Origin of the Stokes Shift: A Geometrical Model of Exciton Spectra in 2D Semiconductors

Fang Yang, M. Wilkinson, E. J. Austin, and K. P. O'Donnell
Department of Physics and Applied Physics, John Anderson Building, University of Strathclyde, Glasgow G4 0NG, Scotland, United Kingdom
(Received 29 July 1992)

We consider the optical absorption and emission spectra of excitons in two-dimensional semiconductors disordered through interface fluctuations. These spectra show a universal behavior exemplified by the fact that the offset of the spectral peaks (the Stokes shift) is proportional to their linewidths over a range of at least 2 orders of magnitude. We introduce a topographical theory of the exciton spectra which models such behavior in terms of statistical properties of a Gaussian random function. The coefficient of proportionality between the Stokes shift and the exciton absorption linewidth is found to be $\gamma = 2/\sqrt{6\pi \ln 2} = 0.553$ by analysis and 0.6 by experiment.

PACS numbers: 71.35.+z, 78.55.Et

During the last decade, there have been many experimental studies of the optical properties of low-dimensional semiconductors. In particular, a large body of experimental work describes absorption and luminescence spectra of samples in which excitons are confined in thin sheets of one material which form two-dimensional quantum wells (QW) when embedded within another material of larger band gap. The prototypical system is the GaAs QW in $Al_xGa_{1-x}As$ [1]. Several other combinations are well characterized, including GaInAs wells in GaAs [2], ZnCdTe wells in ZnTe [3], and ZnSe wells in ZnS [4]. A common feature of these 2D systems is the large inhomogeneous broadening of the exciton transitions, compared with those in pure bulk materials, by random fluctuations in the well width. An inverse square dependence of the energies on the well width for single electrons and holes in wells of infinite depth is a result of elementary quantum mechanics, while fluctuations in well width are an inevitable consequence of crystal growth in finite time. A theory of exciton line broadening in terms of monolayer fluctuations has been developed from the work of Weisbuch *et al.* [5], in particular by Singh, Bajaj, and Chaudhuri [6], although the practice of inferring microstructure from luminescence line shapes has recently been criticized by Warwick *et al.* [7]. However, there is general agreement that the linewidth provides a "rough and ready" indicator of sample quality. A further feature of QW spectra is the shift of the peak of the emission spectrum below that of the absorption or (more usually) excitation spectrum. This "Stokes shift" is also widely used as a negative indicator of sample quality.

In this Letter, we introduce an elementary treatment of the optical spectra and of the Stokes shift in terms of a topographical model of the exciton energy distribution. We choose to ignore the detailed statistics of the microstructure of QW samples in favor of a general theory: the justification of this approach is as follows. Figure 1 compares the absorption (solid curve) and emission spectrum (dashed curve) of a ZnCdTe-ZnTe multiple QW sample [3]. The absorption is dominated by a large peak due to direct creation of excitons (hydrogenically bound electron-hole pairs). The luminescence peak is due to the recombination of some of these pairs. Similar examples abound in the literature. Figure 2 summarizes data from many such experiments [8–14] by plotting the Stokes shift against the linewidth (full width at half maximum) of the exciton absorption peak. We show that the striking near universality of the line shapes and shifts can be understood quantitatively by using a very simple model involving the topography of a Gaussian random function.

The inhomogeneous broadening is due to two types of disorder: fluctuations in the widths of the layers of lower band-gap material, and substitutional alloy disorder. We assume that the former effect is dominant: Evidence supporting this assumption comes from the fact that the inhomogeneous broadening in multilayers (typically a few tens of meV) is usually much larger than that in comparable bulk alloys (typically a few meV). The random

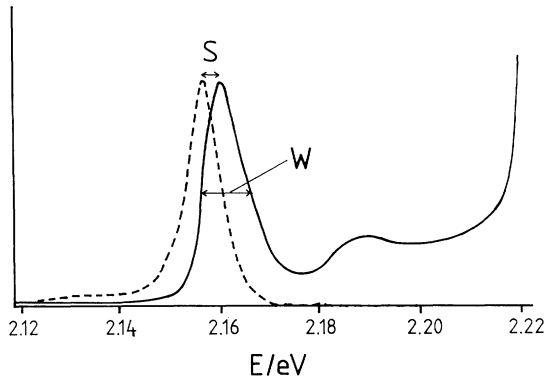

FIG. 1. The experimental optical absorption (solid line) and emission (dashed) spectrum of a ZnCdTe multiple QW, showing the Stokes shift S and absorption half-width W. (This diagram is reproduced with permission by the author [3].)

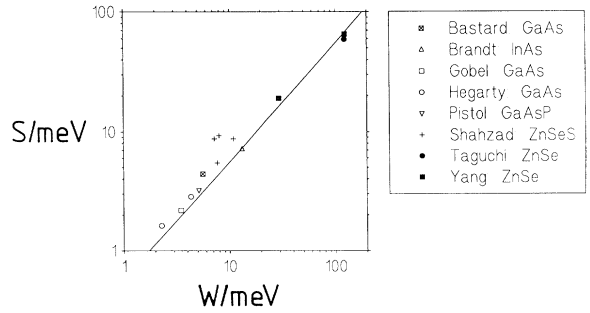

FIG. 2. A representative selection of data from the literature [8–14] compares values of S and W. The solid line has a slope given by the theoretical estimate $\gamma = 0.553$.

fluctuations in the widths of the potential wells cause a random variation in the energy of an exciton as a function of position in the plane of confinement, due to variations in the confinement energy. The energy $E(x, y)$ of a stationary exciton at position (x, y) contains a contribution $K/\bar{w}^2$ due to confinement in the z direction, where $\bar{w}(x, y)$ is the well width averaged over the spatial extent of an exciton centered at (x, y). Because of the random fluctuations in the well width $E(x, y)$ can be modeled by a random function. We assume that the effective well width $\bar{w}(x, y)$ has fluctuations over a length scale which is large compared to other relevant dimensions such as the size of the excitons: The good agreement between our theory and experiment will provide strong evidence supporting this assumption. Short-ranged fluctuations of the well width, on length scales smaller than the exciton radius, are suppressed because of the averaging over the range of the exciton wave function. The fluctuations in $E(x, y)$ are therefore dominated by the long-ranged fluctuations in the well width.

We will model the fluctuations of the effective well widths $\bar{w}(x, y)$ by a Gaussian random function. As well as facilitating the mathematical analysis, this is physically well motivated because the fluctuations of the well widths are the result of uncorrelated random events in the deposition of the multilayers: the central limit theorem implies that the sum of many random processes is Gaussian distributed. If the fluctuations $\delta\bar{w}$ of the effective well width are small compared to its mean value $\langle\bar{w}\rangle$, the fluctuations $\delta E = -2K\delta\bar{w}/\langle\bar{w}\rangle^3$ of the exciton energies $E(x, y)$ are proportional to $\delta\bar{w}$, and therefore also Gaussian distributed. We will assume a Gaussian distribution of the exciton energies throughout this paper.

The probability of creating an exciton is independent of position, implying that the absorption peak will be proportional to the probability distribution of the function $E(x, y)$. Most exciton absorption spectra, such as that of Fig. 1, can be fitted reasonably well by a Gaussian curve, providing experimental justification for the Gaussian random function model. The Stokes shift of the luminescence peak is due to the exciton losing energy by excitation of phonons before it decays. Because the exciton lifetimes are much larger than the relaxation time for emission of phonons [15], we assume that the exciton will have relaxed into a local minimum of the potential energy function $E(x, y)$ when it decays. The luminescence spectrum therefore reflects the distribution of heights of local minima of the Gaussian random function $E(x, y)$; in this Letter we assume that the luminescence spectrum is proportional to the unweighted distribution of heights of minima. A more detailed discussion of the model, together with a more refined analysis, will be published elsewhere [16].

The problem of counting minima of Gaussian random functions of one variable was considered by Rice [17] in an analysis of electrical noise. An extension to counting stationary points of a two-dimensional function was described by Longuet-Higgins [18,19] in the context of studies of the surface of the ocean. We now extend Longuet-Higgins' calculation to obtain the *distribution* of heights of minima of a Gaussian random function $f(x, y)$. We assume that the joint probability distribution $P(f, f_x, f_y, f_{xx}, f_{yy}, f_{xy})$ of the function f, its first derivatives $f_x = \partial f/\partial x$, $f_y = \partial f/\partial y$, and its second derivatives f_{xx}, f_{yy}, f_{xy} at a given point (x, y) is known. Let $\mathcal{N}(f)df$ be the number of minima of $f(x, y)$ per unit area with the height of the minimum between f and $f + df$. By a simple adaptation of the arguments leading to equation (2.4.13) of Ref. [18], we find the density of minima with height f can be written in the form

$$\mathcal{N}(f) = \int_{-\infty}^{\infty} df_{xx} \int_{-\infty}^{\infty} df_{yy} \int_{-\infty}^{\infty} df_{xy}\; P(f, 0, 0, f_{xx}, f_{yy}, f_{xy})\; D\; \Theta(D)\; \Theta(T)\;. \tag{1}$$

Here $D = \det(\tilde{M}) = f_{xx}f_{yy} - f_{xy}^2$ and $T = \mathrm{tr}(\tilde{M}) = f_{xx} + f_{yy}$ are the determinant and trace of the Hessian matrix $\tilde{M}$ of the function f, and $\Theta(x)$ is the unit increasing step function. The step functions are used to select regions where both eigenvalues of $\tilde{M}$ are positive, implying that the stationary point selected by setting $f_x = f_y = 0$ is a minimum.

We now consider how to evaluate this density of minima for a Gaussian random function with isotropic statistics. We can assume without loss of generality that the mean value of the function is zero. A Gaussian random function of two variables $f(x, y)$ with mean value zero can be generated by convolution of a white noise function $W(x, y)$ with a smoothing function $F(x, y)$:

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx'\; dy'\; F(x - x', y - y')\; W(x', y')\;. \tag{2}$$

The white noise function has the properties $\langle W(x, y)\rangle =$

0, $\langle W(x,y)W(x',y')\rangle = \delta(x-x')\delta(y-y')$, where the angle brackets denote either an ensemble or a spatial average. In the context of the problem considered here it is natural to assume that the statistical properties of $f(x,y)$ are isotropic: this can be obtained by choosing the smoothing function F to be a function of $r = (x^2+y^2)^{1/2}$ only, i.e., $F(x,y) = F(r)$.

In order to utilize (1) we will require the joint probability distribution of the function and its first and second derivatives. A standard result shows that the joint probability distribution of a set of N correlated Gaussian random variables $(X_1,...,X_N) = \mathbf{X}$ is given by

$$P(X_1,\ldots,X_N) = \left[(2\pi)^N \det(\tilde{C})\right]^{-1/2} \times \exp(-\tfrac{1}{2}\mathbf{X}^T\tilde{C}^{-1}\mathbf{X}) , \quad (3)$$

where $\tilde{C}$ is the covariance matrix, with elements $C_{ij} = \langle X_i X_j\rangle$. We can assume without any loss of generality that $\langle f\rangle = 0$, $\langle f^2\rangle = 1$, $\langle f_x^2\rangle = \langle f_y^2\rangle = 1$: these conditions can always be satisfied by scaling f, x, and y, respectively. It is convenient to choose the elements of the vector $\mathbf{X} = (X_1,\ldots,X_6)$ in (3) as follows,

$$X_1 = f,\ X_2 = f_x,\ X_3 = f_y,\ X_4 = \tfrac{1}{2}(f_{xx}+f_{yy}) ,\quad X_5 = \tfrac{1}{2}(f_{xx}-f_{yy}),\ X_6 = f_{xy} \quad (4)$$

in which case, after some elementary calculations, we find that the covariance matrix $\tilde{C}$ takes the nearly diagonal form

$$\tilde{C} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 2a & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a \end{pmatrix} , \quad (5)$$

where we have written $a = \langle f_{xy}^2\rangle$, and used the assumption of isotropy. Note that $D = X_4^2 - X_5^2 - X_6^2$ and that $T = 2X_4$, so that the region of integration selected by the step functions in (1) is a cone in X_4, X_5, X_6 space, with vertex at the origin, with its principal axis along the X_4 axis. In the X_i variables, only $X_1 = f$ and $X_4 = \frac{1}{2}T$ are correlated, and their joint distribution function is

$$p(f,X_4) = \frac{1}{2\pi\sqrt{2a-1}}\, e^{-f^2/2}\, \exp\left[-\frac{(f+X_4)^2}{2(2a-1)}\right] . \quad (6)$$

The density of minima is therefore

$$\mathcal{N}(f) = \frac{1}{(2\pi)^2 a}\int_0^\infty dR\, p(f,R)\int\int_{\mathcal{A}} dx\, dy\, (R^2-x^2-y^2)\, e^{-(x^2+y^2)/2a} , \quad (7)$$

where $\mathcal{A}$ is the disk $x^2+y^2 < R^2$, and for clarity we have changed the names of X_4, X_5, X_6 to R, x, y. The integral can be evaluated analytically using some results from Gradsteyhn and Ryhzik [20]: The result is

$$\mathcal{N}(f) = \frac{1}{(2\pi)^2\sqrt{2a-1}}\left[\frac{a\sqrt{2\pi a(2a-1)}}{\sqrt{3a-1}}\exp\left(\frac{-3af^2}{2(3a-1)}\right)\mathrm{erfc}\left(\sqrt{\frac{a}{2(2a-1)(3a-1)}}f\right)\right.$$
$$\left. -\tfrac{1}{2}\sqrt{2\pi(2a-1)}(f^2-1)\exp(-f^2/2)\mathrm{erfc}\left(\frac{f}{\sqrt{2(2a-1)}}\right) - (2a-1)f\exp\left(\frac{-af^2}{2a-1}\right)\right] , \quad (8)$$

where $\mathrm{erfc}(x)$ is the complementary error function [21]. The zeroth, first, and second moments of (8) are

$$M_0(a) = \frac{a}{2\pi\sqrt{3}},\quad M_1(a) = -\frac{8\sqrt{\pi a}}{3(2\pi)^2},$$
$$M_2(a) = \frac{1}{2\pi}\left(\frac{1+3a-1}{3\sqrt{3}}\right) . \quad (9)$$

The probability distribution of the heights of minima is obtained by dividing (8) by M_0 and the mean $\mu(a)$ and variance $V(a)$ of this distribution take the surprisingly simple forms

$$\mu(a) = -\frac{4}{\sqrt{3\pi a}} ,$$
$$V(a) = 1 + \frac{1}{a}\left(\sqrt{3} - \frac{16}{3\pi} - \frac{1}{3}\right) . \quad (10)$$

The distribution of heights of local minima therefore depends on the smoothing function $F(r)$ (or equivalently, on the correlation function or the power spectrum of the random function) through a single dimensionless parameter $a = \langle f_{xy}^2\rangle\langle f^2\rangle/\langle f_x^2\rangle^2$. We must now consider the physics of the problem in a little more depth in order to obtain an estimate for a.

It is clear that surface diffusion plays an important role in the growth of semiconductor multilayer samples of good quality, because a random deposition of atoms could never produce structures in which the well widths are defined to the monolayer accuracies obtained in the best samples. Equation (2) provides a simple model of the role of surface diffusion in smoothing out irregularities of the layers: $W(x,y)$ models the initial, highly irregular, pattern of surface deposition, $f(x,y)$ is the final smoothed layer width, and $F(x,y)$ is the diffusion kernel, which is Gaussian. This simple model therefore suggests the use of a Gaussian smoothing function, which implies that $a = 1$.

Equations (8)–(10) were checked by means of a numerical simulation using a Gaussian smoothing function $F(r)$. This simulation used a discretization of the integral relation (2) for generating a Gaussian random function; the value of the function $f(x,y)$ on a lattice point

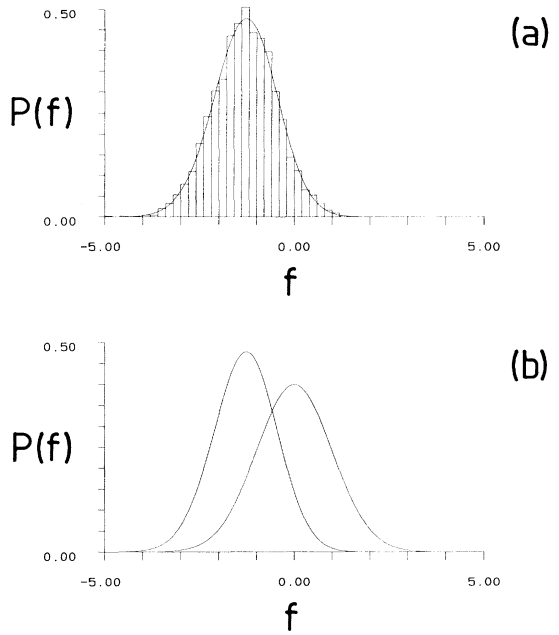

FIG. 3. (a) Theoretical distribution of heights of minima (8) for $a = 1$, compared with numerical simulation. (b) The theoretical distribution is plotted along side the Gaussian distribution of the function: These functions simulate, respectively, the luminescence and absorption peaks illustrated in Fig. 1.

$(x = n\delta, y = m\delta)$ is given by

$$f_{nm} = \delta^2 \sum_{n'} \sum_{m'} F[(n-n')\delta, (m-m')\delta]\, W_{n'm'} , \quad (11)$$

where the $W_{n'm'}$ are a set of uncorrelated random numbers with mean value zero and unit variance, and the lattice constant is $\delta \ll 1$. We used a Gaussian smoothing function $F(x, y) = (1/\pi)e^{-(x^2+y^2)/2}$. With this choice of F, we have $\langle f^2 \rangle = 1$, $\langle f_x^2 \rangle = 1/2$, and $a = 1$. We regard f_{nm} to be a minimum if it is less than the values of f at the eight nearest lattice points. The data in Fig. 3(a) show the results of a simulation in which the grid spacing was $\delta = 1/4$ and a total of 2×10^6 function values were computed, of which 5672 were found to be minima; the expectation value is 5743. The distribution of heights of minima shown in Fig. 3(a) fits the analytical curve very well: the mean and variance of -1.3104 and 0.7037 are in very good agreement with the theoretical values found by substituting $a = 1$ into Eq. (10). The distribution of heights of minima is plotted for $a = 1$ together with the original Gaussian function in Fig. 3(b): The similarity to the experimental excitonic absorption and emission curves is striking.

We finally obtain a numerical comparison with experiment by calculating the ratio of the Stokes shift S to the absorption full width at half maximum W. The constant of proportionality $\gamma = S/W$ is determined using the mean value of the distribution of heights of minima, given by (10), as our estimate of the Stokes shift: we find $\gamma = 2/(6\pi \ln 2)^{1/2}$ when $a = 1$. This constant has numerical value 0.553, while the best fit of a straight line to the experimental values shown in Fig. 2 gives 0.60.

In conclusion, we have characterized a near-universal relationship between the excitonic absorption and emission spectra of two-dimensional semiconductors. We have shown that a satisfactory account of this relationship is obtained from a consideration of purely classical localization of excitons in the minima of a plane Gaussian random potential. The comparison presented here assumes a Gaussian spatial correlation function, but other physically reasonable forms for the correlation function give similar values for γ [16].

The award of research grants by the U.K. Science and Engineering Research Council and the University of Strathclyde is gratefully acknowledged.

[1] R. Dingle, Festkoerperprobleme **15**, 21 (1975).
[2] J. Y. Marzin, M. N. Charasse, and B. Sermage, Phys. Rev. B **31**, 8298 (1985).
[3] R. P. Stanley, Ph.D. thesis, Trinity College, Dublin, 1991 (unpublished).
[4] Fang Yang, B. Henderson, K. P. O'Donnell, P. J. Wright, and B. Cockayne, Appl. Phys. Lett. **59**, 2142 (1991).
[5] C. Weisbuch, R. Dingle, A. C. Gossard, and W. Wiegmann, Solid State Commun. **38**, 709 (1981).
[6] J. Singh, K. K. Bajaj, and S. Chaudhuri, Appl. Phys. Lett. **44**, 805 (1984).
[7] C. A. Warwick, W. Y. Jan, A. Ourmazd, and T. D. Harris, Appl.Phys.Lett. **56**, 2666 (1990).
[8] G. Bastard, C. Delalande, M. H. Maynadier, P. M. Frijlink, and M. Voos, Phys. Rev. B **29**, 7042 (1984).
[9] O. Brandt, H. Lage, and K. Ploog, Phys. Rev. B **45**, 4217 (1992).
[10] E. Göbel, in *Excitons in Confined Systems* (Springer, Heidelberg, 1988).
[11] J. Hegarty, L. Goldner, and M. D. Sturge, Phys. Rev. B **30**, 7346 (1984).
[12] M-E. Pistol and X. Liu, Phys. Rev. B **45**, 4217 (1992).
[13] K. Shahzad, D. J. Olego, C. G. van de Walle, and D. A. Cammack, J. Lumin. **46**, 109 (1990).
[14] T. Taguchi and Y. Yamada, Mater. Res. Soc. Symp. Proc. **161**, 199 (1990).
[15] J. Hegarty and M. D. Sturge, J. Opt. Soc. Am. B **2**, 1143 (1985).
[16] M. Wilkinson, Fang Yang, E. J. Austin, and K. P. O'Donnell, J. Phys. Condens. Matter (to be published).
[17] S. O. Rice, Bell. Syst. Tech. J. **24**, 46 (1945).
[18] M. S. Longuet-Higgins, Philos. Trans. R. Soc. London A **249**, 321 (1957).
[19] M. S. Longuet-Higgins, Philos. Trans. R. Soc. London A **250**, 157 (1957).
[20] I. S. Gradsteyhn and I. M. Ryhzik, *Table of Integrals, Series and Products* (Academic, New York, 1980), English ed.
[21] M. Abromowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).