}/\Delta_i=(2\Delta n_i/n_i)^{-1}$ agree well with those, obtained from beating dC/dV_g (V_g) oscillations.

Fourier analysis of SdH oscillations in $Hg_xCd_{1-x}Te$ with $E_g>0$ at $n_s=8\times10^{11}$ cm^{-2} was used recently in Ref. 20. The authors did not observe the spin splitting. This may be due to the insufficient resolution of Fourier transforms. Our analysis of experimental data in Ref. 20 shows the sharp splitting in the first excited subband and gives the ratio $\Delta_{bi}/\Delta_i=5.5$ in agreement with the above represented results for $Hg_xCd_{1-x}Te$ with $E_g>0$.

III. THEORETICAL CONSIDERATION

For the purpose of studying subband states in the materials with a small gap, we may neglect the contributions from remote bands (including Γ_7 band) and the free-electron term and use 6×6 Kane $\mathbf{k}\cdot\mathbf{p}$ model with a ''dispersionless'' heavy-hole branch (the heavy-hole effective mass $m_{hh}=\infty$). Because of the axial symmetry, with respect to the dispersion in a 2D plane, the basis used in Ref. 5 may be chosen. Eliminating the heavy branch valence-band envelope functions, the following set of coupled differential equations for the envelope wave functions of the Γ_6 $(f_{1,4})$ and light branch of Γ_8 $(f_{2,5})$ bands is obtained:[5]

$$f'_{1,4}=Af_{2,5}\pm\frac{k}{2}f_{1,4}$$

$$f'_{2,5}=Bf_{1,4}\mp\frac{k}{2}f_{2,5} \tag{1}$$

where

$$A=-(\tfrac{3}{2})^{1/2}\frac{E_+}{s\hbar},\quad B=(\tfrac{3}{2})^{1/2}\frac{1}{s\hbar}(E_- -\tfrac{3}{4}s^2\hbar^2k^2/E_+),$$

$$E_\pm=E-V(z)\pm ms^2,$$

$V(z)$ is the surface potential, z axis is directed to the bulk normal to the surface, k is 2D wave vector. The energies are measured upward from the middle of the gap in the bulk. The prime on the symbols hereafter means d/dz. The set (1) may be decoupled and reduced to the separated equations of the second order. There are two possibilities. One, used in many works beginning from Ref. 1, is to express the functions $f_{2,5}$ corresponding to Γ_8 band through Γ_6 band functions $f_{1,4}$. This leads to the equations

$$\left[\hat{k}_z^2+i\left(\frac{A'}{A}\right)\hat{k}_z+AB+\frac{k^2}{4}\pm\left(\frac{A'}{A}\right)\frac{k}{2}\right]f_{1,4}(k,z)=0, \tag{2}$$

which determine the subband spectra of 2D electrons of Γ_6 band (called hereafter Kane ''s electrons'') in the surface quantum well on Kane semiconductors with a positive gap. For the description of 2D carriers of a light branch of a Γ_8 band, we must use another way of eliminating the components corresponding to Γ_6 band. As a result, we obtain

$$\left[\hat{k}_z^2+i\left(\frac{B'}{B}\right)\hat{k}_z+AB+\frac{k^2}{4}\mp\left(\frac{B'}{B}\right)\frac{k}{2}\right]f_{2,5}(k,z)=0. \tag{3}$$

If we are interested in 2D electron subbands on semiconductors with $E_g<0$ (in the following ''p electrons''), we must simply make the inversion of the bands in Eq. (3) (changing the sign of energy and performing the charge-conjugation operation $V\rightarrow -V$).

It should be emphasized that unlike the ''squared'' Dirac equations, the ''squared'' Kane equations (2) and (3), for electrons and ''positrons'' (light holes), are not transformed one into another by charge-conjugation operation, i.e., are essentially nonequivalent. This is caused by the interaction with heavy-hole branch. The equivalency (symmetry) is restored only when we neglect spin effects described by the second and the fifth terms in Eqs. (2) and (3), i.e., in KG approximation. The nonequivalency of the equations for ''s-'' and ''p-electron'' subbands is retained in UR limits $E_g\equiv 0$ also, i.e., the spin effects for electron subbands are different for the cases $E_g=+0$ and $E_g=-0$. It must be noted that if subband states are described by Eq. (2) [bounded surface ''s electrons'' in semiconductor with $E_g>0$ in attractive potential $V(z)$], the continuum of bulk states obey Eq. (3) and vice versa in the case of surface ''p electrons'' at $E_g<0$.

The equations (2) and (3) contain the first derivative of the wave function and so have no simple physical interpretation. In the following, we use the approach based on the transformation of matrix equations into Schrodinger-like equation.[31,32] Such a line of attack, as applied to semiconductors, was used in Ref. 1 for ''s-'' subband calculations and in Ref. 21 for the demonstration of nonresonant character of 2D states in the surface layer on zero-gap semiconductor in Dirac approximation. Because the terms with $\hat{k}_z$ in (2) and (3) contain the logarithmic derivative, they can be eliminated by simple substitutions,

$$f_s^\pm=\sqrt{A}f_{1,4};\quad f_p^\pm=\sqrt{B}f_{2,5},$$

and subband effective-mass equations take the standard form

$$\left[\hat{k}_z^2-\frac{2ms^2}{s^2\hbar^2}(E^{\rm eff}-U_{s,p})\right]f_{s,p}^\pm=0, \tag{4}$$

with effective energy

$$E^{\rm eff}=\frac{E^2-m^2s^4-s^2\hbar^2k^2}{2ms^2}, \tag{5}$$

and with the effective potential [for every state $E(k)$, the effective energy $E^{\rm eff}$ and potential $U_{s,p}(E,z,k)$ are of their own]

$$U_{s,p}(E,z,k)\equiv U=U_0+U_{\rm so}+U_r. \tag{6}$$

We single out three parts in this effective potential: the ''Klein-Gordon'' term independent of ''spin,'' which is the same for ''s-'' and ''p-'' Kane electrons and for Dirac-like electrons (''D- electrons''),

$$U_0=(2VE-V^2)/2ms^2, \tag{7}$$

and spinlike terms: the spin-orbit term

$$U_{\rm so}=\pm\frac{s^2\hbar^2}{4ms^2}ga_{s,p}\frac{V'}{E_+}k, \tag{8}$$

and the ''resonant'' term describing ''spin-interband'' interaction, arising from the mixing of Γ_6 and Γ_8 bands by electric field [the second terms in Eqs. (2) and (3)],

$$U_r=\frac{s^2\hbar^2}{2ms^2}\left[\frac{3}{4}b_{s,p}\left(a_{s,p}\frac{V'}{E_+}\right)^2+\frac{1}{2}a_{s,p}\frac{V''}{E_+}\right]. \tag{9}$$

The term U_r describes the effect which is usually treated as Zener resonant tunneling in the surface electric field.[7,9,34] Distance-independent 2D kinetic term $s^2\hbar^2k^2/2ms^2$ in the expression for effective energy $E^{\rm eff}$ is an analog of the centrifugal potential in the corresponding spherically symmetric problem for Dirac electrons[31] and formally can be included in the effective potential.

The equations for ''s'' and ''p electrons'' differ by coefficients $a_{s,p}$ and $b_{s,p}$ in the expressions for the potentials U_r and $U_{\rm so}$. For ''p electrons,'' they are

$$a_p=\left(1+\frac{3s^2\hbar^2k^2}{4E_-^2}\right)\Big/\left(1-\frac{3s^2\hbar^2k^2}{4E_+E_-}\right), \tag{10}$$

$$b_p=\frac{1}{3}+\frac{2}{3}\left[\frac{1}{1+\dfrac{E_-}{E_+}}+\frac{1}{a_p-1}\right]^{-1}+\frac{2}{3a_p}\left[1+2\frac{E_+(a_p-1)}{E_-\left(a_p+\dfrac{E_-}{E_+}\right)}\right], \tag{11}$$

$g=1$, while for ''s electrons,'' $a_s=b_s=g=1$. The equation for Dirac electrons is just the same as for Kane ''s electrons,'' but with one exception, $g=2$. It is easily seen that for $ms^2 \gg E\text{-}V$, Eq. (4) is reduced into a Schrödinger equation, with the original energy E and the potential V.

Because of the singularity in ''spin potential'' U_r+U_{so} at the point $z=z_+$, corresponding to the condition $E_+(z=z_+)=0$, for ''s'' and ''D electrons'' (at $k=0$ for ''p electrons'' also), there is an infinitely high barrier in the effective potential U for the states $E(k)$ lying below the gap ($E<-ms^2$). In the equation for ''p electrons,'' the similar barrier at $k\neq 0$ arises from singularity in a_p at the point $z=z_a$, given by $V(z_a)=E-\sqrt{m^2s^4+3s^2\hbar^2k^2/4}$. The existence of such ''nonpenetrative'' infinite potential walls, separating the surface states from the volume states, dictates the boundary conditions $f_{s,p}(z_+)=0$ for envelope functions describing both bound and continuum states. On the other hand, this barrier ensures the stationary character of 2D states at any ms^2 and k although for $E<-ms^2$ they are in ''resonance'' with valence band states. For ''p'' subbands, they are in ''resonance'' with the heavy holes just for $E<+ms^2$. It should be noted, that even when the spin effects are ignored, almost all 2D states are stationary (or well-defined quasistationary states), owing to the barrier existing in ''Klein-Gordon'' potential U_0, as it is shown in Ref. 21.

The resonant term U_r causes the additional attraction to the surface, especially at small 2D wave vectors k when the influence of kinetic term $s^2\hbar^2k^2/2ms^2$ in E^{eff} is weak. This leads to the narrowing of the effective potential well and as a result to the forcing of the subband levels upward. For ''s'' and ''D electrons,'' the influence of U_r is the same. Because in the classically accessible region a_p, $b_p \geqslant 1$ (the factor b_p is close to 1), such an effect at $k\neq 0$ is more pronounced for ''p electrons.''

The spin-orbit splitting effect for ''p electrons'' is also expected to be more pronounced as compared with ''s electrons.'' If the spin-orbit term is small and may be treated as perturbation (this is so indeed), the spin-orbit splitting in the case of ''D electrons'' is to be twice as large as for Kane ''s electrons,'' because of the additional factor $g=2$ in U_{so}. The latter is in accord with the magnitudes of ''effective g factors'' for these two cases: the ratio of spin splitting to orbital splitting in magnetic field for ''D electrons'' is twice as large. Equation (4), for ''s,'' ''p,'' and ''D electrons,'' coincide only for $k=0$ (subband bottom) and in extreme ''ultrarelativistic'' subband limit ($k\rightarrow\infty$).

In the following, we shall confine our consideration by the semiclassical approximation both for the calculation of the surface potential $V(z)$ and for the quantization of Eq. (4). The direct integration of Eqs. (4) and WKB quantization lead to very similar results. However, the WKB approach permit obtain the analytical approximations (see below). The high accuracy of semiclassical quantization is caused by the following circumstance: as it is known, the appreciable error in WKB treatment is expected for the ground state, but the electrons in the lowest subband appear to be confined in the region where the potential is close to linear, when the semiclassical and exact treatment give practically the same results. The semiclassical self-consistent potential V is obtained in the frame of quasirelativistic modification of the Thomas-Fermi method.[19,35] This method seems to be a reasonable compromise between the accuracy and the ease of calculations. In the KG approach, this method leads to the results which are quite close to those obtained by direct self-consistent calculations.[5,6,19,35]

For the self-consistent potentials, it is convenient to modify the Bohr-Sommerfeld quantization rules passing to the integration over potential

$$\int_{V(z=0)}^{V(k_z=0)} k_z(E,k,V)\left(\frac{dV}{dz}\right)^{-1}dV=\pi(i+\tfrac{3}{4}), \tag{12}$$

because, in this case, we do not need the knowledge of explicit form $V(z)$. At the same time, the first Poisson integral (dV/dz) as a function of V is calculated by simple integration.

IV. PSEUDOULTRARELATIVISTIC APPROACH

Let us consider in this and the next sections more fully the subband spectrum and usually measured parameters (n_i, m_{ci}) in extreme limit $E_g\equiv 0$. On the one hand, the peculiarities due to quasirelativistic character of discussed systems (including the spin effects and their specific for ''s,'' ''p,'' and ''D electrons'') are most distinctly pronounced in the UR limit. On the other hand, this approach corresponds to the typical experimental situation at small gaps, as it was discussed above.

In the considered case, the bulk dispersion is of neutrino type $E=s\hbar k$ and the local bulk concentration as a function of chemical potential μ is $n(z)=(\mu(z)/s\hbar)^3/3\pi^2$. If we take into account only the charge of electrons in the Poisson equation (it is satisfied at bulk chemical potential $\mu_b=0$ and $T=0$), we obtain at given band bending $\mu_s-\mu_b$,

$$\frac{dV}{dz}=-\frac{d\mu}{dz}=\frac{\beta V^2}{2\pi s\hbar};\quad V(z)=-\mu(z)=-\frac{\mu_s}{(1+z/z_s)}, \tag{13}$$

where the characteristic length $z_s=2\pi s\hbar/\beta\mu_s$, $\beta=\sqrt{8\pi\alpha/3}\approx 1$ (Ref. 21), $\alpha=e^2/s\hbar\chi$ is a modified fine-structure constant, χ is a dielectric permittivity. The effective potential then takes the simple form

$$U_{so}=\pm\left(\frac{\beta}{2\pi}\right)\frac{s\hbar}{4ms^2}\,g a_{s,p}\,\frac{V^2}{E_+}\,k,$$

$$U_r=\left(\frac{\beta}{2\pi}\right)^2\frac{1}{2ms^2}\left[\frac{3}{4}\,b_{s,p}\left(a_{s,p}\,\frac{V^2}{E_+}\right)^2+a_{s,p}\,\frac{V^3}{E_+}\right],$$

$$b_{s,p}=1+\frac{4}{3}\,\frac{(1-a_p^{-1})}{(1+a_p)}\approx 1,$$

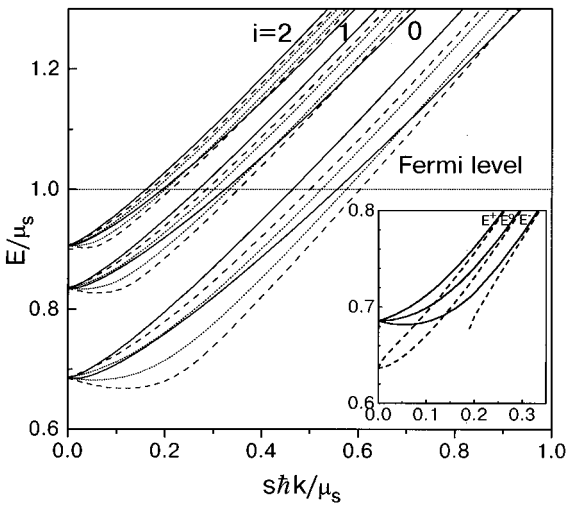

FIG. 4. Dispersion relations in dimensionless units for Kane ''*p* electrons'' (full curves), Kane ''*s* electrons'' (dotted curves), and ''Dirac electrons'' (broken curves) in the UR limit $E_g=0$. The inset shows dispersions for the states near the ground subband bottom for Kane ''*s* electrons,'' with (full curves) and without (broken curves) the resonant term. The curves $E^0(k)$ are calculated without the U_{so} term.

[Eq. (4) behaves nonsingular for $ms^2=0$, because of a multiplier $2ms^2/s^2\hbar^2$ at the term $(E^{\text{eff}}-U)$]. Because the scale invariance of the PU approach, with respect to μ_s, remains valid also when the spin effects are taken into account, the subband dispersions are described by the universal relationship between dimensionless (normalized to μ_s) energy E/μ_s and momentum $s\hbar k/\mu_s$. The results for the first three subbands are plotted in Fig. 4 (in the UR limit, there is an infinite number of filled 2D subbands). The spin-orbit splitting $\Delta E_{so}=E^+-E^-$ for ''*s*'' and ''*p* electrons'' at small k is linear on k [see Fig. 5(a)] as in the phenomenological Rashba model. However, with increasing k, the calculated $\Delta E_{so}(k)$ dependencies deviate sufficiently from linear ones, peaking already at $k\leqslant k_F/2$. For the ''*p* electrons,'' the linear portion of $\Delta E_{so}(k)$ curves is, in fact, absent and the maximum of ΔE_{so} takes place at $k\sim 3k_F$. In the region of the large wave vectors, the splitting is slowly decreasing with increasing k for all three cases ($\Delta E_{so}\to 0$ at $k\to\infty$). The latter fact shows that not only the linear approach

$$E_i^{\pm}(k)-E_i^{\pm}(0)=\frac{\hbar^2k^2}{2m_i}\pm d_ik \tag{14}$$

(m_i-subband-edge mass, α_i-Rashba parameter), but the Dirac-like approximation[30]

$$E_i^{\pm}(k)-E_i^{\pm}(0)=\sqrt{s^2\hbar^2k^2+m_i^2s^4\pm 2m_is^2\alpha_ik}-m_is^2 \tag{15}$$

predicting the saturation in ΔE_{so} at large k, are not valid for the description of spin-split subband dispersions. If we are going to consider Rashba parameter α_i as depending on the wave vectors, its value on Fermi level for ''*s*'' and ''*D* electrons'' (determined in Ref. 30) is nearly twice as much as that at small k, when the linear approximation holds. Spin-splitting for ''*D* electrons'' at any k is indeed twice as large

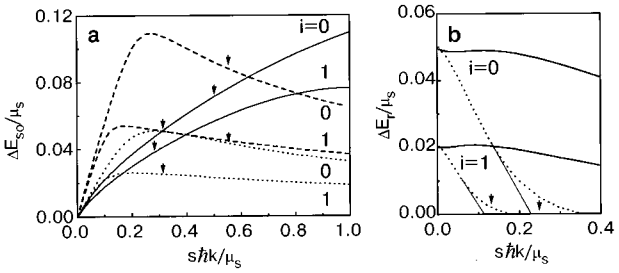

FIG. 5. Dimensionless spin-orbit splitting (a) and resonant shift at $U_{so}=0$ (b), as functions of the dimensionless 2D wave vector for Kane ''*p*'' (full curves), Kane ''*s*'' (dotted curves), and ''Dirac electrons'' (broken curves) in the UR limit. The arrows on (a) show Fermi wave vectors k_{Fi}^0, calculated at $U_{so}=0$. The arrows in (b) show in the KG approach, critical wave vectors k_{nr} (Ref. 21), beginning from which the Zener tunneling is impossible, because of the nonconservation of transverse quasimomentum.

as for ''*s* - electrons [see Fig. 5(a)]. For ''*p* electrons,'' ΔE_{so} for the states near the Fermi level is close to ΔE_{so} for Dirac electrons. Thus, the spin-orbit effects are more pronounced in the Kane semiconductors with inverted bands in the line with the experimental observations of the Sec. II.

As to resonant effects, the term U_r for ''*s*'' and ''*D* electrons'' has but a minor effect at large k. However, when k is decreasing, the pole in U_r at z_+ (for states with $E<0$) is moved to the surface, whereas the turning point z_i is moved from the surface because of mixing of motion in confinement direction with that in 2D plane (it is clear that in any case z_i is nearer to the surface than z_+). In PU limit for all subbands, the spatial dimension of semiclassical wave functions in the z direction for $k=0$ is approximately twice as large as that for $k=k_F$ (this phenomena may be called ''quasirelativistic confinement''). As a result, z_+ is brought to z_i and the positive corrections to the energy from resonant term U_r increase. The pole at z_+ corresponds to the condition $V(z_+)=E$, whereas z_i for small k is determined from the approximate condition $V(z_i)\approx E-s\hbar k$ (KG approach). At $k\to 0$, the turning point z_i is maximally close to z_+ and the energy shift reaches maximum. For the important case of the states on the Fermi level $E=V(z_+)=0$ and pole z_+ in the spin potential goes to infinity, so the influence of U_r is weak for ''*s*'' and ''*D* electrons'' [see Fig. 6(a)]. For ''*p* electrons,'' the distance between z_i and the pole at z_a [$V(z_a)=E-\sqrt{3}s\hbar k/2$] is almost independent of k from $k=0$ to $k\approx 3k_F$ ($z_i-z_a\approx 0.6z_s$ for $i=0$) and so at large k is essentially less than for ''*s* electrons.'' The contribution of spin terms to U for states on Fermi level is almost of the same importance as for subband bottom [see Fig. 6(b)]. The distance z_i-z_a rapidly increases beginning with $k>4k_F$ and only at $k\approx 10k_F$ the pole z_a goes to infinity.

The corrections to the energy $\Delta E_r=E^0-E^{00}$, due to resonant term (here E^0 and E^{00} are the subband energies calculated at $U_{so}=0$ and $U_{so}=U_r=0$, respectively), are plotted on Fig. 5(b) as the functions of 2D wave vector k. For ''*s*'' and ''*D* electrons,'' k values, starting from which the contribution from U_r becomes small, correlate with critical wave vectors k_{nr} determined in the KG approach in Ref. 21, starting from which Zener tunneling is impossible, because of the nonconservation of transverse quasimomentum. The elec-

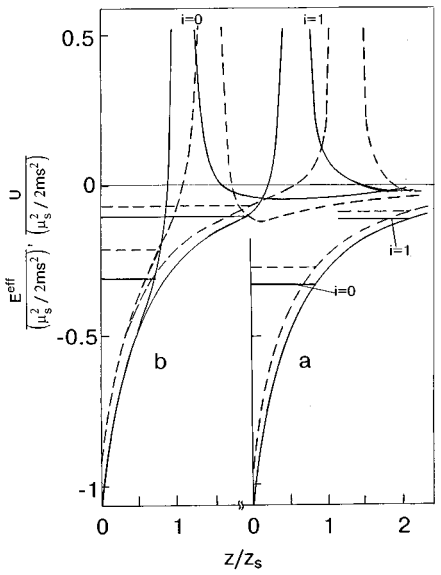

FIG. 6. Dimensionless effective potential $U(z)$ and effective energies $E^{\rm eff}$ (horizontal lines) for the states on Fermi level in the UR approach. (a) is for Kane ''*s* electrons;'' (b) for Kane ''*p* electrons.'' The full curves are for E^- branch, broken curves are for E^+. For ''*s* electrons,'' the effective potentials for the $i=0$ and 1 subband are close.

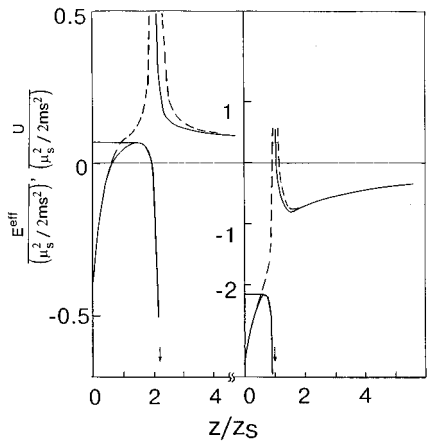

FIG. 7. Dimensionless effective potential $U(z)$ and effective energies $E^{\rm eff}$ (horizontal lines) of ground subband for E^- branch near its cutoff [occurring when $U_r(z)=0$], calculated in the UR limit with (broken curves) and without (full curves) the resonant term. (a) is for Kane ''*s* electrons'' ($s\hbar k_t/\mu_s=0.192$, $E/\mu_s=0.681$); (b) for Kane ''*p* electrons'' ($s\hbar k_t/\mu_s=1.832$, $E/\mu_s=2.086$).

trons in excited subbands are localized in a more wide space region and averaged electric field and the potential U_r which they see are less. As a result, the resonant shift ΔE_r decreases when the subband index i is decreasing, i.e., ''spin-interband'' interaction diminishes the interband energies $E_i(k)-E_j(k)$ at small k.

The resonant term has a more drastic impact on the low-energy dispersion branch E^-. The dispersions for Kane ''*s* electrons,'' with and without an allowance for the resonant term, are given in the insert in Fig. 4 for the energy near the ground subband bottom. As it may be seen, the branch E^- terminates on the side of small enough $k<k_t$, if we ignore resonant term U_r (there exists a formal similarity between this phenomena and the motional binding effect[26,27]). Such a behavior may be understood if we consider the effective potential form at $U_r=0$. Because of the negative singularity for the E^- branch in the spin-orbit potential $U_{\rm so}$ at $E_+(z_+)=+0$ in the effective potential $U_0+U_{\rm so}$, there is an infinitely deep well at $z\leqslant z_+-0$ and infinitely high barrier at $z\geqslant z_++0$. The surface quantum well near $z=0$ is separated from that centered at the $z=z_+$ infinite well by the barrier [see Fig. 7(a)]. At large enough k, the energy level $E_{\rm eff}$ lies below the barrier top and the subband states exist. As k is decreasing the level $E_{\rm eff}$ shifts upwards and at some critical values k_t (for ground subband $s\hbar k_t/\mu_s=0.19$ and 0.29 for ''*s*'' and ''*D* electrons,'' respectively) becomes tangent to the top of the barrier. In the semiclassical picture, this situation corresponds to the simultaneous reduction in both $k(z)$ and $dk(z)/dz$ to zero. At smaller values of k the levels $E^{\rm eff}$ fall into infinite well centered at $z=z_+$ (in some sense, a peculiar kind drop to the center). The total number of the states remains the same as in the E^+ branch. For the E^+ branch, the opposite situation is realized; there is infinitely high barrier at $z\leqslant z_+-0$ and well at $z\geqslant z_++0$, so the surface states are separated from the well at any k. For ''*p* electrons,'' similar behavior [see Fig. 7(b)] is due to the pole in a_p (for ground subband $s\hbar k_t/\mu_s=1.84$). For ''*s*'' and ''*D* electrons,'' the cutoff wave vectors k_t are less than Fermi wave vectors k_F For ''*p* electrons,'' the E^- branch appears only at $k\gg k_F$. The omitting of the resonant term U_r in Eq. (4) [or of the second terms in Eqs. (2) and (3)] leads rather to the imaginary than real effects. The ''spin-interband'' interaction described by this term prevents the formation of ''nonphysical'' potential well at $z=z_+$ and hence eliminates the cutoff in E^- dispersion branch.

Taking into account U_r is important yet in another respect. In the KG approach, $U_r=U_{\rm so}=0$ the subband dispersions at small k obey the subquadratic relationship $E(k)-E(0)\sim k^\gamma$ ($\gamma<2$) and cannot be approximated by (15) independent of the k subband-edge effective masses m_i (at the KG limit $\alpha_i=0$). At the same time, an allowance for U_r the $E(k)-E(0)$ dependencies for ''*s*'' and ''*D* electrons'' at $U_{\rm so}=0$ are quadratic at small k (see Fig. 8) and the approach (15) with m_i determined in this region describes well the subband dispersions in wide regions of the wave vectors from the subband bottom to the Fermi level. The dispersions for ''*p* electrons,'' at small k remain close to those in the KG approach.

From the standpoint of correlation with experiment, the Fermi wave vectors $k_{Fi}^{\pm}$ are of particular interest, because they determine both the fundamental periods of oscillations in magnetic fields $\Delta_i=(\Delta_i^+ +\Delta_i^-)/2$, $\Delta_i^{\pm}=2e/c\hbar(k_{Fi}^{\pm})^2$ and the beating periods $\Delta_{bi}=(\Delta_i^+ -\Delta_i^-)/2$. Because of the scale invariance of the UR limit, $k_{Fi}^{\pm}$ is linear on μ_s

$$k_{Fi}^{\pm}=\frac{\mu_s c_{ki}^{\pm}}{s\hbar}$$

with the scale coefficients $c_{ki}^{\pm}$ presented for ''*s*,'' ''*p*,'' and ''*D* electrons'' in Table I together with the values c_{ki}^0 calculated neglecting spin effects.[21] For the same μ_s, the Fermi wave vectors $k_{Fi}^{\pm}$ and consequently total subband occupations $n_i=n_i^+ +n_i^-$ ($n_i^{\pm}=(k_{Fi}^{\pm})^2/4\pi$) for ''*p* electrons'' are

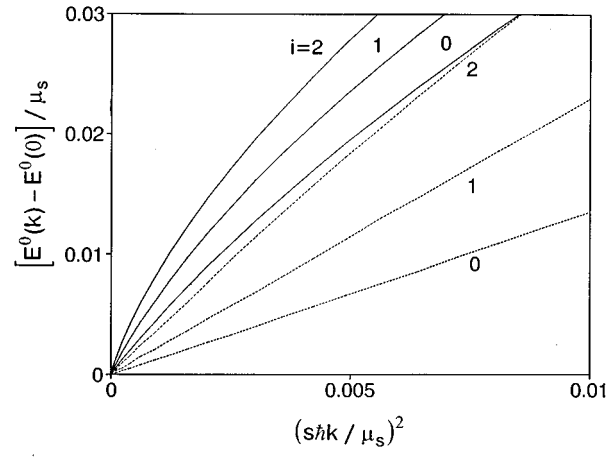

FIG. 8. UR dispersion relation in dimensionless units for Kane ''*p*'' (full curves) and Kane ''*s*'' subbands (dotted curves) near subband bottoms calculated at $U_{so}=0$.

less than for ''*s*'' and ''*D* electrons.'' Nevertheless, the relationship between $k_{Fi}^{\pm}$ for different subbands is the same for all three cases. As a result, the subband occupations n_i are equal for ''*s*,'' ''*p*,'' and ''*D*'' subbands at one and the same total surface density $n_s=\Sigma n_i$. The scale coefficients $c_{mi}^{\pm}$, which give the relation between cyclotron mass and surface potential $m_{ci}^{\pm}s^2=c_{mi}^{\pm}\mu_s$ are also different for ''*s*,'' and ''*p*,'' and ''*D* electrons'' (Table I). However, for the same subband occupations, the ''average'' cyclotron masses for ''*s* electrons'' are close to those for ''*D* electrons'' and are 5–6 % less than for ''*p* electrons.'' The latter difference, however, is, in fact, within the experimental error of m_{ci} measurements based on the analysis of the temperature dependence of the oscillation amplitudes. Hence, both the electron distribution on subbands and the average cyclotron masses as the functions of n_s are described practically by the same as that for ''*s*-,'' ''*p*-,'' and ''*D*-electrons'' universal dependencies, which coincide with the dependencies deduced from the KG approach. This is in agreement with the conclusion of Ref. 21.

V. ANALYTICAL APPROXIMATIONS

The WKB method is remarkable for our purposes, because the simple analytical expressions for subband parameters, including the parameters characterizing spin-orbit splitting, for ''*s*'' and ''*D* electrons'' may be derived. The quasiclassical expression for the *z* component of wave vectors in the UR limit for states on the Fermi level takes the simple form

$$k_z^{\pm}=\frac{[V^2(1+\eta_r)-s^2\hbar^2k^2\pm\eta_s s\hbar kV]^{1/2}}{s\hbar}, \qquad (16)$$

where the constant $\eta_r=(\beta/4\pi)^2$ arises from the resonant term and $\eta_s=g\beta/4\pi$ from the spin-orbit term in (4). The left-hand side in Eq. (12), in this case, is exactly integrable and we obtain the following transcendental equation for the scale coefficients $c_{ki}^{\pm}$,

$$-\sqrt{1+\eta_r\mp\eta_s c_k-c_k^2}+\sqrt{1+\eta_r}\,\ln\left|\frac{2\sqrt{1+\eta_r\mp\eta_s c_k-c_k^2}+2(1+\eta_r)\mp\eta_s c_k}{c_k\sqrt{4(1+\eta_r)+\eta_s^2}}\right| \mp\frac{\eta_s}{4}\left[\pi-2\arcsin\frac{2c_k\pm\eta_s}{\sqrt{4(1+\eta_r)+\eta_s^2}}\right]=\frac{\beta}{2}\left(i+\frac{3}{4}\right). \qquad (17)$$

At $\eta_r=\eta_s=0$, Eq. (17) is reduced to Eq. (4) from Ref. 21, which has been obtained without an allowance for spin effects. Because of $\eta_r\ll 1$, taking resonant effects into account leads to very small correction on $c_{ki}^{\pm}(<1\%)$.

The Fermi wave vectors $k_{Fi}^{\pm}$, ''split'' by the spin-orbit interaction, may be found also as the corrections to the magnitudes calculated in the KG limit $\eta_r=\eta_s=0$. Using small parameter $(\eta_s c_k)$ for Taylor series expansion of the terms in the left-hand side of Eq. (17), it can be shown that such a correction is reduced to the changing of factor $\frac{3}{4}$ in the right-hand side of Eq. (17), written for the unperturbated problem

TABLE I. The scale coefficients c_{ki} and c_{mi} giving the relations, respectively, between subband Fermi wave vector $k_{Fi}=c_{ki}\mu_s/s\hbar$ and surface potential μ_s and between subband cyclotron mass $m_{ci}=c_{mi}\mu_s/s^2$ and μ_s in pseudoultrarelativistic limit $ms^2=0$. $\Delta n_i/n_i=(n_i^- - n_i^+)/(n_i^- + n_i^+)$ is the relative difference of the occupations of spin-split subbands.

	''*s* electrons''			''*p* electrons''			''*D* electrons''			KG approach
i	c_{ki}^+	c_{ki}^-	$\Delta n_i/n_i$	c_{ki}^+	c_{ki}^-	$\Delta n_i/n_i$	c_{ki}^+	c_{ki}^-	$\Delta n_i/n_i$	c_{ki}^0
0	0.524	0.575	0.094	0.464	0.556	0.180	0.500	0.604	0.186	0.547
1	0.300	0.333	0.104	0.268	0.323	0.186	0.285	0.352	0.206	0.315
2	0.179	0.200	0.111	0.160	0.194	0.191	0.169	0.212	0.220	0.188
3	0.108	0.121	0.113	0.097	0.118	0.195	0.102	0.129	0.228	0.113
4	0.065	0.073	0.115	0.059	0.071	0.197	0.062	0.078	0.234	0.068
	c_{mi}^+	c_{mi}^-		c_{mi}^+	c_{mi}^-		c_{mi}^+	c_{mi}^-		c_{mi}^0
0	0.638	0.664		0.567	0.745		0.621	0.670		0.648
1	0.403	0.431		0.357	0.474		0.388	0.443		0.414
2	0.255	0.278		0.225	0.302		0.243	0.289		0.264
3	0.160	0.176		0.141	0.190		0.151	0.184		0.166
4	0.099	0.110		0.087	0.119		0.093	0.116		0.103

($\eta_r=\eta_s=0$) to the factor $\frac{3}{4}\mp g[1/8-c^0_{ki}/4\pi]\approx\frac{3}{4}\mp g/8$. For scale coefficients $c^{\pm}_{ki}$, the simplest approximation can be used

$$c^{\pm}_{ki}\approx c^0_{ki}\exp\mp\frac{g\beta}{16}\left(1-\frac{2c^0_{ki}}{\pi}\right), \tag{18}$$

where

$$c^0_{ki}=c_0\left(\exp-\frac{\beta i}{2}+\frac{1}{3}\exp\frac{3\beta i}{2}\right),$$

$$c_0=2\exp-\left(1+\frac{3\beta}{8}\right)\approx 0.50 \tag{19}$$

are the scale coefficients in the KG limit.

The values $\Delta n_i/n_i$ characterizing the spin-orbit splitting and determining the experimentally measured beating periods are given by

$$\frac{\Delta n_i}{n_i}\approx\frac{g\beta}{8}\left(1-\frac{2c^0_{ki}}{\pi}\right)\approx\frac{g\beta}{8}. \tag{20}$$

It is easy to verify that the values given by the expressions (18) and (20) are in very good agreement with magnitudes presented in Table I.

VI. DISCUSSION AND COMPARISON

The general peculiarities of the above considered ''ultrarelativistic'' spectrum remain valid for real 2D systems based on narrow-gap semiconductors. However, at small surface density n_s, the quantitative results may differ. In this case, the surface potential μ_s is close to or smaller than E_g [this corresponds to subband occupations $n_i<(E_g c^0_{ki})^2/2\pi s^2\hbar^2$] and (or) the contribution of depletion layer charge or electrons of continuum to the surface potential is large (at high doping level).[19]

As to the influence of doping level on the spin effects, it is essentially different for the inversion channels and accumulation layers. In a former case, the additional electric field, due to the charge in depletion layer, magnifies the contribution of spin-dependent terms to the effective potential. This leads to the increase of both the spin-orbit splitting and resonant shift of the levels. Under the usual treatment, the latter effect is interpreted as the increasing of the tunnel exchange between 2D and volume states, owing to the thinning of the depletion layer at an increasing doping level. Another specific feature of narrow-gap semiconductors is exhibited in the electron accumulation layers. The electron gas in the bulk in such semiconductors of n type at any donor concentrations N_D is degenerated, due to the smallness of the critical for Mott transition donor concentration. The contribution of the degenerated electron gas of continuum to the screening of surface electric field leads to such specific effects as the ''two-dimensionalization'' of electrons at zero surface electric fields[22–24] and motional binding.[26,27] Unlike the case of the inversion layers, the spin effects in the accumulation layers are decreasing at decreasing subband occupations, especially in the motional binding regime. However the discussion of these results is beyond the scope of the present work.

In contrast to the influence of the doping level, the increase of gap E_g results in the weakening of spin effects both in inversion and accumulation layers. This is evident from the form of effective potential $U(z)$ because, with increasing E_g, the values E_+ and a_p^{-1} increase and the contribution of spin like potentials U_{so} and U_r to U decreases (the pole in U at z_+ is rapidly shifted inside the semiconductor for states, which are in resonance with the states of continuum). The relative difference of occupations $\Delta n_i/n_i$ is plotted in Fig. 9 as a function of ms^2/μ_s for $\mu_b=0$ (i.e., for accumulation layers on ''nondoped n-type semiconductors''). For ''p electrons,'' $\Delta n_i/n_i$ is less sensitive to the changes of the gap than for ''s'' and ''D electrons.'' Even at $ms^2\approx\mu_s/2$, the changes in $\Delta n_i/n_i$ for the ground subband of ''p electrons'' are less than 15% from those obtained in the UR limit. A strong increase with the subband index i is observed for all the three cases.

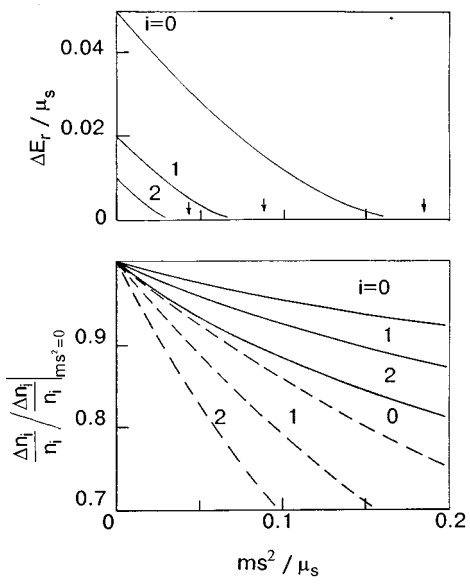

FIG. 9. $(\Delta n_i/n_i)/(\Delta n_i/n_i)|_{ms^2=0}$ and dimensionless resonant shift $\Delta E_r/\mu_s$ of subband bottoms ($k=0$) for accumulation layers at $\mu_b=0$ and $T=0$, as functions of normalized gap ms^2/μ_s. Full curves, for Kane ''p electrons;'' broken curves, for Kane ''s electrons'' [$\Delta E_r(k=0)$ for ''s,'' ''p,'' and ''D electrons'' are the same]. The arrows mark the values of ms^2/μ_s at which the subbands bottom fall outside the resonance with continuum states.

In the framework of the WKB method, we can find the corrections to the UR limit (for ''s'' and ''D electrons''), resulting from nonzero E_g. In the first order on ms^2/μ_s, the expression for Fermi wave vector is transformed to the form

$$k^{\pm}_{Fi}=c^{\pm}_{ki}(ms^2)(\mu_s+ms^2)/s\hbar, \tag{21}$$

where the scale coefficients $c^{\pm}_{ki}(ms^2)$ may be expressed by their magnitudes calculated at $ms^2=0$ [Eqs. (17) and (18)],

$$c^{\pm}_{ki}(ms^2)=c^{\pm}_{ki}|_{ms^2=0}\exp\pm\frac{g\beta}{16\pi}\left(\frac{2}{c^0_{ki}}-c^0_{ki}\right)\frac{ms^2}{\mu_s}. \tag{22}$$

In the same linear on the ms^2/μ_s approach, the expression for $\Delta n_i/n_i$ has the form

$$\frac{\Delta n_i}{n_i}=\frac{\Delta n_i}{n_i}\bigg|_{ms^2=0}\left[1-\frac{(2-(c^0_{ki})^2)}{c^0_{ki}(\pi-2c^0_{ki})}\frac{ms^2}{\mu_s}\right]. \tag{23}$$

This gives the correct description of the linear on the ms^2/μ_s portion of the curves in Fig. 9 for ''s'' and ''D electrons.'' The dependencies as obtained from (23) so presented on Fig. 9 characterize the changes of $\Delta n_i/n_i$ with ms^2 at the same band bending μ_s. But it is easy to understand that these dependencies remain valid in the case of the same subband occupations, which can be experimentally controlled. Indeed, according to (21), the different band bendings correspond to the same subband occupations n_i in the materials with different ms^2 [$\mu_s(ms^2)=\mu_s(0)-ms^2$]. However, the arising correction to (23) is of the second order, because $ms^2/(\mu_s-ms^2)\approx ms^2/\mu_s-(ms^2/\mu_s)^2$. It should be noted that without an allowance for spin-orbit splitting [$g=0$ in (22)], the corrections to the scale coefficients are zero up to and including the second order on ms^2/μ_s. As a result, the distribution of the electrons on the subbands is almost insensitive to ms^2. This explains the adequacy of the UR approach $ms^2=0$ for the description of $n_i(n_s)$ dependencies even at $ms^2\approx\mu_s$. For such a small band bending the corrections caused by the doping level are more pronounced.

The effect described by the resonant term U_r exhibits essentially more strong variations with E_g. This coincides with the result of the usual treatment. The latter associates the decreasing of ΔE_r with drastic reductions of the probability of interband tunneling caused by the increasing of both the height ($\sim E_g$) of ''classical'' barrier for interband tunneling and its width ($\sim E_g/eF$, where F is the strength of the electric field) at increasing E_g. The shift ΔE_r as a function of $|E_g|$ is shown in Fig. 9 for the states at subband bottom, where the ''resonant'' effect is the most pronounced and the same for ''s,'' ''p,'' and ''D electrons.'' The arrows in Fig. 9 mark the values of E_g at which the subband bottoms $E_i(0)$ are brought into the energy gap between Γ_6 and Γ_8 bands. As seen from Fig. 9, the shifts ΔE_r, due to U_r, go to zero just at magnitudes E_g when subbands fall outside the resonance with continuum states. This justifies the above interpretation of the term U_r in the effective potential as responsible for the Zener mixing. It must be stressed that the U_r term at the same time forbids the resonant broadening of the levels, in contrast to the usual treatment (Zener tunneling). Owing to the infinitely high barrier (see above) induced by this term, the subband surface states are not resonant in an ordinary sense, even if they coincide in energy with continuum states. In this sense, the term U_r is more properly called the ''antiresonance term.'' It appears that the resonant tunneling is strictly forbidden, owing to the existence of some conserved quantity of helicity type as for Dirac equation.

The calculated subband occupations for parameters of $H9$ and $K11$ samples using Eqs. (4) and (12) are plotted in Fig. 3, together with experimental data. Good agreement is observed for the sample with both direct and inverted band structures practically in the whole investigated range of $n_s<10^{13}$ cm^{-2}. In the scale of Fig. 3, the calculated dependencies $n_i^{\pm}(n_s)$ for both samples almost coincides with the UR results. The differences are detectable in the $\Delta n_i/n_i(n_i)$ dependencies presented in Fig. 10. At high n_i, the calculations for sample with $E_g<0$ are close to those in the UR limit whereas, in the case of $E_g>0$, they give the somewhat lesser values of $\Delta n_i/n_i$, because ''s'' subbands are more sensitive to the gap magnitude. On the other hand, for the H9 sample at $n_i<2\times10^{12}$ cm^{-2}, $\Delta n_i/n_i$ is larger as compared with UR values, because of the contribution of the depletion field. For the $K11$ sample with the smallest doping level, such deviations are noticeable only in the range of $n_i<3\times10^{11}$ cm^{-2}.

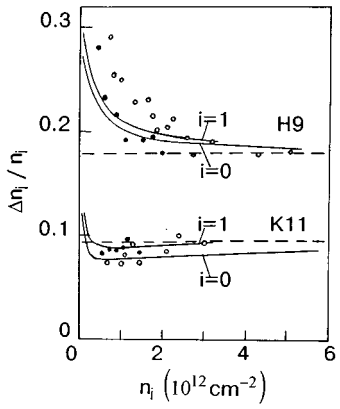

FIG. 10. Calculated (curves) and measured (points) dependencies of $\Delta n_i/n_i$ on subband occupations n_i, for $H9$ and $K11$ samples. Full curves are the theory. Broken lines are UR approximations for $i=0$. The experimental results: ●—$i=0$; ○—$i=1$.

The experimental data also show the increase of $\Delta n_i/n_i$ with decreasing n_i, but for the $H9$ sample, the increase falls in the region of the larger occupations than is predicted by the theory. Such a discrepancy cannot be caused by the doping nonuniformity, because the relations between the occupations of different subbands n_i/n_j, which are very sensitive to doping, are in accordance with Hall measurement of N_A-N_D. In addition, the above data are repeatable on the different MOS capacitors. The discrepancy may be partly attributed to several approaches used in calculations, such as the ignoring of the remote bands, the use of vanishing boundary conditions for the envelope functions at $z=0$ and WKB approximation. As for the spin-off band Γ_7 its influence at small gap may be detectable only at large n_s. However, in this region, the accordance between experiment and theory is just observed. For small n_i, there is the possibility of comparing our calculations for ''s'' subbands with the self-consistent calculations, using different boundary conditions, as it is done in Ref. 9 for the inversion layer on $Hg_xCd_{1-x}Te$, with $N_A=3\times10^{17}$ cm^{-3}. We can determine the Fermi wave vectors from the presented in this work dispersion, for two spin branches of ground subband $E_0^{\pm}(k)$ at $n_0=2\times10^{11}$ cm^{-2}. It gives the values $\Delta n_0/n_0\approx0.25$ at zero boundary conditions and $\Delta n_0/n_0\approx0.31$ at Marques and Sham boundary conditions.[3] For the same parameters, the treatment in the framework of the method used here gives $\Delta n_0/n_0\approx0.29$. Hence, both methods lead to the close results even for small subband occupation and large enough acceptor concentration.

The discrepancy between theory and experiment at small n_i in Fig. 10 can be caused also by high experimental error in the determination of $\Delta n_i/n_i$ at strong spin-orbit splitting. In this case, Δn_i and $n_i^{\pm}$ are of the same order of magnitude and the beating periods are close to the fundamental periods. Together with the small number of the oscillations observed at low n_i, this causes a large error in both Δn_i and n_i determined from Fourier spectra. Furthermore, at small occupations n_i only low Landau levels are observed in oscillations.

The periodicity of the oscillations on the reciprocal magnetic field and the cosine form of the state density in the magnetic field may be broken under these conditions. This also deteriorates the reliability of the data extracted from Fourier transforms.

ACKNOWLEDGMENTS

This work was partially supported by a grant from the High Education Committee of Russian Federation and by the program ''Universities of Russia.''

[1] F. Ohkawa and Y. Uemura, J. Phys. Soc. Jpn. **37**, 1325 (1974).
[2] Y. Takada, K. Arai, and Y. Uemura, *Lecture Notes in Physics*, Proceedings of the Fourth International Conference on Physics of Narrow Gap Semiconductors, Linz, Austria, 1981 (Springer-Verlag, Berlin, 1982), Vol. 152, p. 101.
[3] G. E. Marques and L. J. Sham, Surf. Sci. **113**, 131 (1982).
[4] F. Malcher, I. Nachev, A. Ziegler, and U. Rossler, Z. Phys. B **68**, 437 (1987).
[5] I. Nachev, Phys. Scr. **37**, 825 (1988).
[6] J. P. Zollner, G. Paasch, G. Gobsch, and H. Ubensee, Phys. Status Solidi B **148**, 611 (1988).
[7] A. Ziegler and U. Rossler, Europhys. Lett. **8**, 543 (1989).
[8] P. Sobkowicz, Acta Phys. Pol. A **75**, 29 (1989).
[9] P. Sobkowicz, Semicond. Sci. Technol. **5**, 183 (1990).
[10] I. Nachev, Semicond. Sci. Technol. **5**, 69 (1990).
[11] B. Freytag, U. Rossler, and O. Pankratov, Semicond. Sci. Technol. **8**, S243 (1993).
[12] T. I. Deryabina, L. P. Zverev, and V. F. Radantsev, Fiz. Tekh. Poluprovodn. **17**, 2065 (1983) [Sov. Phys. Semicond. **17**, 1319 (1983)].
[13] W. Q. Zhao, F. Koch, J. Ziegler, and H. Maier, Phys. Rev. B **31**, 2416 (1985).
[14] V. F. Radantsev, T. I. Deryabina, L. P. Zverev, G. I. Kulaev, and S. S. Khomutova, Zh. Eksp. Teor. Fiz. **88**, 2088 (1985) [Sov. Phys. JETP **61**, 1234 (1985)].
[15] W. P. Kirk, P. S. Kobiela, R. A. Shiebel, and M. A. Reed, J. Vac. Sci. Technol. A **4**, 2132 (1986).
[16] J. Singleton, F. Nasir, and R. J. Nicholas, J. Phys. C **19**, 35 (1986).
[17] V. F. Radantsev, Pis'ma Zh. Eksp. Teor. Fiz. **46**, 157 (1987) [JETP Lett. **46**, 197 (1987)].
[18] F. Nasir, J. Singleton, and R. J. Nicholas, Semicond. Sci. Technol. **3**, 654 (1988).
[19] V. F. Radantsev, Fiz. Tekh. Poluprovodn. **22**, 1796 (1988) [Sov. Phys. Semicond. **22**, 1136 (1988)].
[20] N. J. Bassom and R. J. Nicholas, Semicond. Sci. Technol. **7**, 810 (1992).
[21] V. F. Radantsev, Semicond. Sci. Technol. **8**, 394 (1993).
[22] O. V. Konstantinov and A. Ya. Shik, Zh. Eksp. Teor. Fiz. **58**, 1662 (1970) [Sov. Phys. JETP **31**, 891 (1970)].
[23] G. A. Baraff and J. A. Appelbaum, Phys. Rev. B **5**, 475 (1972).
[24] An-zhen Zhang, J. Slinkman, and R. E. Doezema, Phys. Rev. B **44**, 10 752 (1991).
[25] W. Zawadzki, S. Klahn, and U. Merkt, Phys. Rev. B **33**, 6916 (1986).
[26] R. E. Doezema and H. D. Drew, Phys. Rev. Lett. **57**, 762 (1986).
[27] M. Kubisa and W. Zawadzki, Semicond. Sci. Technol. **8**, S246 (1993).
[28] Yu. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).
[29] R. Wollrab, R. Sizmann, F. Koch, J. Ziegler, and H. Maier, Semicond. Sci. Technol. **4**, 491 (1989).
[30] V. F. Radantsev, Zh. Eksp. Teor. Fiz. **96**, 1793 (1989) [Sov. Phys. JETP **69**, 1012 (1989)].
[31] Ya. B. Zel'dovich and V. S. Popov, Usp. Fiz. Nauk **105**, 403 (1971) [Sov. Phys. Usp. **14**, 673 (1972)].
[32] A. B. Migdal, *Fermions and Bosons in Strong Fields* (Moscow, Nauka, 1978); A. B. Migdal, V. S. Popov, and D. N. Voskresenskii, Zh. Eksp. Teor. Fiz. **72**, 834 (1977) [Sov. Phys. JETP **45**, 436 (1977)].
[33] J. Luo, H. Munekata, F. F. Fang, and P. J. Stiles, Phys. Rev. B **41**, 7685 (1990).
[34] W. Brenig and H. Kasai, Z. Phys. B **54**, 191 (1984).
[35] T. Ando, J. Phys. Soc. Jpn. **54**, 2676 (1985).