Plasmon and magnetoplasmon excitation in two-dimensional electron space-charge layers on GaAs

E. Batke
AT&T Bell Laboratories, Murray Hill, New Jersey 07974-2070
D. Heitmann
Max-Planck-Institut für Festkörperforschung, D-7000, Stuttgart 80, Federal Republic of Germany
C. W. Tu
AT&T Bell Laboratories, Murray Hill, New Jersey 07974-2070
(Received 2 June 1986)

The excitation of two-dimensional (2D) plasmons in $Al_xGa_{1-x}As$-GaAs heterostructures has been investigated with far-infrared transmission spectroscopy. Grating couplers of high efficiency and samples of high mobility ($\mu \approx 1.8\times10^5$ cm^2/V s) make it possible to study the 2D plasmon dispersion up to plasmon wave vectors $q \gtrsim 2\times10^5$ cm^{-1}. Thus we can test the theoretically predicted second-order q corrections to the dispersion due to nonlocal effects, finite thickness of the space-charge layer, and correlations. In perpendicular magnetic fields, a strong nonlocal interaction of the 2D magnetoplasmon resonance with harmonics of the cyclotron resonance $n\omega_c$ ($n=2,3,4,\ldots$) is observed. Excellent agreement is found with predictions of a classical nonlocal theory.

I. INTRODUCTION

In metal-oxide-semiconductor (MOS) and in heterostructure systems, electrons can be confined in very narrow one-dimensional potential wells.[1] The characteristic dynamic intraband excitations in these quasi-two-dimensional (2D) systems are plasmons and, with a magnetic field, cyclotron and magnetoplasmon resonances. Of particular interest in this review are the collective longitudinal intraband excitations, the 2D plasmons. 2D plasmons have been studied both theoretically[2,3] and experimentally, in particular for the Si MOS system,[4–6] and are reviewed in several articles, e.g., Refs. 7–9. Recently 2D plasmon experiments in GaAs—heterolayers using Raman,[10] infrared emission,[11] and transmission[12] spectroscopy have also been performed. This research has also been extended to plasmon excitation in multiple-quantum-well systems.[13]

The plasmon frequency ω_L for an exactly 2D electron gas in the long-wavelength limit ($k_F \gg q \gg \omega/c$) and in a local approximation[2,3] is given by

$$\omega_L^2 = \frac{N_s e^2 q}{2\bar{\epsilon}(\omega,q)\epsilon_0 m_p} . \tag{1}$$

Here N_s is the surface charge density, q, the plasmon wave vector, $\bar{\epsilon}(\omega,q)$, the effective dielectric function, and m_p, the plasmon mass. In the random-phase approximation (RPA) the plasmon mass for an isotropic nonparabolic system is equivalent to the effective 2D band-structure mass.[14,15]

With increasing q, corrections to the dispersion (1) become important due to nonlocal effects,[2] correlations[16–19] and the finite thickness of realizable 2D systems.[18,20,21] So far, for the experimental conditions on the Si MOS system[5–7,14,22] with the limited experimentally achievable wave vectors, these corrections are small and could not be extracted within the experimental accuracy. For GaAs, however, at similar values of q, these corrections are more pronounced, mainly because of the smaller effective mass and the lower valley degeneracy.[12]

We report here on extended far-infrared (FIR) transmission spectroscopy studies on $Al_xGa_{1-x}As$-GaAs heterostructures. High-mobility samples and efficient grating couplers allow us to investigate nonlocal and finite-thickness corrections on the 2D plasmon dispersion (Sec. VI). Related to the latter we also discuss briefly intersubband resonance experiments in these samples (Sec. V). For the magnetoplasmon dispersion we show extended results, in particular at large wave vectors, of the nonlocal interaction[12] with harmonics of the cyclotron resonance (CR) (Sec. VII). Here we also gain important information on the origin of the experimental plasmon linewidth. In Sec. IV we discuss the experimental CR linewidth and an anomalous excitation strength for the CR amplitude in the near field of the grating coupler. We precede these topics by briefly summarizing some theoretical results on the plasmon dispersion (Sec. II) and by describing the experimental techniques and the sample characterization (Sec. III).

II. THEORETICAL REMARKS

The sample cross section is shown schematically in Fig. 1. The sample has a sandwich structure with a GaAs buffer layer (dielectric function ϵ_{Ga}) at the bottom followed by an undoped $Al_xGa_{1-x}As$ spacer layer (dielectric function ϵ_{Al}, thickness d_s) a doped $Al_xGa_{1-x}As$ layer (ϵ_{Al},d_{Al}), and a cap layer of GaAs (ϵ_{Ga},d_c). Periodic stripes with periodicity a of highly conducting Ag serve as a grating coupler to couple the normally transmitted FIR radiation with plasmons of wave vector $q = 2\pi m/a$ ($m=1,2,3,\ldots$). If we assume that this grating screens the plasmon electromagnetic fields perfectly and we neglect the small difference for ϵ_{Al} and ϵ_{Ga} in the region d_c, then the effective dielectric function for the plasmon frequency [Eq. (1)] is[3]

$$\bar{\epsilon}(\omega,q) = \tfrac{1}{2}[\epsilon_{Ga} + \epsilon_{Al}\coth(qd)] , \tag{2}$$

with $d = d_c + d_{Al} + d_s$.

With increasing wave vector q several corrections become important which change the dispersion relation in Eq. (1). Four correction terms have been proposed

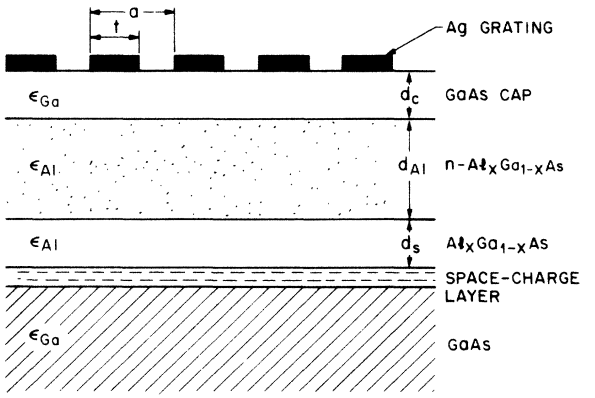

FIG. 1. Sample configuration of a $Al_xGa_{1-x}As$-GaAs heterostructure with grating coupler of periodicity a and strip width t. The quasi-two-dimensional electron space-charge layer is separated from the grating coupler by a distance $d=d_c+d_{Al}+d_s$. d_c, d_{Al}, and d_s are the thicknesses of a GaAs cap, doped $Al_xGa_{1-x}As$ layer, and undoped $Al_xGa_{1-x}As$ spacer, respectively. ϵ_{Ga} and ϵ_{Al} characterize the dielectric functions of GaAs and $Al_xGa_{1-x}As$, respectively.

theoretically. The first corrects for the finite thickness of the space-charge layer. The finite thickness affects the effective Coulomb interaction potential of the charged carriers in the channel, thus changing the effective dielectric function from its exactly 2D result [Eq. (2)].[1,18,20,21,23] According to Ref. 23 the effective dielectric function for a quasi-two-dimensional system with a screening gate is

$$\bar{\epsilon}(q,\omega)=\frac{\epsilon_{Ga}}{\mathcal{F}(q/b)}\ , \tag{3}$$

with

$$\mathcal{F}(x)=(1+x)^{-6}\left[\tfrac{1}{8}(33+54x+44x^2+18x^3+3x^4)+\frac{2\epsilon_{Ga}}{\epsilon_{Ga}+\epsilon_{Al}\coth(qd)}\right].$$

In deriving (3) a space-charge layer wave function $\psi(z)\sim ze^{-zb/2}$, perpendicular to the interface, has been assumed. $\mathcal{F}(q/b)$ is a form factor, depending on the average extent $1/b$ of the wave function in the z direction. The limit $1/b\to 0$ recovers the 2D result [Eq. (2)]. For $1/b\neq 0$ the plasmon frequency depends upon N_s as well as on the depletion charge N_{depl}, which is related to the background doping level of the semiconductor. The finite thickness reduces the plasmon frequency compared to the 2D result.

A further correction term lowering the local plasmon frequency is related to the confined perpendicular motion of the space-charge layer which for finite q is coupled to the intraband plasmons. For the approximation of a system of two confined states E_1 and E_2 and with subband separation $E_{21}=E_2-E_1$ the correction has been calculated in Ref. 24 and reads, neglecting the so-called depolarization shift[1]

$$\omega_p=\omega_L\left[1-\frac{2\bar{\epsilon}\epsilon_0 q}{e^2}\frac{v_{1112}^2(q\to 0)N_s}{E_{21}}\right]. \tag{4}$$

$v_{1112}(q)$ is the matrix element of the Coulomb interaction[25] averaged over the extent of the space-charge layer for both subbands.

Correlations also reduce the plasmon frequency. Electron correlations, influencing the plasmon frequency, have been discussed theoretically by several authors.[16–19] The results depend slightly on the models. If we use the result of Ref. 16 the plasmon frequency is modified as

$$\omega_p=\omega_L\left[1-\frac{1}{4}\frac{q}{k_F}\right], \tag{5}$$

where k_F is the Fermi wave vector.

The fourth correction term to which we will refer as the "nonlocal" correction is positive and leads to an increase in plasmon frequency.[2,16,17,19–21] In the long-wavelength limit it arises from a series expansion of the Lindhard polarization[2] in the wave vector. If we consider only the second-order term this modification can be expressed as

$$\omega_p=\omega_L\left[1+\frac{3}{4}\frac{q}{q_s}\right],\qquad q_s=\frac{g_v me^2}{2\pi\bar{\epsilon}(\omega,q)\epsilon_0\hbar^2}\ , \tag{6}$$

where q_s is the 2D screening wave vector,[1] and g_v the valley degeneracy. For a given q nonlocal corrections are much larger for GaAs than for Si because of the smaller mass and lower valley degeneracy.

In a magnetic field B perpendicular to the 2D plane, the magnetoplasmon frequency ω_{mp} is shifted in the local approach[3] by

$$\omega_{mp}^2=\omega_L^2+\omega_c^2\ . \tag{7}$$

Here $\omega_c=eB/m_c$ is the cyclotron frequency. Nonlocal corrections cause an interaction of the plasmon resonance with harmonics $n\omega_c$ ($n=2,3,4,\ldots$) of the cyclotron frequency resulting in a splitting of the magnetoplasmon dispersion at the intersections with $n\omega_c$. The amount of the splitting is governed by the parameter $(qv_F/\omega_c)^2$, in which v_F is the Fermi velocity. Since v_F is related to the plasmon frequency ω_L via N_s we can rewrite this parameter for $n=2$ as 6 (q/q_s) analogous to the nonlocal effect for $B=0$ T [Eq. (6)]. This interaction has been treated theoretically (e.g., Refs. 26–29) and has recently been observed in $Al_xGa_{1-x}As$-GaAs heterostructures.[12] We will discuss it in more detail in the course of the presentation of our experimental results.

III. SAMPLE CHARACTERIZATION AND EXPERIMENTAL SETUP

The samples are $Al_{0.28}Ga_{0.72}As$-GaAs modulation-doped single heterostructures grown by molecular-beam epitaxy (MBE) on GaAs semi-insulating substrates. Their cross section is shown in Fig. 1. The thickness of the GaAs buffer layer is about 1.3 μm. On top of a thin spacer layer ($d_s\approx 35$ Å) of undoped $Al_{0.28}Ga_{0.72}As$ is grown a Si-doped layer of $Al_{0.28}Ga_{0.72}As$ ($d_{Al}\approx 530$ Å, $n_{Si}=1\times 10^{18}$ (cm^{-3}). Both d_s and d_{Al} were estimated from the $Al_xGa_{1-x}As$ growth rates. The total thickness

$d_s + d_{Al}$ is verified by cross-section transmission electron microscopy to be 550 ± 20 Å. A protective cap layer of Si-doped GaAs of 570-Å thickness completes the layer sequence. After alloying indium contacts to the 2D channel, the cap layer was partly removed by reactive ion etching. The residual thickness of the cap layer d_{Ga} was determined by Auger electron sputtering to about 230 ± 80 Å. At this stage we characterized the samples by Shubnikov–de Haas (SdH) and cyclotron resonance (CR) measurements. The mobilities derived from the van der Pauw method are about 1.8×10^5 cm^{2}V^{-1}s^{-1}. The charge density is $N_s = (6.7 \pm 0.4) \times 10^{11}$ cm^{-2}. The spread in density reflects measurements from different samples including density gradients and slight N_s drifts caused by the cool-down process. Due to the persistent photoeffect the charge density in the 2D channel can be increased by illumination up to about 9.5×10^{11} cm^{-2}. As a side effect of light exposure a parallel conductivity caused by free carriers in the $Al_{0.28}Ga_{0.72}As$ layer is induced. However, if not otherwise noted all measurements here are performed at the initial charge density $N_s = 6.7 \times 10^{11}$ cm^{-2}.

To excite longitudinal collective oscillations of the two-dimensional electron gas (2D EG) with FIR radiation a grating coupler[4] is necessary (Fig. 1). A grating of periodicity $a \ll \lambda_{FIR}$ spatially modulates the incident radiation and couples to the collective oscillation with wave vectors $q = (2\pi m/a)$ $(m = 1,2,3,\ldots)$. For this purpose Ag gratings with periodicities of submicron dimensions were prepared on top of the samples using holographic interferometry and liftoff techniques. A second set of samples was prepared with a continuous thin metal layer underneath the grating. No influence of a Schottky barrier formation on N_s has been found in measuring the SdH effect in these samples. Fluctuations in the density are within the margins given above. In the following sections we will discuss the results of two representative samples with parameters listed in Table I.

The FIR measurements are done in transmission with a rapid scan Fourier transform spectrometer[30] in combination with a wave-guide system immersed together with a superconducting solenoid in a bath cryostat. The sample is mounted in the center of the magnet while the detector is located about 15 cm below the coil. Both are cooled by He exchange gas to a temperature of 4.3 K. A scatter filter is placed in front of the sample to avoid heating of the 2D EG by radiation and to operate the Ge:Ga composite bolometer under background limited conditions. Winston cone optics is used to focus the FIR radiation onto the sample and to guide the transmitted signal to the detector.

The quantity we extract from our measurements is the relative change in transmission

$$\frac{\Delta T}{T} = \frac{T(X) - T(Y)}{T(X)} ,$$

where T is the transmission depending on X and Y. X and Y can be either two different magnetic fields or two different charge densities. In the small signal approximation $\Delta T/T$ is proportional to the difference in the real parts of the 2D high-frequency conductivities[31] dependent on X and Y. For analytical expressions of the high-frequency conductivity we refer the reader to the Refs. 4, 5, 29, and 30. We will see in the following that for our experimental conditions we are already beyond the limits of the small signal approximation. We then use exact Fresnel coefficients of the system, taking into account the different dielectric layers and the quasi-two-dimensional electron gas.

TABLE I. Parameters of samples nos. 1 and 2. a, d, and R_G are the grating periodicity, total distance between the grating, and the 2D channel and sheet resistance of the continuous gate, respectively.

Sample no.	a (Å)	$d = d_s + d_{Al} + d_c$ (Å)	R_G (Ω/□)
1	8720±70	800±100	
2	6810±50	800±100	800

IV. CYCLOTRON RESONANCE EXCITATION WITH AND WITHOUT GRATING COUPLER

Novel features are observed in cyclotron resonance excitation measured on samples with grating couplers. Figure 2 shows the cyclotron resonance for sample no. 1 before and after preparation of the grating coupler. We have fitted the experimental CR line shape, obtained on samples without grating coupler, using exact Fresnel coefficients (Fig. 2) to obtain information about the scattering time and cyclotron mass. For our high-density and high-mobility samples we have to take into account the following facts. In the so-called linear approximation (see above) $\Delta T/T$ is proportional to the real part of the dynamic conductivity $\Delta T/T = A\,\mathrm{Re}\sigma(\omega,B)$, with

$$A = -2/[1 + (\epsilon_{Ga})^{1/2} + \sigma_G/\epsilon_0 c]$$

and σ_G the average conductivity of the gate. For $\tau = 7 \times 10^{-12}$ s, which corresponds to a dc mobility of about 180 000 cm^2/V s, and $N_s = 6.7 \times 10^{11}$ cm^{-2}, one calculates in this approximation a full width at half maximum (FWHM) of $\Delta\nu_c \approx 1.5$ cm^{-1} and a maximum signal

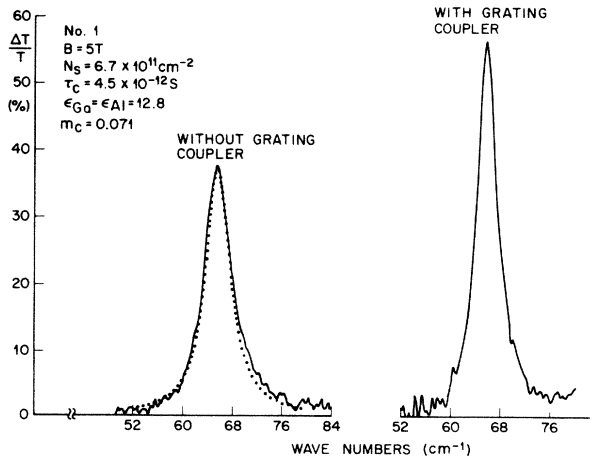

FIG. 2. Experimental cyclotron resonances for sample no. 1 with and without grating coupler. The dotted line is the line shape calculated with exact Fresnel coefficients and the listed parameters. The grating coupler enhances the CR amplitude and decreases the CR linewidth.

$(\Delta T/T)_{max}$ of about 310% for the active circular mode. The latter demonstrates that one is far beyond the limits of the linear approximation. Using exact Fresnel coefficients, one calculates that with increasing N_s the maximum signal saturates if it approaches 100% and the linewidth is broadened and no longer determined by τ alone. Deviations from the linear approximation are well known, see, e.g., Ref. 1, and have tentatively been observed in Ref. 32. The line shape calculated using Fresnel coefficients for linearly polarized incident radiation for samples without grating coupler is shown in Fig. 2 by a dotted line. The experimental CR amplitude is about 37% with a FWHM of $\Delta\nu_c=(5.5\pm0.3)$ cm^{-1} in the magnetic field range from 2.5 to 9 T. Best fits were obtained with a cyclotron mass $m_c=(0.0710\pm0.0005)m_e$ and a scattering time τ_c of about 4.5×10^{-12} s which is 1.5 times smaller than calculated from dc measurements ($\tau_{dc}\approx7\times10^{-12}$ s). The relation between τ_c and τ_{dc} is in good agreement with theoretical predictions for the case of short-range scatterers.[33] For magnetic fields smaller than 2.5 T a decrease in cyclotron mass is found with decreasing magnetic field. In addition a slight increase in CR amplitude and decrease in FWHM is observed. We could study the CR only down to about 1.7 T with a sufficient signal-to-noise ratio since our instrument posed a low-frequency limit of about 15 cm^{-1}. The cyclotron mass obtained at this particular field is $(0.0690\pm0.0005)m_e$. The magnetic field dependence of the CR might be related to the nature of the scattering processes involved. Scattering processes can introduce electron-electron interactions which can shift the CR position and induce harmonic CR transitions.[34] We would like to point out that, for the samples used here, we have not observed any indication of a harmonic cyclotron resonance.

A remarkable observation is made for cyclotron resonance with unpolarized radiation transmitted through samples with grating couplers, which are known to polarize the radiation linearly in the far field. We find that the signal increases to about 58% (Fig. 2), which is above the maximum value of 50% expected for $\Delta T/T$ on a sample without a grating coupler. The CR position is unshifted but its FWHM is reduced to $\Delta\nu_c\approx4$ cm^{-1}. This observation does not depend upon the direction of the magnetic field and there is no circularly polarized FIR radiation induced by our sample holder which could simulate such a result. We attribute the enhanced CR excitation to near field influences of the grating coupler. Deteriorating the grating efficiency reduces the enhanced CR amplitude significantly. For a quantitative discussion a rigorous theory for the grating coupler is required which includes the complex optical anisotropy of the sample. It is threefold anisotropic, due to the 2D system itself and the linear grating and is gyrotrop due to the magnetic field. Such a theory is presently not available.

V. GRATING COUPLER-INDUCED INTERSUBBAND RESONANCE IN SPACE-CHARGE LAYERS ON GaAs

The depletion charge governs the width of the 2D channel and thus the finite-thickness correction [Eq. (3)]. To obtain information about N_{depl} we have measured the intersubband resonance (ISR) transitions in our samples which are known to be strongly dependent on N_{depl}.[1] Due to the spherical symmetry of the GaAs conduction-band ISR transitions for $B=0$ T can only be excited with an electric field vector perpendicular to the interface. Such an electric field component perpendicular to the interface is provided by our efficient grating couplers[6,35,36] in their near field.

Figure 3 shows the ISR for sample no. 1 obtained in transmission for radiation incident perpendicular to the surface. In Fig. 3 we have evaluated

$$\frac{\Delta T}{T}=\frac{T(N_s)-T(N_{s\,sat})}{T(N_s)}\ .$$

The initial charge density N_s of 6.7×10^{11} cm^{-2} has been increased by a light emitting diode (LED) to its saturation value of $N_{s\,sat}=9.5\times10^{11}$ cm^{-2} (see Sec. III). ISR associated with the charge density $N_{s\,sat}=9.5\times10^{11}$ cm^{-2} is clearly observed in our experiment. The ISR signal is superimposed on a background of Drude type, showing a decrease in $\Delta T/T$ with increasing wave number. This Drude-type absorption is caused by carriers in the doped $Al_{0.28}Ga_{0.72}As$ layer which are also responsible for the parallel conductivity observed in dc-transport measurements. The resonance profile is of accumulation type[37] with a sharp increase on its low-energy side and a long high-energy tail, characteristic for a not dominating contribution of the depletion charge to the band bending at the interface. The resonance position is at about 400 cm^{-1} or equivalently at 50 meV. To estimate the depletion charge, we have compared the resonance position with calculations for electron energy levels in GaAs space-charge layers.[38] From this comparison we derived a value $N_{depl}\approx2.6\times10^{11}$ cm^{-2}. This value represents an upper limit, since we have not considered the depolarization field and excitonlike effect, which cause a shift be-

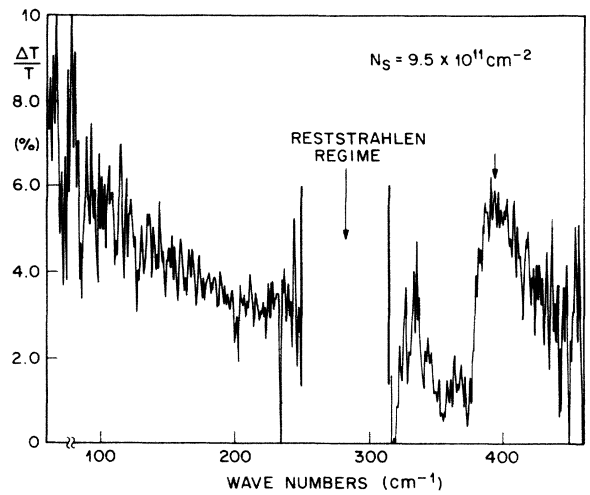

FIG. 3. Grating-induced intersubband resonance for sample no. 1. The arrow marks the resonance position for $N_{s\,sat}=9.5\times10^{11}$ cm^{-2}. The ISR is superimposed on a Drude-type background absorption caused by free carriers in the doped $Al_{0.28}Ga_{0.72}As$ layer.

tween the ISR transition energies and the actual subband separations.[1] No ISR is observed associated with the charge density $N_s = 6.7 \times 10^{11}$ cm^{-2}. Most likely for $N_s = 6.7 \times 10^{11}$ cm^{-2} the ISR transition is hidden in the frequency regime associated with the reststrahlen band of GaAs. Due to our thick substrates (≈ 0.3 mm), a regime spanning from 250 to 320 cm^{-1} is not transparent for FIR radiation (Fig. 3). However, SdH measurements give no evidence that a second subband is populated, and we conclude that in this case the subband separation is larger than the Fermi energy of 22.6 meV. On the other hand the ISR at 9.5×10^{11} cm^{-2} of 50 meV sets an upper limit for the subband separation at 6.7×10^{11} cm^{-2}. Assuming the ISR to be within the reststrahlen band of GaAs, we can estimate the depletion charge to be approximately 1.5×10^{11} cm^{-2} with an uncertainty of 0.5×10^{11} cm^{-2}. Although there is a large uncertainty in knowing N_{depl} we will show in the following that this does not affect the finite-thickness correction significantly.

VI. 2D PLASMON RESONANCES AT ZERO MAGNETIC FIELD

Experimental results for sample nos. 1 and 2 are shown in Fig. 4. In Fig. 4(a) we have evaluated

$$\frac{\Delta T}{T} = \frac{T(B=0) - T(B=8\text{ T})}{T(B=8\text{ T})} .$$

Two well-pronounced plasmon resonances of amplitudes 33% and 5%, respectively, are present. They correspond to plasmons with wave vectors $q_1 = 2\pi/a$ and $q_2 = 4\pi/a$, respectively. In Fig. 4(b) we show

$$\frac{\Delta T}{T} = \frac{T(N_s) - T(N_s + dN_s)}{T(N_s + dN_s)} ,$$

which for $dN_s \ll N_s$ is qualitatively the derivative of T with respect to N_s. N_s has been changed by a short light pulse from a LED. In this spectrum a third resonance due to plasmon excitation at wave vector $q_3 = 6\pi/a$ is also resolved for sample no. 1. Due to a lower grating-coupler efficiency for sample no. 2 only one resonance is present in Fig. 4(c). This trace is again qualitatively a derivative with respect to N_s. This time the density change is induced by a small voltage applied between the gate and the 2D channel. Column 3 of Table II lists the experimental resonance positions for both samples, including their experimental uncertainties.

In order to compare the experimental plasmon frequencies with theoretical values we first have to discuss how accurate an absolute plasmon frequency can be determined within the knowledge of the experimental parameters. First we determined for sample no. 1 that the assumption of an ideally screening gate is valid. We evaporated onto the grating coupler of this sample an additional layer of Cr of thickness 5 nm and resistivity 1 $k\Omega/\square$. We found no change in the plasmon resonance position, indicating perfect screening. This can also be deduced from the fact that three plasmon resonances are observed in Fig. 4(b), which assures that the ratio t/a (Fig. 1) is close to one. Only then an efficient excitation of higher spatial Fourier components of the FIR field is possible.[7]

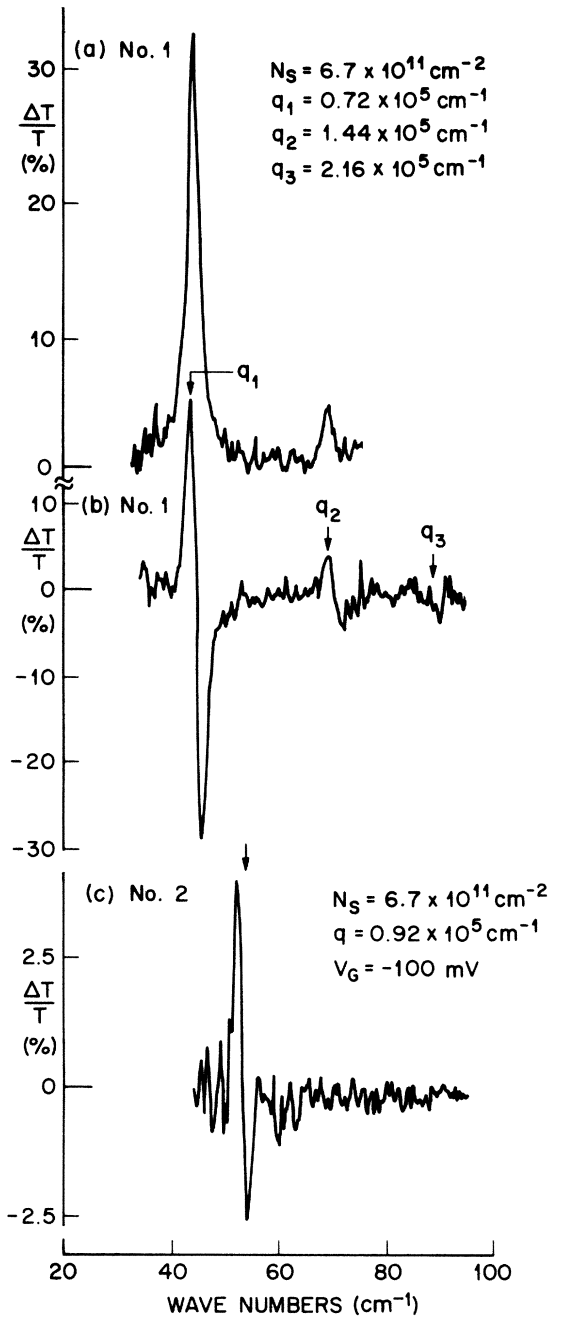

FIG. 4. Experimental plasmon resonances for a space-charge layer in GaAs with $N_s = 6.7 \times 10^{11}$ cm^{-2}. The arrows mark plasmon resonance positions. For the upper trace (a) transmissions at $B = 0$ T and $B \geq 8$ T have been ratioed to determine $\Delta T/T$ for sample no. 1. The middle trace (b), qualitatively the derivative of the upper trace, is obtained by changing N_s slightly via the persistent photoeffect. The lower trace (c) is for sample no. 2 with a thin metal layer underneath the grating. A voltage V_G applied between the gate and the space-charge layer is used to change N_s and to extract $\Delta T/T$.

The plasmon frequency depends on N_s, N_{depl}, d, q, ϵ_{Ga}, and ϵ_{Al}. We have made a great effort to determine these parameters, yet their uncertainties still limit the compar-

TABLE II. Experimental $\nu_{p\,\mathrm{expt}}$ and theoretical $\nu_{p\,\mathrm{loc}}$ plasmon frequencies calculated in classical local approximation for $\epsilon_{Al}=11.0$ and $\epsilon_{Ga}=12.8$. For the other parameters see Table I. The given inaccuracy $\Delta\nu_{p\,\mathrm{loc}}$ for $\nu_{p\,\mathrm{loc}}$ results from the uncertainties of N_s, d, and q. $\nu_{p\,\mathrm{nonloc}}$, ν_{pd}, and $\nu_{p\,\mathrm{corr}}$ are calculated plasmon frequencies including the nonlocal effect, the finite-thickness correction, and correlations, respectively.

Sample no.	q (10^5 cm^{-1})	$\nu_{p\,\mathrm{expt}}$ (cm^{-1})	$\nu_{p\,\mathrm{loc}}$ (cm^{-1})	$\Delta\nu_{p\,\mathrm{loc}}$ (cm^{-1})	$\nu_{p\,\mathrm{nonloc}}$ (cm^{-1})	ν_{pd} (cm^{-1})	$\nu_{p\,\mathrm{corr}}$ (cm^{-1})
1	0.72	44±0.5	43.0	±2.5	44.5	42.7	42.6
	1.44	69±1	69.1	±2.8	72.9	66.9	67.9
	2.16	88±2	87.6	±3.2	94.3	82.9	85.3
2	0.92	53±1	51.4	±2.6	53.4	50.7	50.8

ison of the absolute plasmon frequency. In the local approach [Eqs. (1) and (2)] the uncertainty in the parameters N_s, d, and q along limit the comparison to an accuracy of about ±2.5 to ±3.2 cm^{-1} (see column 5 of Table II). Thus, inspite of high-mobility samples and a careful characterization, the accuracy is limited to gain information on the higher-order q corrections from the absolute plasmon frequency. However, if we compare resonance positions for different wave vectors q, the uncertainties of the parameters act on the resonance position for different q in the same direction, at least in a first approximation. So comparing plasmon frequencies at different wave vectors allows us to analyze effects on the plasmon dispersion in a quantitative way.

The parameters determining the plasmon frequency which could not be measured independently here are ϵ_{Ga}, ϵ_{Al}, and m_p. ϵ_{Ga} is known fairly accurately at low temperatures[39] and we have adopted a value of 12.8. To approximate the plasmon mass m_p we have substituted it by the cyclotron mass at low magnetic fields $m_c=0.069m_e$. Since ϵ_{Al} is not known at low temperatures for the frequency regime of interest here we have treated it as a variational parameter to reproduce the resonance position for sample no. 1 and q_2 with the classical local Eq. (1). From this procedure we have obtained a value $\epsilon_{Al}=11.0$ for the average dielectric function for the three-layer system of GaAs cap and $Al_{0.28}Ga_{0.72}As$ layers (Fig. 1). The theoretical local plasmon frequencies for all experimental wave vectors calculated with the discussed values of ϵ_{Ga}, ϵ_{Al}, and m_p and the measured parameters listed in Table I are summarized in column 4 of Table II. In the following we will discuss the size of the higher-order q correction terms to the local plasmon dispersion.

In Fig. 5 the local plasmon dispersion [Eqs. (1) and (2)] is shown by a solid line. Both plasmon dispersions taking into account the nonlocal effect [Eqs. (1), (2), and (6), marked as curve 1] and the finite-thickness correction [Eqs. (1) and (3), marked as curve 2] are shown by dashed lines. At a wave vector of 2.16×10^5 cm^{-1} the nonlocal correction accounts for an increase in plasmon frequency of 6.7 cm^{-1} (see column 6 of Table II). If the nonlocal correction would be the only q-dependent correction it should lead to a clear increase with respect to the classical local plasmon frequency. The calculated correction is so large that it could be well seen within the experimental accuracy. In the magnetoplasmon experiments, to be discussed in Sec. VII, we directly observed the same nonlocal interaction, in excellent agreement with theory, so we know that this mechanism is present. However, here the predicted large plasmon frequency enhancement cannot be deduced from our experiment and we conclude that the nonlocal correction must be canceled due to other processes. Indeed, Fig. 5 demonstrates, that in particular the finite-thickness correction lowers the plasmon frequency significantly. We calculated the finite-thickness correction for $N_{depl}=1.5\times10^{11}$ cm^{-2} and found the result to be rather insensitive to the absolute value of N_{depl} in the range from 0.5×10^{11} cm^{-2} to 3×10^{11} cm^{-2}. Variations are smaller than 1.3 cm^{-1} even at the largest q values shown. A major fraction of the nonlocal correction is compensated by the finite-thickness correction, which at $q=2.16\times10^5$ cm^{-1} accounts for a decrease in the local plasmon frequency of about 4.6 cm^{-1} (column 7 of Table

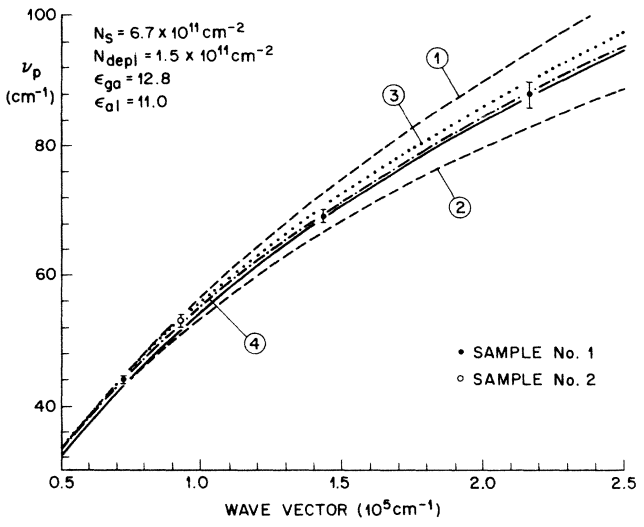

FIG. 5. Theoretical and experimental plasmon dispersion. The solid line is the classical local plasmon dispersions [Eqs. (1) and (2)]. The curves marked 1–4 are defined as follows: curve 1, plasmon dispersion including nonlocal correction [Eqs. (1), (2), and (6)]; curve 2, plasmon dispersion including finite-thickness effect [Eqs. (1) and (3)]; curve 3, plasmon dispersion including nonlocal and finite-thickness corrections combined [Eqs. (1), (3), and (6)]; and curve 4, plasmon dispersion including all correction term [Eqs. (1), (3), (5), and (6)].

II). Taking the two correction terms into account we show the resulting plasmon dispersion in Fig. 5 with a dotted line [Eqs. (1), (3), and (6), marked as curve 3]. The net result is positive and leads to a small increase with respect to the local plasmon frequency. At $q=0.72\times10^5$ cm^{-1} and $q=2.16\times10^5$ cm^{-1} this increase ranges between 1.3 and 2.4 cm^{-1}, respectively. Taking the nonlocal, finite-thickness, and many-body correction into account, we arrive at a plasmon dispersion shown in Fig. 5 by a dash-dotted line (marked as curve 4). The total correction is negligibly small and positive for all wave vectors with respect to the classical local plasmon dispersion. Because of the smallness of the correlation term it cannot be resolved within the experimental accuracy but we can clearly demonstrate the cancellation of the nonlocal plasmon frequency enhancement by the finite-thickness correction. In our analysis we have neglected the correction term due to coupling to the intersubband modes [Eq. (4)]. Since we are in the nonresonant limit $\hbar\omega_p \ll E_{21}$ this correction is negligible. Even for the highest wave vector the expected correction is lower than 1%. Recently resonant ISR-plasmon coupling has been studied experimentally in Ref. 40.

In the following we will discuss the linewidth of the plasmon resonance. From the spectra in Fig. 4 we find for sample no. 1 and q_1 an experimental linewidth of $\Delta\nu_p=2.5$ cm^{-1}. This is to our knowledge the smallest linewidth that has ever been observed for 2D plasmon resonances in semiconductor space-charge layers. Nevertheless, it is roughly three times larger than expected from the dc mobility ($\tau=7\times10^{12}$ s, $\Delta\nu_{\text{theory}}=0.76$ cm^{-1}).

Several processes can be responsible for the large experimental linewidth. The first is broadening due to inhomogeneities in the sample. A 6% variation of N_s over the active sample area is enough to explain the observed broadening. A variation of d also causes inhomogeneous broadening, directly via the plasmon dispersion and indirectly via a smaller N_s for a thinner doped $Al_xGa_{1-x}As$ layer. To check these assumptions we have blocked most of the active sample area and measured only a small spot of 2 mm in diameter at different locations within the 25 mm^2 original area. We found no difference in the resonance position and linewidth. Therefore, if inhomogeneity is the main reason for the broadening, then the relevant scale is smaller than the diameter given above.

Several other processes must be considered. For our experiments a strong coupling to the plasmon resonance is observed on sample no. 1. More than 30% of the transmitted FIR radiation is coupled within the plasmon excitation, implying that the plasmon resonance is strongly coupled to radiation fields. This should lead to additional damping processes resulting in line shifts and a broadening of the resonance. In the small signal approximation it has been calculated[8] that for the conditions here with $\omega/c \ll q$ such influences are expected to be small. However, the question arises whether this approximation can still be applied for 30% signal coupling. It would be helpful to have more accurate theories on the linewidth and dispersion of plasmons coupled strongly to radiation fields which are presently not available. In any case, we think that the strong coupling cannot be the main reason for the broadening, since similar plasmon linewidths are also observed for couplers with less efficiency, e.g., sample no. 2.

For high-mobility samples other processes become also important. We have calculated that ohmic losses in the grating coupler affect the plasmon linewidth. Assuming an ohmic resistivity of 1 $k\Omega/\square$ causes a 100% broadening of the plasmon linewidth. From the evaporation of additional Cr layers mentioned above, we can conclude that this effect is not the main contribution to the observed linewidth. A similar influence, however, would also arise from dielectric losses in the $Al_xGa_{1-x}As$ layer, which are not known for the doped region.

Recently the 2D plasmon dispersion and linewidth for GaAs heterostructures has been investigated theoretically[16] where, in addition to nonlocal, correlations, and finite-thickness corrections, scattering processes are also included within a self-consistent memory function approach. Impurity scattering couples the 2D plasmon resonance with different types of excitations, in particular plasmons of different wave vectors and the electron-hole continuum. Dissipative and reaction processes influence both the linewidth and the resonance frequency. Thus, if the resonance positions and linewidths are interpreted in terms of cyclotron or plasmon masses and scattering times, these quantities become frequency and wave-vector dependent. However, at the relative high charge densities, effects here are expected to be small.[41] This is also confirmed by the fact that within our uncertainty in determining the FWHM (≈0.25 cm^{-1}) we do not observe a difference in the FWHM of the resonance with q_1 and $q_2=2q_1$, for sample no. 1. We conclude, within the experimental accuracy we cannot extract details of this theory beyond the nonlocal and finite-thickness corrections discussed above.

VII. MAGNETOPLASMON EXCITATION IN GaAs

Figure 6 shows plasmon resonances measured in magnetic fields B perpendicular to the interface of sample no. 1. Plasmon resonances are observed for $q_1=2\pi/a$ and $q_2=4\pi/a$. Both resonances shift with increasing magnetic field B to higher frequencies. At certain magnetic fields, $B\approx1.7$ T for q_1, a second resonance appears at lower wave numbers and increases in intensity with increasing B, whereas the original resonance observed at $B=0$ T decreases in intensity. Both resonances repel each other, indicating a resonant interaction of different modes. A similar antilevel crossing is also observed for the magnetoplasmon resonance with q_2 in the field regime $B\approx2.2$–3.5 T. In Fig. 7 we plotted the experimental cyclotron and plasmon resonance positions. The cyclotron resonance frequency ω_c increases linearly with B. The experimental magnetoplasmon resonances in general follow the classical magnetoplasmon dispersion [Eq. (7)] and approach the cyclotron resonance at high magnetic fields. However, when the magnetoplasmon dispersion crosses harmonics $n\omega_c$ ($n=2,3,4,\ldots$) of the cyclotron frequency, a strong interaction of plasmons and harmonic cyclotron resonance is observed. This results in a splitting of the magnetoplasmon dispersion and an exchange of the

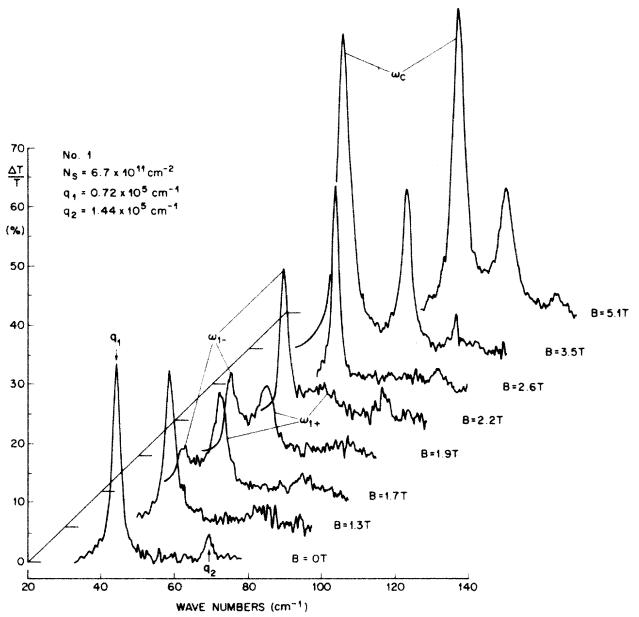

FIG. 6. Experimental plasmon resonances as a function of a perpendicular magnetic field B for sample no. 1. $\Delta T/T$ has been obtained by ratioing spectra for two different magnetic fields. For clarity the CR, marked ω_c, is shown only for the two higher magnetic fields. The plasmon resonance for $B=0$ T and q_1 (q_2) is marked by a downward (upward) pointing arrow. Due to interaction with the harmonic CR the nonlocal effect splits the plasmon resonance with q_1 into two branches, marked ω_{1-} and ω_{1+}, at around 1.9 T.

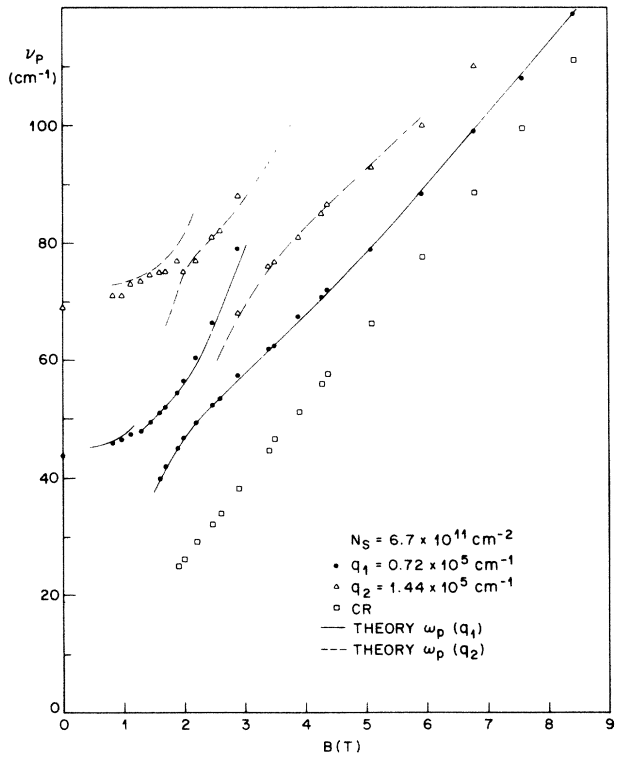

FIG. 7. Experimental plasmon resonance positions as a function of the magnetic field for two wave vectors q_1 and $q_2=2q_1$. Solid and dashed lines are calculated using Eq. (7) of Ref. 29. We have used the parameters listed in Table I and a scattering time of $\tau_p=7\times10^{-12}$ s. Experiment and theory are in excellent agreement.

excitation strength discussed above for Fig. 6. Excitation of harmonic cyclotron resonance is strictly forbidden in a parabolic isotropic translationally invariant system. Electron wave functions in a magnetic field are the wave functions of a harmonic oscillator, with energy levels $E_n=\hbar\omega_c(n+\frac{1}{2})$ and the optical matrix elements for $\Delta n\neq\pm1$ are zero. The dynamic spatial modulation of the charge density that characterizes the plasmon excitation breaks the isotropy and thus the selection rules given above are no longer valid. In several theoretical papers this nonlocal interaction of plasmons and harmonic CR has been calculated (e.g., Refs. 26–29). The strength of the interaction is governed by the parameter $[(v_Fq/\omega_c)]^2$. As was pointed out in Ref. 29, the small effective mass and the small valley degeneracy make the splitting for GaAs very large. For $n=2$, $(v_Fq/\omega_c)^2$ is 0.26 for q_1 and 0.56 for q_2, whereas in previous experiments on Si (Ref. 5) it is 0.005.

This nonlocal interaction was first reported in Ref. 12 for a wave vector $q=0.56\times10^5$ cm^{-1} and for $2\omega_c$. Here we can observe the splitting for $q_1=0.72\times10^5$ cm^{-1} and $q_2=1.44\times10^5$ cm^{-1}. The experiment shows that the splitting increases with increasing wave vector q, which is a direct consequence of the nonlocality. For the large value of q_2 the effect is strong enough that a splitting has also been observed for $3\omega_c$ (Fig. 7). In Ref. 29 the nonlocal plasmon excitation was calculated, including finite values of the scattering time τ. The latter is important to describe the amplitude of the excitation. With this theory we have calculated the theoretical curves in Fig. 7. Using the same parameters as in Table I the theory reproduces excellently the experimental magnetoplasmon dispersion without any fitting parameter. The nonlocal splitting for both q_1 and q_2 is in exact agreement with theory. The theory also predicts, that the splitting is much smaller for $3\omega_c$ as it is observed in the experiment. With the same parameters and formalism we have also calculated the dependence of the amplitude on the magnetic field B. There is good agreement with the theory for the ω_+ branch of the magnetoplasmon dispersion, as shown in Fig. 8(a). The enhanced experimental amplitude in the region of 3 T for the ω_- branch might have the same origin that causes the variation in linewidth and will be discussed for Fig. 8(b) below. The experimental linewidth of the magnetoplasmon resonance increases with B, except for the crossing regime at $2\omega_c$. There the ω_- branch exhibits an asymmetric profile and a variation of the half width as shown in Fig. 8(b). The linewidth of the ω_+ branch increases with B, whereas the linewidth of the ω_- branch shows a minimum at $B\approx2.5$ T and then increases with B. Assuming a B-independent τ, the local magnetoplasmon theory predicts an increase of the magnetoplasmon linewidth, due to the increase for the slope of the magnetoplasmon dispersion. For $\tau=7\times10^{-12}$ s $\Delta\nu_p$ is 0.75 cm^{-1} at $B=0$ T and 1 cm^{-1} at $B=3.5$ T. The nonlocal plasmon theory predicts that in the crossing regime the linewidths of both the ω_+ and ω_- branch increase,

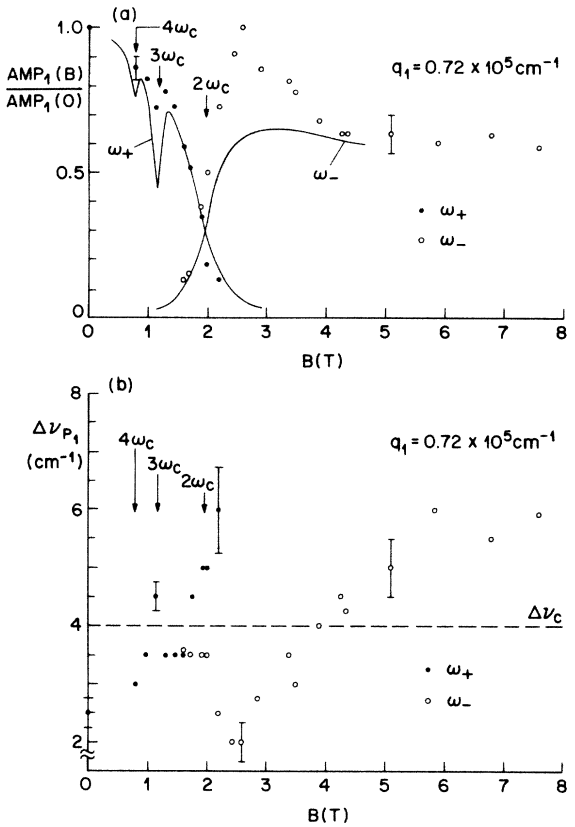

FIG. 8. Normalized magnetoplasmon amplitudes (a) and FWHM (b) for a wave vector of $q_1 = 0.72 \times 10^5$ cm^{-1} as a function of the magnetic field B. The dashed line indicates the FWHM of the CR measure on samples with grating couplers. Solid lines are calculated using Eq. (7) of Ref. 29. The parameters used are the same as described for Fig. 7. Positions of CR harmonics are marked by arrows. The overall agreement between experiment and theory for the normalized amplitudes is good for the ω_+ branch, but not for the ω_- branch between 2 and 4 T. In this magnetic field regime also the FWHM of the ω_- resonance shows an anomaly.

e.g., $\Delta\nu_{p-} = 1.1$ cm^{-1} at $B = 1.5$ T and $\Delta\nu_{p+} = 1.1$ cm^{-1} at 2 T. Asymmetric profiles are only found if the magnetoplasmon resonance coincides with higher harmonics of the CR and the nonlocal splitting is smaller than the linewidth of the magnetoplasmon resonance. Nevertheless this effect cannot be the reason for the observed anomalies in the magnetoplasmon line shape. We rather attribute the experimentally observed B dependence of the magnetoplasmon resonance to inhomogeneity and/or grating-coupler influences.

Calculations of the nonlocal magnetoplasmon resonance linewidth show that an inhomogeneity of N_s leads to a stronger broadening of the ω_+ branch as compared with the ω_- branch. Inhomogeneous broadening is stronger in the regime where the resonance is plasmonlike and thus dependent on N_s and d, whereas it is smaller in the regime where the mode is CR-like and thus not dependent on N_s and d. Inhomogeneous broadening is most likely the dominant effect for the plasmon resonance linewidth at $B = 0$ T (Sec. VI). Thus, its influence is also expected at finite magnetic fields. Furthermore the grating coupler might also affect the observable linewidth. In Sec. IV we demonstrated that the grating coupler causes a decrease in FWHM and an increase in amplitude of the CR. Here it could also influence the CR-like ω_- branch more than the plasmonlike ω_+ branch. Surprisingly the FWHM of the ω_- branch at $B \approx 2.5$ T is about half the FWHM of the grating-coupler-influence CR [Fig. 8(b)]. Qualitatively the same line-shape anomalies, discussed here for the plasmon resonance with q_1, have also been observed for the resonance with $q_2 = 2q_1$.

As discussed above, nonlocal corrections on the magnetoplasmon dispersion are small in Si(100) MOS structures.[5,22] However, a strong interaction with harmonics of the CR is also observed. It has been shown that this interaction is induced by scatterers which break the translational invariance of the 2D system.[42] This mechanism induces the excitation of so-called "subharmonic" CR in B-sweep experiments[32] on Si samples for $q = 0$. This experiment is equivalent to an observation of harmonic CR in frequency sweeps at fixed B. For our GaAs samples we observe harmonic CR only with grating couplers. The excellent agreement of the nonlocal theory with our experiment also demonstrates that the observed splitting is of nonlocal origin and not of the scatterer-induced type discussed for the experiments on Si.[42] From the experimental nonlocal magnetoplasmon dispersion in Fig. 7, in particular, in the region where the resonance is dominantly CR-like, an effective harmonic cyclotron mass $m_{c_n} = eB/n\omega_c$ can be extracted. Both m_{c_2} and m_{c_3} agree, within less than 2% with the cyclotron mass $m_c = 0.071$ for magnetic fields larger 2.5 T. On the contrary, for scatterer-induced harmonic CR on Si, m_{c_n} increases due to electron-electron interaction.[42] The importance of electron-electron interactions on a possible m_{c_n} $(n \geq 2)$ enhancement for GaAs is unknown, since scatterer-induced harmonic CR has not yet been confirmed experimentally. However, considering the fact that for the same charge densities, due to the smaller effective mass, the average kinetic energy for GaAs is larger than for Si, compared to the average interaction energy, a smaller m_{c_n} enhancement is expected.

VIII. CONCLUSIONS

We have investigated the 2D plasmon dispersion in high-mobility $Al_{0.28}Ga_{0.72}As$-GaAs heterojunctions. In perpendicular magnetic fields a strong interaction with harmonics of the CR is observed, resulting in a splitting of the magnetoplasmon dispersion. This splitting increases with wave vector q and is in excellent agreement with a quasiclassical nonlocal theory. The plasmon dispersion for $B = 0$ T has been investigated for plasmon wave vectors q larger than 2×10^5 cm^{-1}. The large $\approx 8\%$ increase of the plasmon frequency due to nonlocal effects at this q is found to be canceled primarily by the effect of the finite spatial extent of the quasi-two-dimensional space-charge layer.

ACKNOWLEDGMENT

We would like to thank H. L. Störmer for many valuable discussions.

[1]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[2]F. Stern, Phys. Rev. Lett. **18**, 546 (1967).
[3]A. V. Chaplik, Zh. Eksp. Teor. Fiz. **62**, 746 (1972) [Sov. Phys.—JETP **35**, 395 (1972)].
[4]S. J. Allen, Jr., D. C. Tsui, and R. A. Logan, Phys. Rev. Lett. **38**, 980 (1977).
[5]T. N. Theis, J. P. Kotthaus, and P. J. Stiles, Solid State Commun. **24**, 273 (1977).
[6]D. Heitmann, J. P. Kotthaus, and E. G. Mohr, Solid State Commun. **44**, 715 (1982).
[7]T. N. Theis, Surf. Sci. **98**, 515 (1980).
[8]A. V. Chaplik, Surf. Sci. Rep. **5C**, 289 (1985).
[9]D. Heitmann, Surf. Sci. **170**, 332 (1986).
[10]D. Olego, A. Pinczuk, A. C. Gossard, and W. Wiegmann, Phys. Rev. B **25**, 7867 (1982).
[11]R. Höpfel, G. Lindemann, E. Gornik, G. Stangl, A. C. Gossard, and W. Wiegmann, Surf. Sci. **113**, 118 (1982).
[12]E. Batke, D. Heitmann, J. P. Kotthaus, and K. Ploog, Phys. Rev. Lett. **54**, 2367 (1985).
[13]J. K. Jain and P. B. Allen, Phys. Rev. Lett. **54**, 2437 (1985).
[14]E. Batke and D. Heitmann, Solid State Commun. **47**, 819 (1983).
[15]E. Batke, D. Heitmann, A. D. Wieck, and J. P. Kotthaus, Solid State Commun. **46**, 269 (1983).
[16]A. Gold, Z. Phys. B **63**, 1 (1986).
[17]D. E. Beck and P. Kumar, Phys. Rev. B **13**, 2859 (1976).
[18]M. Jonson, J. Phys. C **9**, 3055 (1976).
[19]A. K. Rajagopal, Phys. Rev. B **15**, 4264 (1977).
[20]J. I. Gerstens, Surf. Sci. **97**, 206 (1980).
[21]R. Z. Vitlina and A. V. Chaplik, Zh. Eksp. Teor. Fiz. **81**, 1011 (1981) [Sov. Phys.—JETP **54**, 536 (1981)].
[22]E. G. Mohr and D. Heitmann, J. Phys. C **15**, L753 (1982).
[23]T. K. Lee, C. S. Ting, and J. J. Quinn, Solid State Commun. **16**, 1309 (1975).
[24]Sankar Das Sarma, Phys. Rev. B **29**, 2334 (1984).
[25]K. Nakamura, H. Ezawa, and K. Watanabe, Phys. Rev. B **22**, 1892 (1980).
[26]T. K. Lee and J. J. Quinn, Phys. Rev. B **11**, 2144 (1975).
[27]N. J. M. Horing and M. M. Yildiz, Ann. Phys. **97**, 216 (1976).
[28]M. L. Glasser, Phys. Rev. B **28**, 4387 (1983).
[29]A. V. Chaplik and D. Heitmann, J. Phys. C **18**, 3357 (1985).
[30]E. Batke and D. Heitmann, Infrared Phys. **24**, 189 (1984).
[31]D. C. Tsui, S. J. Allen, Jr., R. A. Logan, A. Kamgar, and S. N. Coppersmith, Surf. Sci. **73**, 419 (1978).
[32]G. Abstreiter, J. P. Kotthaus, J. F. Koch, and G. Dorda, Phys. Rev. B **14**, 2480 (1976).
[33]T. Ando, J. Phys. Soc. Jpn. **38**, 989 (1975).
[34]T. Ando, Phys. Rev. Lett. **36**, 1383 (1976).
[35]E. Batke, D. Heitmann, and E. G. Mohr, Physica **117&118B**, 643 (1983).
[36]D. Heitmann and U. Mackens, Phys. Rev. B **33**, 8269 (1986).
[37]A. D. Wieck, E. Batke, D. Heitmann, and J. P. Kotthaus, Phys. Rev. B **30**, 4653 (1984).
[38]F. Stern and Sankar Das Sarma, Phys. Rev. B **30**, 840 (1984).
[39]K. S. Champlin and G. H. Glover, Appl. Phys. Lett. **12**, 231 (1968).
[40]S. Oelting, D. Heitmann, and J. P. Kotthaus, Phys. Rev. Lett. **56**, 1846 (1986).
[41]A. Gold (private communication).
[42]T. Ando, Solid State Commun. **27**, 895 (1978).