Diamagnetism as a probe of exciton localization in quantum wells

K. J. Nash and M. S. Skolnick
Royal Signals and Radar Establishment, St. Andrews Road, Great Malvern, Worcestershire WR14 3PS, United Kingdom

P. A. Claxton and J. S. Roberts
Science and Engineering Research Council Central Facility for III-V Semiconductors, Department of Electronic and Electrical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom

(Received 27 September 1988)

Exciton photoluminescence (PL) is an important technique for the characterization of quantum wells (QW's). We discuss the effect of localization on the diamagnetic energy shift of an exciton in a QW. It is shown how the diamagnetism of an electron in the ground state, with arbitrary geometry, depends on the dimensions of its wave packet. We consider the properties of free excitons in QW's, introducing an effective electron-hole interaction. We use dimensional analysis to relate the finite-barrier problem to the simpler case of perfect confinement. For a bound exciton, localization in the plane of the QW causes the diamagnetism to be smaller than for a free exciton. The effect of localization is not important if the range is much larger than the free-exciton Bohr radius. The exciton diamagnetic shift is reduced by localization in 10 and 20 Å (In,Ga)As/InP QW's grown by solid-source molecular-beam epitaxy. Uncertainty about the value of the free-exciton diamagnetic shift limits the sensitivity of the diamagnetism as a probe of exciton localization. Despite this, the method still provides valuable information on exciton localization, particularly when combined with studies of the phonon sideband of PL.

I. INTRODUCTION

Low-temperature photoluminescence (PL) is one of the most important techniques for the characterization of quantum wells (QW's).[1] The PL spectrum is dominated by excitonic recombination.[2,3] Most of the information available from low-temperature PL originates in the study of excitons and often includes comparison with the excitonic states observed in absorption-type experiments such as photoluminescence excitation (PLE) spectroscopy, photoconductivity, or electroreflectance.[4,5] The exciton diamagnetic shift is one quantity that is frequently studied.[6–17] In this paper we consider how the diamagnetic shift is related to the spatial extent of the excitons observed in PL.

Except in QW's of very high quality, emission occurs at a lower photon energy than the excitonic absorption.[1] This "Stokes shift" arises because at low temperature the excitons that recombine radiatively are bound in disorder-induced states in the low-energy tail of the exciton band.[4] These states observed in PL have a very low density and therefore do not play a significant role in absorption processes.

Some properties of the luminescent excitons can be determined from the PL spectrum. The exciton localization energy is equal to the Stokes shift. Information on the spatial extent of these bound excitons, in the plane of the QW, is less direct. Both the strength of the phonon satellites of exciton PL,[18,19] and the exciton diamagnetic shift, are determined by the dimensions of the bound-exciton wave packet, albeit in a complicated manner.

The defects that bind excitons in a QW do not necessarily have rotational symmetry about an axis normal to the plane of the QW, but are likely to be elongated in one direction of this plane. In the preparation of III-V compound semiconductors by molecular-beam epitaxy (MBE) or metalorganic chemical vapor deposition (MOCVD), the crystallographic growth surface is usually (001). The point symmetry of the unreconstructed (001) surface is C_{2v}. Therefore the [110] and [$\bar{1}$10] directions in the growth surface are not related by symmetry, although these same directions are symmetry related in the bulk crystal. The low symmetry of the (001) surface is relevant, even in thick layers of material, when defects have been incorporated at the growth surface. For example, the bound-exciton PL of certain defects in thick layers of MBE GaAs is strongly polarized parallel to [$\bar{1}$10].[20,21] Each distinct center with strongly polarized emission is a complex of a small number of defect atoms, incorporated at the growth surface with [$\bar{1}$10] preferred to [110] as the direction of the electric dipole.

The effective potential that binds an exciton in the plane of a QW arises not from point defects, but from fluctuations in well width[22] or alloy composition.[23] Here, too, the inequivalence of the [110] and [$\bar{1}$10] directions of the growth surface can be important. For example, reflection high-energy electron diffraction (RHEED) studies[24] have shown that the MBE growth of GaAs proceeds by the formation of monolayer islands elongated along the [$\bar{1}$10] direction. The width fluctuations that always occur in a QW are likely to be elongated. We do not have specific evidence for such elongation in our QW's but, for generality, we discuss defects of arbitrary shape as well as the more frequently considered case in which the defect has rotational symmetry about an axis parallel to the growth direction.

We begin our study of diamagnetism by considering, in Secs. II and III, the diamagnetic energy shift of a single electron in a QW, and show how this is related to the geometry of the electronic wave packet. Section IV discusses exciton diamagnetism. In Sec. IV A we use dimensional analysis to express the properties of free excitons in QW's with finite barriers in terms of results previously calculated for perfect confinement in the QW (i.e., infinite barriers). In Sec. IV B we argue that the diamagnetic shift is always smaller for a localized exciton than for a free exciton. Section IV C considers the effect of high magnetic field, and Sec. IV D discusses nonparabolicity of the electron subbands.

In Sec. V we present experimental studies of the exciton PL for a series of (In,Ga)As/InP QW's grown by solid-source MBE.[25] In the narrower QW's the exciton diamagnetism is smaller than the free-exciton value, estimated using position-dependent effective-mass theory. We attribute this reduction of the diamagnetism to exciton localization. In Sec. VI we discuss the use of diamagnetism as a probe of exciton localization.

II. DIAMAGNETISM OF A BOUND ELECTRON IN EFFECTIVE-MASS THEORY

In effective-mass theory, the Hamiltonian for a single electron in a uniform magnetic field $\mathbf{B}$ is[26]

$$H_{\text{eff}}=[\mathbf{p}+e\,\mathbf{A}(\mathbf{r})]\cdot\frac{1}{2m_e^*(\mathbf{r})}[\mathbf{p}+e\,\mathbf{A}(\mathbf{r})] -eV_e(\mathbf{r})+\frac{g\mu_B}{\hbar}\mathbf{S}\cdot\mathbf{B}\ , \tag{1}$$

where $\mathbf{B}=\nabla\wedge\mathbf{A}$, the charge on the electron is $-e$, $\mu_B=e\hbar/(2m_e)$, V_e is the potential for electrons, and the other symbols have their usual meanings. The spin-orbit term in H_{eff} is ignored. The usual kinetic energy term[26] has been generalized for a position-dependent effective mass $m_e^*(\mathbf{r})$. This modified kinetic energy is Hermitian, gauge invariant, has the correct form for $m_e^*=$const, and gives "current-conserving" boundary conditions[27,28] at a discontinuity in m_e^*.

The most general time-independent form of the vector potential $\mathbf{A}(\mathbf{r})$ is

$$\mathbf{A}=\mathbf{A}_0+\nabla\chi\ , \tag{2}$$

where $\mathbf{A}_0(\mathbf{r})$ represents one choice of potential, and $\chi(\mathbf{r})$ is an arbitrary function. The eigenvalues ${}_iE$ ($i=0,1,2,\ldots$) of the Hamiltonian H_{eff} are gauge invariant, i.e., independent of $\chi(\mathbf{r})$. Let us consider the magnetic field as a perturbation and, following Griffith,[29] write

$$H_{\text{eff}}=H_0+H_{\text{I}}+H_{\text{II}}\ ,$$

where

$$H_0=H_{00}+\frac{g\mu_B}{\hbar}\mathbf{S}\cdot\mathbf{B}\ ,$$

$$H_{\text{I}}=\frac{e}{2}\left[\mathbf{p}\cdot\frac{1}{m_e^*}\mathbf{A}+\mathbf{A}\cdot\frac{1}{m_e^*}\mathbf{p}\right]\ ,$$

$$H_{\text{II}}=\frac{e^2}{2m_e^*}A^2\ .$$

The Hamiltonian

$$H_{00}=\mathbf{p}\cdot\frac{1}{2m_e^*}\mathbf{p}-eV_e(\mathbf{r})$$

applies when $\mathbf{B}=\mathbf{0}$. For a given direction of the magnetic field $\mathbf{B}$, let us express the eigenvalues ${}_iE$ of the full Hamiltonian H_{eff} as a power series in B, in which ${}_iE_n$ is the term nth order in B. Let the eigenfunctions of H_{00} be $|i\rangle\equiv\psi_i(\mathbf{r})$, corresponding to eigenvalues ${}_iE_0$. Griffith[29] has shown that for the case of a bound state of $V_e(\mathbf{r})$, ${}_iE_n$ is gauge invariant for $n\le 2$. His proof is for the case $\nabla\cdot\mathbf{A}_0=\nabla^2\chi=0$ with $m_e^*=$const, but is readily generalized to arbitrary $\mathbf{A}_0(\mathbf{r})$, $\chi(\mathbf{r})$, and $m_e^*(\mathbf{r})$.

The nth-order energy ${}_iE_n$ is easily evaluated. The spin term in H_0 contributes precisely $\pm\frac{1}{2}g\mu_B B$ to the energy. ${}_iE_1$ (the Zeeman energy) is the sum of this spin energy and the first-order perturbation due to H_{I}. ${}_iE_2$ is the sum of two terms:

$${}_iE_2={}_iE_2^1+{}_iE_2^2\ , \tag{3}$$

where ${}_iE_2^1$ is the first-order perturbation due to H_{II} and ${}_iE_2^2$ is the second-order perturbation due to H_{I}. Griffith[29] shows that ${}_iE_1$ and ${}_iE_2$ are gauge invariant, but that the separate contributions ${}_iE_2^1$ and ${}_iE_2^2$ to ${}_iE_2$ are not. ${}_iE_2$ is the diamagnetic energy, and we shall consider the bounds that can be placed on its value.

The first contribution to the diamagnetism is

$${}_iE_2^1=\left\langle i\left|\frac{e^2}{2m_e^*}A^2\right|i\right\rangle>0\ . \tag{4}$$

The second contribution to ${}_iE_2$ is the "second-order Zeeman term,"

$${}_iE_2^2=\sum_{j(\neq i)}\frac{|\langle i|H_{\text{I}}|j\rangle|^2}{{}_iE_0-{}_jE_0}\ . \tag{5}$$

For the ground state ($i=0$), the well-known result of second-order perturbation theory gives ${}_0E_2^2\le 0$. Substitution in Eq. (3) yields

$${}_0E_2\le{}_0E_2^1\ . \tag{6}$$

The value of ${}_0E_2^1$ depends on the choice of gauge; the smaller ${}_0E_2^1$ is, the better it is as an estimate of ${}_0E_2$, and the smaller the magnitude of ${}_0E_2^2$ is. In Appendix A we show that there exists a choice of $\mathbf{A}(\mathbf{r})$ such that ${}_0E_2={}_0E_2^1$. Thus, for any fixed $\mathbf{A}_0(\mathbf{r})$, the minimum value of ${}_0E_2^1$ with respect to variations in $\chi(\mathbf{r})$ is the exact value of the ground-state diamagnetic energy ${}_0E_2$. Also, applying Eq. (4) to the ground state gives ${}_0E_2>0$.

III. EXAMPLES OF DEFECT STATES AND THEIR DIAMAGNETISM

A. Axial or spherical symmetry

If the defect has axial symmetry, i.e., $V_e(\mathbf{r})$, $m_e^*(\mathbf{r})$ have an axis of rotational symmetry that lies parallel to $\mathbf{B}$, then $\mathbf{B}\cdot\mathbf{L}$ is a constant of the motion. We choose the Dingle gauge $\mathbf{A}=-\frac{1}{2}\mathbf{r}\wedge\mathbf{B}$, with the origin on the defect symmetry axis. H_1 is equal to $e/(2m_e^*)\mathbf{B}\cdot\mathbf{L}$.

The $B=0$ ground state $|0\rangle$ has magnetic quantum number $m_l=0$. Thus $(\mathbf{B}\cdot\mathbf{L})|0\rangle\equiv 0$ and so

$$\langle i|e/(2m_e^*)\mathbf{B}\cdot\mathbf{L}|0\rangle=0 .$$

It follows that the second-order Zeeman term ${}_0E_2^2$ vanishes, and the diamagnetism is equal to the ground-state expectation value

$${}_0E_2^1=\left\langle\frac{e^2}{2m_e^*}A^2\right\rangle=\frac{e^2B^2}{8}\left\langle\frac{1}{m_e^*}(x^2+y^2)\right\rangle \tag{7}$$

(defining the direction of $\mathbf{B}$ as the z direction).

B. States extended in one dimension in the plane normal to magnetic field

Suppose the confining potential $V_e(\mathbf{r})$ and the effective mass $m_e^*(\mathbf{r})$ are functions of x and z only, with magnetic field $\mathbf{B}$ parallel to z. With $B=0$, the stationary states are

$$\psi(x,y,z)=e^{ik_y y}\phi(x,z) .$$

Griffith's theorem does not apply to these unbound states. The perturbation theory is still gauge invariant, but there is the complication that the lowest unperturbed state of the continuum does not necessarily evolve into the lowest perturbed state when B is increased from zero. This property depends on the choice of gauge.

With the Landau gauge $\mathbf{A}=(0,B(x-x_0),0)$, k_y remains a good quantum number for $B\neq 0$. If x_0 is chosen so that $\langle(x-x_0)/m_e^*\rangle=0$ then, to second order in B, the state with $k_y=0$ remains lowest for $B\neq 0$. Perturbation theory is valid for small B if the unperturbed ground-state wave function is confined in the x direction by $V_e(x,z)$.

With this choice of gauge, the second-order Zeeman term vanishes for the state with $k_y=0$, and the diamagnetic shift is equal to

$${}_0E_2^1=\frac{e^2B^2}{2}\left\langle\frac{1}{m_e^*}(x-x_0)^2\right\rangle . \tag{8}$$

C. Application to multilayer systems

We consider an electron bound in an imperfect QW, with $\mathbf{B}$ normal to the QW interfaces. In expressions for the diamagnetism such as Eq. (8), it is convenient to take the effective mass m_e^* outside the expectation value. We introduce an approximation which makes this possible even when m_e^* is position dependent. Then we express the diamagnetic energy shift in terms of a length a_d.

In (homogeneous) 2D or 3D systems, m_e^* is independent of position; but in a layered structure (with interfaces perpendicular to z) the barrier material will have a larger effective mass than the well, at least in a type-I material system.[30] The discontinuity in effective mass means that the effective barrier height depends on the in-plane wave vector.

For a free electron in a QW the Schrödinger equation is separable. When $B=0$ the solutions are $\Psi_{ke}(z_e)\exp(i\mathbf{k}\cdot\boldsymbol{\rho}_e)$, where $(H_{ke}-E)\Psi_{ke}(z_e)=0$,

$$H_{ke}=-\frac{\hbar^2}{2}\frac{d}{dz_e}\frac{1}{m_e^*(z_e)}\frac{d}{dz_e}+\left[\frac{\hbar^2k^2}{2m_e^*(z_e)}-eV_e(z_e)\right] , \tag{9}$$

$V_e(z_e)$ is the potential due to the QW for $\mathbf{k}=\mathbf{0}$, and $\boldsymbol{\rho}_e=(x_e,y_e)$. The wave functions $\Psi_{ke}(z_e)$ are henceforth assumed to be normalized.

The in-plane mass for the electron at $\mathbf{k}=\mathbf{0}$, m_{pe}, is determined from Eq. (9) as

$$\frac{1}{m_{pe}}=\int_{-\infty}^{\infty}dz_e|\Psi_{ke}(z_e)|^2\frac{1}{m_e^*(z_e)} \tag{10}$$

by using the Hellmann-Feynman theorem to find the derivative of the eigenvalue of Eq. (9) with respect to k^2. The k dependence of $\Psi_{ke}(z)$ implies nonparabolicity of the in-plane dispersion relation; we ignore this k dependence and the resulting nonparabolicity, setting $k=0$ in $\Psi_{ke}(z)$. Then we approximate the ground-state wave function ψ_0 as the product $\psi_0(x,y,z)=\Psi_{0e}(z)\Psi_{\parallel}(x,y)$. Thus,

$$\left\langle\frac{1}{m_e^*}f(x,y)\right\rangle=\frac{1}{m_{pe}}\langle f(x,y)\rangle ,$$

where $f(x,y)$ is a function of coordinates in the plane of the QW.

It is convenient to express the diamagnetic energy ${}_0E_2$ in terms of the "diamagnetic length" a_d, defined by

$${}_0E_2=\frac{e^2}{4m_{pe}}B^2a_d^2 . \tag{11}$$

The results of Secs. III A and III B can be written in terms of a_d, noting that these expressions derived using m_{pe} are exact only if m_e^* is position independent.

For the case of axial symmetry,

$$a_d^2=\tfrac{1}{2}\langle x^2+y^2\rangle .$$

For the hydrogenic donor, the analytic solution in three dimensions[31] and in two dimensions[32] leads to the well-known results $a_d=a_B$ in 3D, and $a_d=\frac{1}{4}\sqrt{3}a_B$ in 2D (for $\mathbf{B}$ perpendicular to the plane), where a_B is the effective Bohr radius $(4\pi\epsilon\epsilon_0\hbar^2)/(e^2m_e^*)$. For states extended in the y direction (perpendicular to $\mathbf{B}$),

$$a_d^2=2(\Delta x)^2$$

where Δx is the uncertainty in x.

D. Defects whose form is unknown

The diamagnetic length a_d can be determined by measurement of the diamagnetic energy shift, and is a length scale related to the dimensions of the wave packet for the electron ground state. In Sec. III C, a_d is calculated from the ground-state wave function ψ_0 for certain simple cases. Now we consider the value of a_d for an arbitrary ψ_0. We take **B** to be parallel to z and choose the origin of our coordinates so that $\langle \mathbf{r}/m_e^* \rangle = \mathbf{0}$. We first treat the case where m_e^* is position independent.

We can obtain a useful inequality by applying the gauge of Sec. III B, $\mathbf{A}=(0,B(x-x_0),0)$, where $x_0=\langle x \rangle$. Now, $\langle A^2 \rangle = B^2(\Delta x)^2$, where Δx is the uncertainty in x; but the second-order Zeeman term does not necessarily vanish, and so

$$(\Delta x)^2 \geq \tfrac{1}{2}a_d^2 \ . \tag{12}$$

We can choose the x axis to lie along an arbitrary $\hat{\mathbf{n}}$ in the plane normal to **B**. If we know the value of the diamagnetism, and hence a_d, Eq. (12) places a lower bound on the spread of the wave function along any direction normal to **B**. It is the best possible lower bound, if there is no other information on the nature of the defect ground state. Equality is achieved in the limit where the defect state is extended in one direction normal to **B** so that $(\Delta x)_{\max} \gg (\Delta x)_{\min}$ (Sec. III B).

Figure 1 shows an example of an elongated defect, together with the vector potential $\mathbf{A}(\mathbf{r})$ for four choices of gauge, and the value of $\langle A^2 \rangle$ for each. The figure shows that the value of $\langle A^2 \rangle$ depends strongly on the choice of gauge, as it does for any defect. It also shows that for the elongated defect, the value of $\langle A^2 \rangle$ for the Dingle gauge does not give a good estimate of the diamagnetic energy, although it gives the exact answer for a defect with axial symmetry (Sec. III A). For any **A** such that $\mathbf{B}=\nabla \times \mathbf{A}$, $[e^2/(2m_{pe})]\langle A^2 \rangle$ is an upper bound on the diamagnetic energy: The smallest value is the best estimate. The actual value of the diamagnetism determines a lower bound on $\Delta(\mathbf{r}\cdot\hat{\mathbf{n}})$, the spread of the wave function for directions $\hat{\mathbf{n}}$ perpendicular to **B**. For the case illustrated in Fig. 1, the diamagnetic length a_d is slightly less than $2^{1/2}\Delta y$.

Suppose we want to calculate the diamagnetism of an electron in the ground state of a particular defect. Equation (12) places an upper bound on a_d and is the best possible upper bound if $(\Delta x)^2_{\min}$ is the only information we have on the defect ground state.

The question of a lower bound on a_d is more problematic. If the defect ground state is intermediate between a state with axial symmetry (Sec. III A) and a state free in one direction normal to **B** (Sec. III B), for example the elongated defect of Fig. 1, then we expect $a_d^2=\lambda(\Delta x)^2_{\min}$ with $1<\lambda<2$. However, defects with isotropic $\langle r_i r_j \rangle$ provide an interesting contrast. The quantity $\langle r_i r_j \rangle$ is a second-rank tensor, and the "isotropy" refers only to directions in the plane perpendicular to **B**. If $\langle r_i r_j \rangle$ is isotropic then $\langle (\mathbf{r}\cdot\hat{\mathbf{n}})^2 \rangle$ is independent of $\hat{\mathbf{n}}$, i.e., $(\Delta x)_{\min}=(\Delta x)_{\max}$. Thus, the Dingle gauge chosen in Sec. III A gives

$$ {}_0E_2^1 = \frac{e^2}{4m_{pe}} B^2 (\Delta x)^2_{\min} \ .$$

However, if the defect lacks axial symmetry (i.e., is anisotropic) then ${}_0E_2^2$ is nonzero in the Dingle gauge, and so $\lambda<1$. An extreme example of an anisotropic defect with isotropic $\langle r_i r_j \rangle$ is shown in Fig. 2. The potential is chosen so that, in the limit $R/W \to \infty$, $\Delta x \propto R$. A small repulsive potential is needed near the origin to prevent the formation, in this limit, of a bound state that is independent of R.[33] It is possible to choose the gauge so that when $R/W \to \infty$, ${}_0E_2^1$ is related only to W (Appendix B). Thus there is no general lower bound on a_d in terms of $(\Delta x)_{\min}$ or $(\Delta x)_{\max}$, though it seems likely that a_d can be given a lower bound in terms of a loosely defined "minimum dimension" of the defect (in this case W), which is also closely related to the actual value of a_d.

Another defect geometry with an interesting diamagnetic shift is the broken ring of Fig. 3. An unbroken ring has axial symmetry, and thus $a_d \approx (R-\frac{1}{2}W)/\sqrt{2}$. However, when the ring is broken the gauge can be chosen so that, when $R/W \to \infty$, ${}_0E_2^1$ is related only to W (Appen-

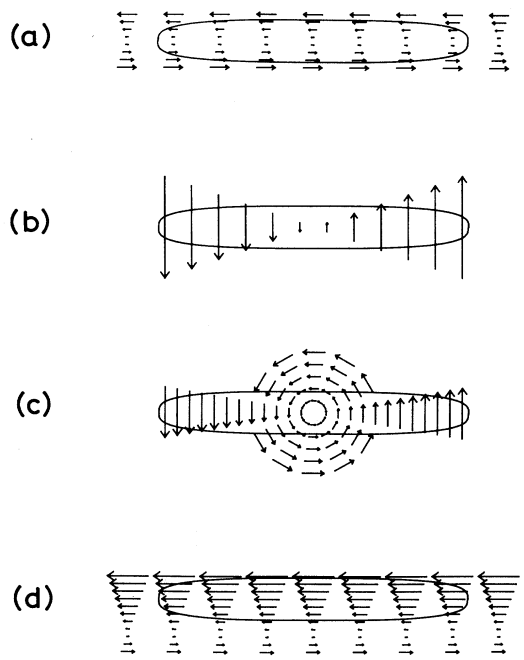

FIG. 1. The electron is confined within the curve (a contour in the x-y plane), which represents the edge of an elongated defect in the plane of a QW. The length and direction of each arrow form a vector proportional to the value of **A** at the midpoint of the length of the arrow. $\mathbf{A}(\mathbf{r})$ is illustrated for four choices of gauge; for each choice we give the value of $\langle A^2 \rangle$. (a) $\mathbf{A}=(-By,0,0)$; $\langle A^2 \rangle = B^2(\Delta y)^2$. (b) $\mathbf{A}=(0,Bx,0)$; $\langle A^2 \rangle = B^2(\Delta x)^2$. (c) $\mathbf{A}=\frac{1}{2}(-By,Bx,0)$; $\langle A^2 \rangle = \frac{1}{4}B^2[(\Delta x)^2 + (\Delta y)^2]$. (d) $\mathbf{A}=(-B(y+y_0),0,0)$; $\langle A^2 \rangle = B^2[(\Delta y)^2 + y_0^2]$. The coordinate system used in these expressions has its origin at the defect center of symmetry, so that the expectation value $\langle \mathbf{r}/m_e^* \rangle = \mathbf{0}$.

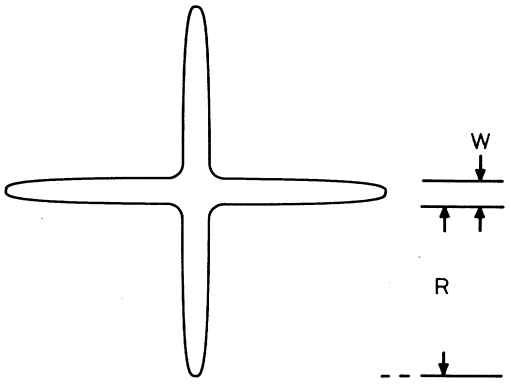

FIG. 2. Edge (defined as in Fig. 1) of an anisotropic defect with isotropic $\langle r_i r_j \rangle$.

dix B). As for the defect of Fig. 2, W is the "minimum dimension" of the defect that is likely to be related to a_d.

If the potential that confines an electron to the broken ring is finite, the wave function will penetrate the barrier which constitutes the gap in the ring. The diamagnetism will be very sensitive to the probability density in the gap.

When m_e^* is position dependent the bound given by Eq. (12) is no longer exact, but is subject to the approximations discussed in Sec. III C.

IV. THEORY OF EXCITON DIAMAGNETISM IN A QUANTUM WELL

A. Free excitons

The spread of the electron and hole wave functions from the QW into the barrier material has an important effect on the properties of the exciton.[34] We introduce an approximate form for the electron-hole interaction which leads naturally to a simple quantification of this spread. The properties of a free exciton in a QW with potential barriers of finite height can then be expressed, using dimensional analysis, in terms of the corresponding results for an exciton perfectly confined in a QW. The latter can be obtained from published calculations.[11,35,36]

The penetration of the wave function into the barrier also affects the in-plane electron mass m_{pe} [Eq. (10)]. For wide QW's, m_{pe} is close to the 3D electron mass in the well material; as the well width decreases and the wave functions penetrate the barrier more and more, m_{pe} increases towards the 3D electron mass in the barrier material (Fig. 4). The free-exciton diamagnetism is very sensitive to m_{pe}.

The hole subbands in a QW (Ref. 37) are much more complicated than the electron subbands because the kinetic energy has the four-band Luttinger form.[38] For $\mathbf{k}=0$ only, the energies and wave functions $\Psi_{kh}(z_h)$ of the heavy-hole subbands are given by the eigensolutions of a Hamiltonian with the same form as Eq. (9).[28,39] The in-plane dispersion of the heavy-hole subbands depends strongly on the mixing of heavy- and light-hole wave functions which occurs for $k \neq 0$, is significantly nonparabolic, and cannot be expressed in a simple form like Eq. (10). We take the lowest heavy-hole subband to have constant in-plane effective mass m_{ph}. Recent work[40,41] which uses the full Luttinger Hamiltonian is briefly discussed in Sec. VI.

Henceforth, the k dependence of $\Psi_{ke}(z)$ and $\Psi_{kh}(z)$ is always neglected: We set $\Psi_k \equiv \Psi_0$ and drop the subscript k. All forms of nonparabolicity are also neglected.

If the QW is neither too narrow nor too wide, the intersubband excitation energies are large compared to the exciton binding energy, and the overlap of the exciton ground-state wave function with subbands other than the lowest can be ignored. The motion of an electron-hole pair in the plane of the QW is governed by the Hamiltonian

$$H=-\frac{\hbar^2}{2m_{pe}}\left[\frac{d^2}{dx_e^2}+\frac{d^2}{dy_e^2}\right]-\frac{\hbar^2}{2m_{ph}}\left[\frac{d^2}{dx_h^2}+\frac{d^2}{dy_h^2}\right]+\frac{e^2}{4\pi\epsilon_{\text{eff}}\epsilon_0}U(\rho)\,, \tag{13}$$

where

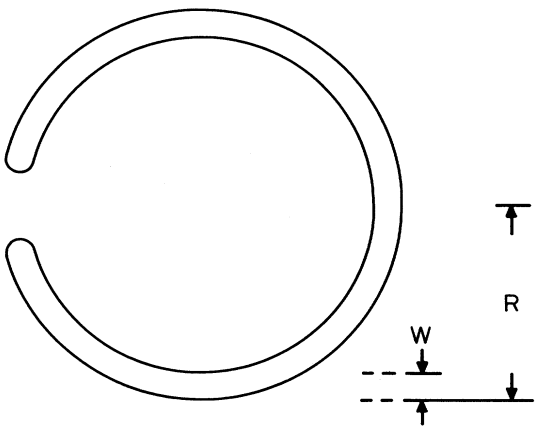

FIG. 3. Edge (defined as in Fig. 1) of the broken-ring defect.

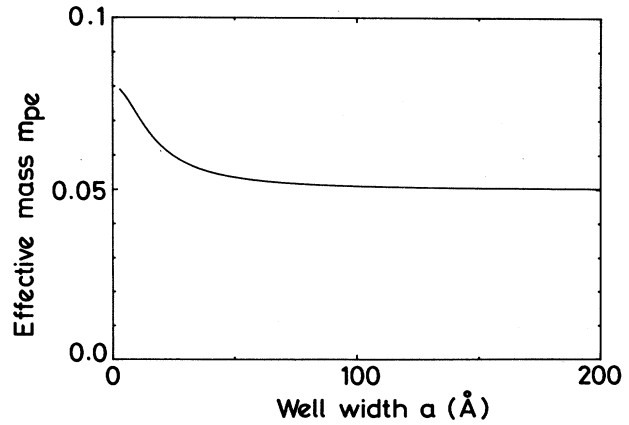

FIG. 4. The effective mass m_{pe} for motion in the lowest electron subband, as a function of well width. Calculated from Eq. (10) for an $In_{0.53}Ga_{0.47}As/InP$ QW.

$\rho=(x,y)=\rho_e-\rho_h,\quad \rho_e=(x_e,y_e),\quad \rho_h=(x_h,y_h)$.

The effective interaction $U(\rho)$ is the 2D Fourier transform of $-f(q)/q$, where[42]

$$f(q)=\int_{-\infty}^{\infty}dz_e\int_{-\infty}^{\infty}dz_h|\Psi_e(z_e)|^2|\Psi_h(z_h)|^2 \times\exp(-q|z_e-z_h|) .$$

Now, following the approximation due to Price,[43] we write $f(q)$ as

$$f(q)=\frac{1}{1+\frac{1}{3}Lq} \tag{14}$$

where

$$\frac{3}{2L}=\int_{-\infty}^{\infty}|\Psi_e(z)|^2|\Psi_h(z)|^2dz . \tag{15}$$

This expression has the correct form at $q=0$ and for large q and is also a good approximation for intermediate q. L is a length, which is equal to the well width a if $\Psi_e(z)$, $\Psi_h(z)$ are perfectly confined in the QW. Thus L can be considered to be an "effective well width," equal to the width of the QW that would have the same electron-hole interaction if perfect confinement were assumed. Figure 5 shows L as a function of well width for an $In_{0.53}Ga_{0.47}As/InP$ QW.

Equation (14) is equivalent in real space to

$$U(\rho)=-\frac{3\pi}{2L}\left[H_0\left[\frac{3\rho}{L}\right]-N_0\left[\frac{3\rho}{L}\right]\right] ,$$

where H_0,N_0 are Struve and Neumann functions of zero order. Thus Price's approximation gives a universal form for the real-space effective interaction, which has the correct behavior for large ρ and the correct singularity at $\rho=0$. It is particularly convenient that this approximate electron-hole interaction has a closed form both in real space and in reciprocal space.

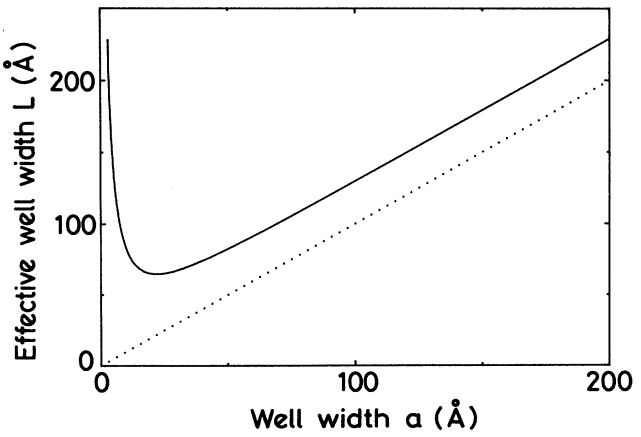

FIG. 5. Effective well width L, defined by Eq. (15), as a function of well width a. Calculated for an $In_{0.53}Ga_{0.47}As/InP$ QW. The dotted line is $L=a$.

The Hamiltonian, Eq. (13), can be written as a kinetic energy term representing the center-of-mass motion, plus the Hamiltonian for the relative motion

$$H_{\rm rel}=\frac{\hbar^2}{\mu}\left[-\frac{1}{2}\left[\frac{d^2}{dx^2}+\frac{d^2}{dy^2}\right]+\frac{1}{a_B}U(\rho)\right] , \tag{16}$$

where a_B is the effective Bohr radius

$$a_B=\frac{4\pi\epsilon_{\rm eff}\epsilon_0\hbar^2}{e^2\mu} , \tag{17}$$

and μ is the reduced mass $m_{pe}m_{ph}/(m_{pe}+m_{ph})$.

In the QW there is a discontinuity in ϵ at each interface. A full treatment of the electron-hole interaction in a QW requires the discussion of image charges. In the present case, the fractional discontinuity in ϵ is small, and we interpolate between the values for large and small well width a with the formula

$$\frac{1}{\epsilon_{\rm eff}}=\int_{-\infty}^{\infty}dz|\Psi_e(z)|^2\frac{1}{\epsilon(z)} .$$

Now, Price's expression for $f(q)$, Eq. (14), is a good approximation; therefore the form factor $f(q)$ for wave functions $\Psi_e(z),\Psi_h(z)$ may be approximated by the $f(q)$ for any other wave functions *which have the same L.* This correspondence in the electron-hole interaction for different wave functions exists whether or not the description of the electron-hole interaction explicitly involves $f(q)$.

The Hamiltonian for the relative motion of the electron and hole in a QW, Eq. (16), thus has the length scale a_B, the energy scale $E_R=\hbar^2/(2\mu a_B^2)$ (the three-dimensional effective Rydberg energy), the dimensionless parameter (L/a_B), and no other parameters. The results of calculations of exciton properties may be generalized by expressing them in terms of a_B, E_R, and L. In particular, the results of calculations that assume perfect confinement of the electron and hole in the QW can be generalized to the case where the penetration of the barrier is treated, by replacing the well width a by the effective width L. Two significant assumptions have been made in reaching this conclusion. The first is that the electron and heavy-hole in-plane dispersion curves are parabolic with effective masses m_{pe} and m_{ph}, respectively (absorbed into the reduced mass μ); a full treatment of nonparabolicity would introduce further dimensionless parameters. The second assumption is that the intersubband excitation energies are large compared to the exciton binding energy. This assumption breaks down both for very narrow and for very wide QW's; again, this breakdown is manifested by the appearance of additional dimensionless parameters.

The method can be applied, for example, to the binding energy E_{1s} of a free exciton in a QW. Both Bastard *et al.*[35] and Shinozuka and Matsuura[36] have calculated E_{1s}/E_R as a function of a/a_B, i.e., have found the function $v(x)$ such that $E_{1s}=E_R v(a/a_B)$ for a QW with perfect confinement. The generalization to the case of finite barriers is $E_{1s}=E_R v(L/a_B)$, with the *same* function $v(x)$. The value of L reflects the spread of the electron

and hole wave functions normal to the plane of the QW. Equation (15) and the a dependence of L (Fig. 5) express quantitatively the "spilling of the wave function" out of a QW, with which Greene *et al.*[34] explain the a dependence of their calculated E_{1s}. Their E_{1s} has its peak value for $a \approx 30$ Å in $GaAs/Ga_{0.7}Al_{0.3}As$ QW's. Rogers *et al.*[10] describe the exciton binding energy with an empirical "dimensionality parameter" D_1. At zero magnetic field $D_1 = v/4$.

The unitary transformation given in Ref. 44 reduces the problem of a free exciton in a magnetic field to a Schrödinger equation in terms of the relative coordinate $\boldsymbol{\rho}$, in which magnetic terms are added to the Hamiltonian for relative motion, Eq. (16). The electron-hole interaction has rotational symmetry about the z axis, and the free-exciton diamagnetic shift ${}_FE_2$ is given by

$$ {}_FE_2 = \frac{e^2B^2}{8\mu}\langle x^2+y^2\rangle $$

[cf. Eq. (7)]. We can define the diamagnetic length a_{dx} for the exciton ground state (whether free or bound) by

$$ E_2 = \frac{e^2B^2}{4\mu}a_{dx}^2 , \qquad (18) $$

where E_2 is the exciton diamagnetic energy. ${}_Fa_{dx}^2$, the free-exciton value of a_{dx}^2, is equal to the ground-state expectation value $\frac{1}{2}\langle x^2+y^2\rangle$ and can be expressed as ${}_Fa_{dx}^2 = a_B^2 w(L/a_B)$. The function $w(L/a_B)$ is obtained, by the method discussed above, from the calculation of Bugajski *et al.*,[11] which assumes perfect confinement, and which uses the trial wave function of Shinozuka and Matsuura.[36] w is equal to $\langle x^2\rangle/a_B^2$, whose square root is plotted in Fig. 4 of Ref. 36. Figure 3 of Bastard *et al.*[35] plots $(2w)^{1/2}$. Figure 1 of Bugajski *et al.*[11] plots a quantity equal to w multiplied by (in their notation) $\frac{1}{2}\gamma^2$. The empirical "dimensionality parameter" D_2, used by Rogers *et al.*[10] to describe the exciton diamagnetic shift, is equal to w.

We calculate the eigenfunctions of Eq. (9) with $In_{0.53}Ga_{0.47}As$ masses $m_e^*/m_e = 0.05$, $m_h^*/m_e = 0.5$, InP masses $m_e^*/m_e = 0.08$, $m_h^*/m_e = 0.6$, conduction-band offset 230 meV, valence-band offset 380 meV.[5] $m_{pe}(a)$, $L(a)$ and $a_B(a)$ are found from Eqs. (10), (15), and (17), respectively, and are given in Table I, together with $w(L/a_B)$. In Sec. V we use these values to calculate the free-exciton diamagnetism, and we compare the results with experiment.

B. Bound excitons

A bound exciton is an exciton that is localized at some defect. The binding potential that localizes the exciton motion in the plane of a QW is most likely to arise from fluctuations in well width[22] or in alloy composition.[23] The former, and probably also the latter, are such in the (In,Ga)As/InP system that a fluctuation that is attractive to electrons is also attractive to holes, in contrast to the Coulomb field of a shallow donor or acceptor.

The potential energy for the two-particle (electron-hole) problem is

$$ V_1(\mathbf{r}_1)+V_2(\mathbf{r}_2)+V_{12}(|\mathbf{r}_1-\mathbf{r}_2|) . $$

When at least one of the particles of the exciton is tightly bound by the defect potential, the effective-mass wave function can be approximated as a product form. The motion of the tightly bound particle (subscript 1) is dominated by the single-particle binding potential V_1; the other particle may also be tightly bound (by V_2) but, even if V_2 is negligible, the particle will at least be bound in a pseudodonor or pseudoacceptor state by the Hartree potential of the other particle (i.e., by V_{12}). The case in which the hole is tightly bound and the electron is bound mainly by the Coulomb attraction of the hole is similar to the isoelectronic donor model of exciton binding at certain point defects in bulk semiconductors,[45] and describes the excitonic recombination in (In,Ga)As/InP QW's grown by atmospheric pressure MOCVD.[19,46]

When the wave function has a product form, the diamagnetism is the sum of separate contributions from the electron and hole, and the analysis of Sec. III can be applied directly to the exciton problem. Thus

TABLE I. Theoretical and experimental parameters for the exciton diamagnetism in (In,Ga)As/InP QW's.

	Well thickness			
	10 Å	20 Å	28 Å	50 Å
$m_{pe}(m_e)$, in-plane electron mass	0.0719	0.0626	0.0584	0.0535
a_B (Å),[a] effective Bohr radius	94.7	111.9	121.9	136.2
L (Å), effective well width	82.7	64.7	66.1	82.4
$w(L/a_B)$[b]	0.50	0.42	0.41	0.43
${}_Fa_{dx}$ (Å),[a] FE diamagnetic length	66	72	77	88
${}_Fa_{dx}$ (Å),[c] FE diamagnetic length	83	90	97	111
${}_{F\infty}E_2$ (μeVT^{-2}),[a] FE diamagnetic energy	28	37	46	65
E_2 (μeVT^{-2}), observed, low field	14	16	45	75
a_{dx} (Å),[d] observed	48	47	77	96

[a]Free-exciton (FE) value, calculated with $m_{ph} = \infty$.
[b]Taken from Bugajski *et al.* (Ref. 11), see Sec. IV A.
[c]Free-exciton (FE) value, calculated with $m_{ph} = m_{pe}/0.3$.
[d]Calculated from E_2 using Eq. (18), the theoretical mass m_{pe}, and $m_{ph} = \infty$.

$$a_{dx}^2 = \frac{m_{ph}}{m_{pe}+m_{ph}} a_{de}^2 + \frac{m_{pe}}{m_{pe}+m_{ph}} a_{dh}^2 ,$$

where a_{de} and a_{dh} are single-particle diamagnetic lengths for the electron and hole, defined by Eq. (11) with masses m_{pe} and m_{ph}, respectively.

When neither particle is tightly bound, i.e., V_1 and V_2 are weak or very extended, the exciton motion is dominated by V_{12}, the electron-hole interaction. The electron and hole are extended over a large region R, but the electron-hole motion is correlated and can be considered as that of a free exciton whose center-of-mass motion is bound within R. Thus as the minimum dimension of R tends to ∞, the bound-exciton diamagnetism approaches that of the free exciton. Electron-hole correlation must start to become important when the charge distributions for both the electron and the hole have dimensions comparable to the radius for the internal motion of a free exciton, i.e., ${}_F a_{dx}$, the free-exciton value of a_{dx}. The effect of localization on the diamagnetism is not significant if the minimum range of localization is much larger than ${}_F a_{dx}$. We therefore suggest that the diamagnetism of a bound exciton is always smaller than that of a free exciton; and, if it is much smaller or if a pseudodonor or pseudoacceptor model is thought to apply, that it can be interpreted in terms of the dimensions of the uncorrelated electron and hole wave packets. The limiting of the diamagnetism by correlation contrasts with the behavior of a single electron bound within the same region R, for which a_d is related to the minimum dimension of R, however large.

An apparent exception to the upper bound on the exciton diamagnetism occurs when an electrostatic potential localizes the electron and hole in separate, diffuse bound states, far apart in the plane of the QW; however, in this case the electron-hole pair can hardly be considered an exciton.

C. High magnetic fields

The second-order perturbation theory for the diamagnetism is valid only if the magnetic length $l_B = (\hbar/Be)^{1/2}$ is somewhat larger than a_{dx}. Thus, a defect with a large value of a_{dx} (and hence a large diamagnetic energy) has a small value for the magnetic field at which second-order perturbation theory breaks down. At high field, bound-exciton energy levels can cross for some binding-center geometries (e.g., the unbroken ring of Sec. III D). In contrast, the free-exciton ground state at high field lies at some (weakly B dependent) binding energy below the lowest free-particle Landau level.[47,48] An exciton bound at a simple defect probably behaves similarly.

D. Nonparabolicity

Throughout this paper we have used position-dependent effective-mass theory. The "wave functions" (envelope functions) are eigenfunctions of a second-order Hermitian operator. As the well width is reduced, the penetration of the barrier by the electron and hole wave functions increases. This means that the electron effective mass for in-plane motion, m_{pe}, becomes more and more barrierlike (Fig. 4) and, for small enough well widths, the effective well width L, which measures the spread of the electron and hole wave functions in the direction perpendicular to the interfaces, increases (Fig. 5). These two trends are important in determining the free-exciton diamagnetism as a function of well width (Sec. IV A). They are expected from simple physical considerations and do not depend specifically on the effective-mass theory.

Conduction-band nonparabolicity is also important[49] and is often considered in connection with diamagnetism (e.g., Ref. 10). It will cause some correction to L and, in particular, to m_{pe}, but no qualitative difference to the dependence of these quantities on well width. The size of the nonparabolic correction to m_{pe} in narrow (In,Ga)As/InP QW's is unclear. The nonparabolicity in bulk $In_{0.53}Ga_{0.47}As$ is larger than predicted from $\mathbf{k}\cdot\mathbf{p}$ theory.[50] However, the effect on the exciton reduced mass μ of the increased m_{pe} in narrow QW's is offset by a reduction in m_{ph} from the large value appropriate in wide QW's towards the smaller value of the narrow-well limit. This fact was first demonstrated by the interband magnetoabsorption experiments of Rogers *et al.*[10] on GaAs/(Al,Ga)As QW's, although the precise value deduced for m_{ph} depends on the method used to fit the excitonic Landau levels at high magnetic field.[10,51] Ossau *et al.*[15] have also argued that μ is insensitive to well width, using the four-band Luttinger Hamiltonian[38] to calculate the dispersion of the heavy-hole subband. It is very difficult to estimate the overall effect of electron nonparabolicity and the complicated structure of the valence band on the free exciton diamagnetic shift.

V. EXPERIMENTAL RESULTS

Photoluminescence (PL) measurements were performed on an (In,Ga)As/InP structure (PMB117) grown by solid-source MBE and discussed elsewhere.[25,52–55] The structure consists of five nominally lattice-matched (In,Ga)As QW's with thicknesses $d = 110$, 50, 28, 20, and 10 Å, grown in that order on the same InP substrate, and separated by 1000-Å barriers of InP. The QW thicknesses were measured by transmission electron microscopy.[56] PL was excited by ~ 50 mW cm^{-2} of 5145-Å Ar^+-laser light and measured at 4.2 K in a magnetic field of up to 9.6 T, directed normal to the QW interfaces.

The 4.2 K PL spectrum at zero magnetic field has been reported in Ref. 25. Electron-hole recombination in each QW gives rise to a single strong peak in the PL spectrum. In the QW's with $d \leq 50$ Å, this is the zero-phonon line (ZPL) of exciton recombination. The spectrum also includes weak LO-phonon satellites of these ZPL's.[18] Coupling to LO modes of both InP and (In,Ga)As is observed. The strengths of these satellites and their dependence on magnetic field are related to the length scale of exciton localization.[18]

The 110-Å QW contains 1.4×10^{11} cm^{-2} free electrons,[54] and in magnetic field the PL spectrum splits into

lines corresponding to different free-electron Landau levels.[25,53] The free electrons have screened out the exciton diamagnetic shift, so this QW will not be discussed further.

Figure 6 shows the energy shift $E(B)$ of the ZPL for each QW with $d \leq 50$ Å. Experimental and theoretical results are presented in Table I. ${}_{F\infty}E_2(B)$, the *quadratic* diamagnetic shift for a free exciton with $m_{ph}=\infty$, is calculated using the theory of Sec. IV A and is plotted in Fig. 6. Since ${}_{F\infty}E_2(B)$ uses an overestimate of m_{ph}, it is expected to form a lower bound for the free-exciton energy shift ${}_FE(B)$ at low fields, but it will exceed all experimental energy shifts in the high-field (quasilinear) regime (Sec. IV C).

The 50-Å QW has a quasilinear energy shift for $B>4.6$ T, but slightly exceeds ${}_{F\infty}E_2(B)$ below this field [Fig. 6(a)]. In the 28-Å QW the quasilinear regime occurs for $B>6.0$ T. The experimental values of $E(B)$ lie below ${}_{F\infty}E_2(B)$ at low fields, but the difference is not large [Fig. 6(b)]. Exciton localization in the 50- and 28-Å QW's is not compact enough to be detected by comparison of the experimental diamagnetic shift with our estimate of the low-field free-exciton diamagnetism. The spread of the electron and hole in the plane of the QW can be represented by the uncertainty in their positions, Δx_e and Δx_h, respectively. The loose binding of the exciton in the 28- and 50-Å QW's is confirmed by analysis of the strength and magnetic field dependence of the phonon satellites,[18] which shows that the bound exciton is localized in a large region ($\Delta x_e, \Delta x_h > 130$ Å). Since this spread is somewhat larger than ${}_Fa_{dx}$ (values of which are given in Table I for $m_{ph}=\infty$ and $m_{ph}=m_{pe}/0.3$), any reduction of the diamagnetic shift due to exciton localization is expected to be small.

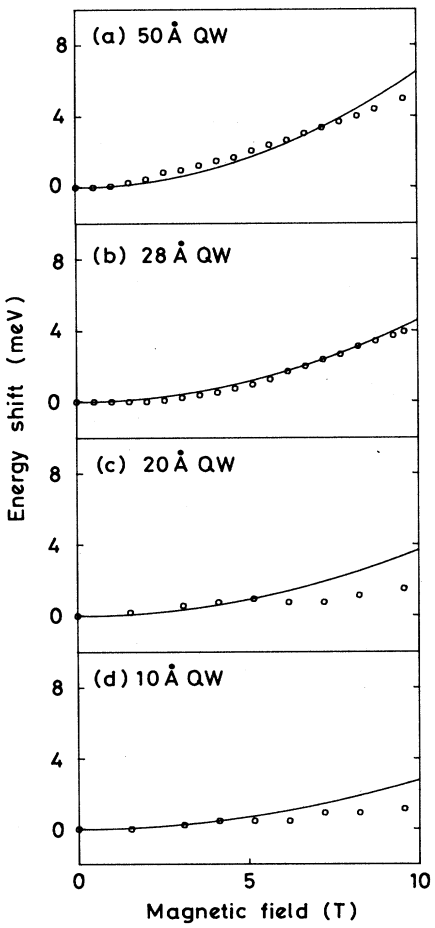

FIG. 6. Energy shift of the exciton PL, measured at 4.2 K, as a function of magnetic field. The circles are the experimental values. The curves are the quadratic free-exciton diamagnetic shifts ${}_{F\infty}E_2$ calculated with the position-dependent effective-mass theory of Sec. IV A (values from Table I). (a) 50-Å QW; (b) 28-Å QW; (c) 20-Å QW; (d) 10-Å QW. For the 50- and 28-Å QW's the behavior is close to that of the free exciton. The small experimental diamagnetism in the 20- and 10-Å QW's is attributed to exciton localization.

The reasonable agreement between ${}_Fa_{dx}$ (for $m_{ph}=\infty$) and the experimental a_{dx} in the 28- and 50-Å QW's is further supported by the high-field energy shift. The onset of the high-field quasilinear behavior of $E(B)$ occurs when $l_B/{}_Fa_{dx}\approx 1.4$ in both QW's, in accord with the qualitative arguments of Sec. IV C.

In both the 10- and 20-Å QW's [Figs. 6(c) and 6(d)] the diamagnetic shift at maximum field (9.6 T) is small (1.2 and 1.5 meV, respectively), reflecting a small a_{dx}, and so the energy shift is expected to have a quadratic form. In both QW's at 9.6 T, the experimental diamagnetic shift E_2 is less than half the value of ${}_{F\infty}E_2$. It is also notable that E_2 drops considerably when the QW width is reduced from 28 to 20 Å (Table I). We attribute the small values of the diamagnetism in the 10- and 20-Å QW's to exciton localization, which causes the diamagnetic length a_{dx} and hence E_2 to be smaller than for the free exciton. The localization is more compact in the narrower QW's because the influence of imperfections in the interfaces is stronger. The phonon satellites are more intense in the 10- and 20-Å QW's than in the 28- and 50-Å QW's, and much less sensitive to magnetic field, confirming that the exciton is more compactly bound than in the wider QW's.[18] Combining the data from diamagnetism and phonon satellite measurements, the exciton localization is calculated to be rather more compact ($\Delta x_e, \Delta x_h < 60$ Å) for the 10- and 20-Å QW's than for the wider QW's.[18]

VI. DISCUSSION

The exciton diamagnetic shift in GaAs/$Al_xGa_{1-x}As$ QW's has been studied in Refs. 6–16. There is some disagreement about the experimental value of the diamagnetic shift in comparable QW's. Now, the diamagnetic energy is not quadratic in B for high magnetic fields; in PL, the diamagnetism can be reduced by exciton localization; and the Al fraction of the barriers may have some effect (see Duggan,[57] who has compared some of these experimental results with calculations for parabolic bands and perfect confinement of the QW wave functions). However, not all discrepancies can be resolved by taking these factors into account.

The sensitivity of the diamagnetism to the exciton reduced mass means that an accurate calculation of the diamagnetism requires a proper description of both the fourfold degenerate valence band and nonparabolicity. Recent calculations by Yang and Sham[40] and by Bauer and Ando[41] of the free-exciton diamagnetic shift have included a full treatment of the four-band Luttinger Hamiltonian.[38] As stated by Yang and Sham,[40] the agreement of their calculation with the experiments of Ref. 6 is poor. The results of Bauer and Ando[41] agree well with experiments of Ref. 12 on 120- and 180-Å GaAs/(Al,Ga)As QW's, although the agreement is best when the energy and position dependence of the effective masses is ignored and the 85:15 band-offset rule is used. Clearly more work is needed, both to understand how to calculate the diamagnetic shift when the hole is described by the Luttinger Hamiltonian, and, in particular, to resolve the apparent discrepancies between different experimental results.

The theoretical techniques used in Refs. 6–16, and in the present work, require the assumption of parabolic dispersion for the holes, in the plane of the QW, with an estimate of the in-plane hole mass m_{ph}. m_{ph} is calculated[37] to be somewhat smaller at $k=0$ than the heavy-hole mass m_h^*; however, due to the large nonparabolicity, the most suitable value for m_{ph} is not clear. The whole range of experimental values reported for GaAs/(Al,Ga)As QW's can be accounted for with choices of m_{ph} that are not implausible. This uncertainty about the value of the free-exciton diamagnetism limits the sensitivity of the diamagnetic shift of exciton PL as a technique for assessing the spatial extent of a localized exciton. This sensitivity could be improved by a better understanding of the free-exciton diamagnetism, either from theoretical considerations or from absorption-type experiments on the same samples.[17] Even without this progress, the diamagnetic shift is a valuable probe of exciton localization, particularly when combined with analysis of the phonon sideband of PL.

APPENDIX A: THE GAUGE IN WHICH THE ELECTRON DIAMAGNETIC ENERGY IS $\langle (e^2/2m_e^*)A^2 \rangle$

Consider first a system whose ground state $|0\rangle$ is bound. $|0\rangle$ is nondegenerate (apart from spin degeneracy) and therefore real, and H_I is an imaginary Hermitian operator. Thus

$$\langle 0|H_I|0\rangle = 0 . \tag{A1}$$

For a bound state, we may apply the results from Sec. II that were derived from Griffith's theorem. Thus, the diamagnetic energy ${}_0E_2$ is equal to ${}_0E_2^1$ if ${}_0E_2^2=0$; using Eq. (5), this condition implies that $\langle i|H_I|0\rangle=0$ for all $i\neq 0$.

$H_I|0\rangle$ has no projection onto any eigenstate of H_{00}, and so

$$H_I|0\rangle = 0 . \tag{A2}$$

We vary $\mathbf{A}(\mathbf{r})$ by taking a fixed $\mathbf{A}_0(\mathbf{r})$, and adding the gradient of an arbitrary function $\chi(\mathbf{r})$ [Eq. (2)]. The condition for ${}_0E_2^2=0$ can be expressed in terms of χ by writing Eq. (A2) in full:

$$\left[\psi_0\nabla\cdot\left[\frac{1}{m_e^*}\nabla\chi\right]+\frac{2}{m_e^*}(\nabla\chi)\cdot(\nabla\psi_0)\right] = -\frac{i}{\hbar}\left[\mathbf{p}\cdot\frac{1}{m_e^*}\mathbf{A}_0+\mathbf{A}_0\cdot\frac{1}{m_e^*}\mathbf{p}\right]\psi_0 . \tag{A3}$$

Following Eq. (6), this equation for χ can also be obtained as the Euler-Lagrange equation for

$$\delta\langle 0|(e^2/2m_e^*)(\mathbf{A}_0+\nabla\chi)^2|0\rangle = 0 ,$$

though with the present method we know that

$$\langle 0|(e^2/2m_e^*)(\mathbf{A}_0+\nabla\chi)^2|0\rangle$$

is not merely stationary but has reached the value ${}_0E_2$. From Eq. (A3),

$$\nabla\cdot\left[\frac{1}{m_e^*}\nabla(\psi_0\chi)\right]-\psi_0\chi\left[\frac{1}{\psi_0}\nabla\cdot\left[\frac{1}{m_e^*}\nabla\psi_0\right]\right]=-\frac{2}{\hbar^2}\phi , \tag{A4}$$

where

$$\phi(\mathbf{r})=\frac{i\hbar}{2}\left[\mathbf{p}\cdot\frac{1}{m_e^*}\mathbf{A}_0+\mathbf{A}_0\cdot\frac{1}{m_e^*}\mathbf{p}\right]\psi_0(\mathbf{r}) .$$

Substituting from the Schrödinger equation $(H_{00}-{}_0E_0)\psi_0=0$, Eq. (A4) becomes

$$-\frac{\hbar^2}{2}\nabla\cdot\left[\frac{1}{m_e^*}\nabla[\psi_0(\mathbf{r})\chi(\mathbf{r})]\right] -[eV_e(\mathbf{r})+{}_0E_0][\psi_0(\mathbf{r})\chi(\mathbf{r})]=\phi(\mathbf{r}) . \tag{A5}$$

Equation (A5) is an inhomogeneous form of the Schrödinger equation in $\psi_0\chi$, with ϕ as the "forcing" term.

$$\int d^3r\,\psi_0^*(\mathbf{r})\phi(\mathbf{r})=\frac{i\hbar}{2}\left\langle 0\left|\left[\mathbf{p}\cdot\frac{1}{m_e^*}\mathbf{A}_0+\mathbf{A}_0\cdot\frac{1}{m_e^*}\mathbf{p}\right]\right|0\right\rangle =0$$

as for Eq. (A1). Thus ϕ has no projection onto ψ_0 and so the solution to Eq. (A5) can be written as an eigenfunction expansion,

$$\psi_0(\mathbf{r})\chi(\mathbf{r})=\sum_{i(\neq 0)}\frac{\int d^3\mathbf{r}'\,\psi_i^*(\mathbf{r}')\phi(\mathbf{r}')}{{}_iE_0-{}_0E_0}\psi_i(\mathbf{r})+\lambda\psi_0(\mathbf{r}) . \tag{A6}$$

If $\langle (e^2/2m_e^*)A^2\rangle$ is to exist we cannot permit a solution of Eq. (A5) that grows exponentially as $r\to\infty$; thus Eq. (A6) is the most general solution, with λ as the only free

parameter. If the potential $V_e(\mathbf{r})$ in H_{00} has no infinite barriers, then the ground-state wave function $\psi_0(\mathbf{r})$ is nonzero everywhere; hence a solution for $\chi(\mathbf{r})$ exists. This solution is unique apart from an arbitrary additive constant arising from the λ term in Eq. (A6). The solution for $\mathbf{A}=\mathbf{A}_0+\nabla\chi$ is unique.

If there are regions from which ψ_0 is excluded by potential barriers of infinite height, the solution $\psi_0\chi$ to Eq. (A5) is also excluded, and χ is undetermined in these regions; but χ is uniquely determined, apart from the λ term, everywhere else, and this is sufficient for any practical purpose. If the infinite barriers in $V_e(\mathbf{r})$ are regarded as a limiting case of finite barriers, then $\chi(\mathbf{r})$ in the barrier regions depends on how the limit is approached.

In the derivation of Eq. (A3) we have assumed that $|0\rangle$ is a bound state. Consider now the case where the motion is free in one direction, the y direction, perpendicular to **B**. We have already found, in Sec. III B, the gauge for which the diamagnetic shift is equal to ${}_0E_2^1$. By direct substitution we can verify for this case that ${}_0E_2^2$ and the orbital part of ${}_0E_1$ vanish and that the vector potential $\mathbf{A}(\mathbf{r})$ satisfies Eq. (A3). Thus Eqs. (A3) and (A5) apply not only to a defect state of arbitrary extent in one direction perpendicular to **B**, but also to the limiting case where the motion in this direction is free.

We have demonstrated the existence of a vector potential for which the second-order Zeeman energy vanishes, and the diamagnetism is given by $\langle(e^2/2m_e^*)A^2\rangle$. The solution Eq. (A6) involves an eigenfunction expansion and is not simpler than the expression Eq. (5) for a nonvanishing second-order Zeeman energy. However, the existence of the solution implies that the *exact* diamagnetism can be found by variation of χ to minimize $\langle(e^2/2m_e^*)A^2\rangle$. This variational principle can be used to calculate the diamagnetism numerically; it requires direct knowledge only of the ground-state density $|\psi_0(\mathbf{r})|^2$ and does not make separate use of the Hamiltonian.

APPENDIX B: A GAUGE FOR THE DEFECTS OF FIGURES 2 AND 3

We present a gauge for the defects of Figs. 2 and 3 that has a low value of $\langle A^2\rangle$, dependent only on W in the limit $R/W\rightarrow\infty$, but not necessarily the lowest possible value. Each defect consists of elongated "arms." We choose **A** to be independent of z, with no z component; and by analogy with Fig. 1(a), to be locally similar to the Landau gauge, vanishing at the center of each arm and pointing along its larger dimension. Following Eq. (2), we first select a convenient $\mathbf{A}_0$, and then determine a suitable χ by considering its contours in the x-y plane.

For the defect of Fig. 2, $\mathbf{A}_0=(0,Bx,0)$, with the origin defined as the symmetry center (so that $\langle\mathbf{r}/m_e^*\rangle=0$, as in Sec. III D). In the region of the "vertical" arms (those parallel to the y axis), for $|y|>W$, we set $\chi=0$; and for the "horizontal" arms (parallel to the x axis) for $|x|>W$, we choose $\chi=-Bxy$. If the electron is confined to the defect by finite potential barriers, then $\langle A^2\rangle$ is insensitive to $\chi(\mathbf{r})$ sufficiently far from the defect, where the electron charge is exponentially small (in the limit of infinite barriers, $\langle A^2\rangle$ is independent of $\chi(\mathbf{r})$ outside the defect). Thus contours of χ that pass through the horizontal arms can be deformed to avoid the vertical arms. Within distance W of the origin, A^2 is comparable to the Landau gauge value and is independent of R. Thus, as $R/W\rightarrow\infty$, the contribution to $\langle A^2\rangle$ from this region tends to zero, with the charge density near the origin. The limiting value of $\langle A^2\rangle$ depends only on W, as required.

For the defect of Fig. 3, a suitable gauge can be expressed very simply. $\mathbf{A}_0=-\frac{1}{2}\mathbf{r}\times\mathbf{B}$ with the origin at the center of the ring. In the region of the defect charge, $\chi=-\frac{1}{2}(R-\frac{1}{2}W)^2B\theta$, where $\theta=\tan^{-1}(y,x)$, $-\pi<\theta<\pi$. Beyond the defect, the contours of χ can be deformed to pass through the gap in the ring, and can be closed.

[1] R. C. Miller and D. A. Kleinman, J. Lumin. **30**, 520 (1985).

[2] C. Weisbuch, R. C. Miller, R. Dingle, A. C. Gossard, and W. Wiegmann, Solid State Commun. **37**, 219 (1981).

[3] M. S. Skolnick, P. R. Tapster, S. J. Bass, N. Apsley, A. D. Pitt, N. G. Chew, A. G. Cullis, S. P. Aldred, and C. A. Warwick, Appl. Phys. Lett. **48**, 1455 (1986).

[4] C. Delalande, M. H. Meynadier, and M. Voos, Phys. Rev. B **31**, 2497 (1985).

[5] M. S. Skolnick, P. R. Tapster, S. J. Bass, A. D. Pitt, N. Apsley, and S. P. Aldred, Semicond. Sci. Technol. **1**, 29 (1986).

[6] J. C. Maan, G. Belle, A. Fasolino, M. Altarelli, and K. Ploog, Phys. Rev. B **30**, 2253 (1984).

[7] N. Miura, Y. Iwasa, S. Tarucha, and H. Okamoto, in *Proceedings of the 17th International Conference on the Physics of Semiconductors, San Francisco, 1984,* edited by J. D. Chadi and W. A. Harrison (Springer, New York, 1985), p. 359.

[8] N. Miura, S. Takeyama, and Y. Iwasa, in *Proceedings of the 18th International Conference on the Physics of Semiconductors, Stockholm, 1986,* edited by O. Engström (World Scientific, Singapore, 1987), p. 715.

[9] H. Sakaki, Y. Arakawa, M. Nishioka, J. Yoshino, H. Okamoto, and N. Miura, Appl. Phys. Lett. **46**, 83 (1985).

[10] D. C. Rogers, J. Singleton, R. J. Nicholas, C. T. Foxon, and K. Woodbridge, Phys. Rev. B **34**, 4002 (1986).

[11] M. Bugajski, W. Kuszko, and K. Reginski, Solid State Commun. **60**, 669 (1986).

[12] W. Ossau, B. Jäkel, E. Bangert, G. Landwehr, and G. Weimann, Surf. Sci. **174**, 188 (1986); and unpublished work.

[13] S. Tarucha, H. Iwamura, T. Saku, H. Okamoto, Y. Iwasa, and N. Miura, Surf. Sci. **174**, 194 (1986).

[14] A. Petrou, G. Waytena, X. Liu, J. Ralston, and G. Wicks, Superlatt. Microstruct. **3**, 133 (1987).

[15] W. Ossau, B. Jäkel, and E. Bangert, in *High Magnetic Fields in Semiconductor Physics,* Vol. 71 of *Springer Series in Solid State Sciences* (Springer-Verlag, Berlin, 1987), p. 213; W. Ossau, B. Jäkel, and G. Weimann, J. Lumin. **40&41**, 723 (1988).

[16] A. S. Plaut, J. Singleton, R. J. Nicholas, R. T. Harley, S. R. Andrews, and C. T. B. Foxon, Phys. Rev. B **38**, 1323 (1988).

[17] D. J. Mowbray, J. Singleton, M. S. Skolnick, N. J. Pulsford, S. J. Bass, L. L. Taylor, R. J. Nicholas, and W. Hayes, Superlatt.

Microstruct. **3**, 471 (1987).
[18]K. J. Nash, M. S. Skolnick, P. A. Claxton, and J. S. Roberts, Phys. Rev. B **39**, 5558 (1989), and unpublished work.
[19]K. J. Nash, M. S. Skolnick, and S. J. Bass (unpublished work).
[20]L. Eaves and D. P. Halliday, J. Phys. C **17**, L705 (1984).
[21]M. S. Skolnick, T. D. Harris, C. W. Tu, T. M. Brennan, and M. D. Sturge, Appl. Phys. Lett. **46**, 427 (1985).
[22]C. Weisbuch, R. Dingle, A. C. Gossard, and W. Wiegmann, Solid State Commun. **38**, 709 (1981).
[23]J. Singh and K. K. Bajaj, J. Appl. Phys. **57**, 5433 (1985).
[24]J. H. Neave, B. A. Joyce, P. J. Dobson, and N. Norton, Appl. Phys. A **31**, 1 (1983).
[25]P. A. Claxton, J. S. Roberts, J. P. R. David, C. M. Sotomayor-Torres, M. S. Skolnick, P. R. Tapster, and K. J. Nash, J. Cryst. Growth **81**, 288 (1987).
[26]C. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1963), Chap. 14.
[27]D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).
[28]G. Bastard, Phys. Rev. B **24**, 5693 (1981).
[29]J. S. Griffith, *The Theory of Transition Metal Ions* (Cambridge University Press, Cambridge, 1961), p. 434.
[30]E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957); see also Fig. 16 of S. Adachi, J. Appl. Phys. **58**, R1 (1985).
[31]For example, P. M. Mathews and K. Venkatesan, *A Textbook of Quantum Mechanics* (Tata McGraw-Hill, New Delhi, 1976), p. 130.
[32]M. Shinada and S. Sugano, J. Phys. Soc. Jpn. **21**, 1936 (1966).
[33]J. M. F. Gunn, Eur. J. Phys. **7**, 29 (1986).
[34]R. L. Greene and K. K. Bajaj, Solid State Commun. **45**, 831 (1983); R. L. Greene, K. K. Bajaj, and D. E. Phelps, Phys. Rev. B **29**, 1807 (1984).
[35]G. Bastard, E. E. Mendez, L. L. Chang, and L. Esaki, Phys. Rev. B **26**, 1974 (1982).
[36]Y. Shinozuka and M. Matsuura, Phys. Rev. B **29**, 3717 (1984).
[37]M. Altarelli, U. Ekenberg, and A. Fasolino, Phys. Rev. B **32**, 5138 (1985).
[38]J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); J. M. Luttinger, *ibid.* **102**, 1030 (1956).
[39]G. Bastard, Phys. Rev. B **25**, 7584 (1982).
[40]S.-R. Eric Yang and L. J. Sham, Phys. Rev. Lett. **58**, 2598 (1987).
[41]G. E. W. Bauer and T. Ando, Phys. Rev. B **37**, 3130 (1988).
[42]T. Ando and Y. Uemura, J. Phys. Soc. Jpn. **37**, 1044 (1974); B. Vinter, Phys. Rev. B **13**, 4447 (1976).
[43]P. J. Price, J. Vac. Sci. Technol. **19**, 599 (1981).
[44]Section 54(d) of W. E. Lamb, Phys. Rev. **85**, 259 (1952); this transformation is used in Refs. 11 and 47.
[45]J. J. Hopfield, D. G. Thomas, and R. T. Lynch, Phys. Rev. Lett. **17**, 312 (1966).
[46]M. S. Skolnick, K. J. Nash, P. R. Tapster, D. J. Mowbray, S. J. Bass, and A. D. Pitt, Phys. Rev. B **35**, 5925 (1987).
[47]O. Akimoto and H. Hasegawa, J. Phys. Soc. Jpn. **22**, 181 (1967).
[48]R. L. Greene and K. K. Bajaj, Phys. Rev. B **31**, 6498 (1985).
[49]For example, S. Yamada, A. Taguchi, and A. Sugimura, Appl. Phys. Lett. **46**, 675 (1985).
[50]C. K. Sarkar, R. J. Nicholas, J. C. Portal, M. Razeghi, J. Chevrier, and J. Massies, J. Phys. C **18**, 2667 (1985); L. G. Shantharama, R. J. Nicholas, A. R. Adams, and C. K. Sarkar, *ibid.* **18**, L443 (1985).
[51]D. C. Rogers, J. Singleton, R. J. Nicholas, and C. T. Foxon, in *High Magnetic Fields in Semiconductor Physics,* Ref. 15, p. 223.
[52]M. Watt, C. M. Sotomayor-Torres, P. D. Hatton, H. Vass, P. A. Claxton, and J. S. Roberts, Superlatt. Microstruct. **3**, 75 (1987).
[53]M. S. Skolnick, K. J. Nash, M. K. Saker, S. J. Bass, P. A. Claxton, and J. S. Roberts, Appl. Phys. Lett. **50**, 1885 (1987).
[54]M. K. Saker, M. S. Skolnick, P. A. Claxton, J. S. Roberts, and M. J. Kane, Semicond. Sci. Technol. **3**, 691 (1988).
[55]D. J. Mowbray, M. S. Skolnick, D. Lee, P. A. Claxton, and J. S. Roberts, Appl. Phys. Lett. **53**, 752 (1988).
[56]N. G. Chew and A. G. Cullis (unpublished work).
[57]G. Duggan, Phys. Rev. B **37**, 2759 (1988).