Near-Field Spectroscopy of a Gated Two-Dimensional Electron Gas

Thesis for the Ph.D Degree

by

Guy Eytan

Department of Physics

The Weizmann Institute of Science

Submitted to the Scientific Council of the Weizmann Institute of Science

July 1998

This research work was carried under the supervision of Prof. Israel Bar-Joseph in the department of physics at the Weizmann Institute of Science.

Acknowledgments

I wish to thank all those who helped me accomplish this research work. In particular, I thank Israel Bar-Joseph, my supervisor, who always provided help, encouragement, good advice and new idea when the project seems to come to a dead end. Special thanks I owe to Yossi Yayon who worked with me for the last critical year and a half as a master student. His joining to the project gave a strong push in all aspects of the research work towards its end. Michael Rappaport deserves a special thank for his extremely useful assistance in all technical issues concerning this research work, and especially for designing the mechanical part of the last version of the microscope. I enjoyed and learned a lot from long discussions with him on general experimental physics.

I want to thank all the students in our group, Shmil, Michael, Eran and the former students Gleb, Guy, Shimshon and Amit for their assistance, useful discussions and enjoyable lunch and coffee breaks.

I also want to thank Yehoda, Aron and all the peaple at the workshop for a lot of homemade mechanical parts they did. I am glad to thank Herman, Yossi, and Solomon for their quick response to my liquid helium requests.

I thank all the staff members of the sub-micron center for there contribution to this research work: Diana Mahalu, Hadas Shtrikman, Vladimir Umansky, Gregori Bunin, Sorina Neagu, Yoram Rotblat, Ofer Dudovitch, and Roby Tsabary. I want to thank all student that were in the sub-micron center during my work for creating a good and stimulating atmosphere.

Finally I wanted to thank my wife, Perach, who encouraged me at the desperate moments along this project and suffered patiently the long nights I have spent in the lab.

Abstract

In the last decade there was a tremendous advance in the new research field of scanning probe microscopy, and various kinds of scanning probe microscopes were developed. One of these microscopes is the near-field optical microscope, which break the resolution limits of conventional optical microscopy. In parallel a huge improvement in semiconductors growth techniques stimulated the research area of low-dimensionality systems, such as the two-dimensional electron gas. In this research work we combine these two research fields and investigate optically the two dimensional electron gas, using near-field optical microscopy.

The first two chapters of this work present the relevant scientific background to the two dimensional electron gas and its optical spectroscopy. Chapter 3 reviews the theoretical and experimental issues related to near field microscopy.

A considerable part of this research work was devoted to design, build and characterize a low temperature optical microscope. The development of the tip-sample interaction sensing method and the characterization of the tip performance are described in chapter 4. In chapter 5 we describe the mechanical structure of the microscope. In this chapter we also describe the optical system and a noise source, which we discover during our measurements. Also we describe in this chapter test measurements that were developed to characterize our experimental system.

In chapter 6 we describe near field photoluminescence measurements of a gated two-dimensional electron gas. We show that as the electron gas is depleted below a certain density, the electrons become singly localized at potential minima induced by the donor positive charge. We show how these electron density fluctuations can be used to image the donors random distribution. In addition we have found local temporal fluctuations in the potential landscape when illuminating the sample with high intensity. These temporal fluctuations are due to charge movements in the donors layer.

1.The two dimensional electron gas 8

1.1 Realization 8

1.2 2DEG characteristic 10

1.3 Localization in 2DEG 11

1.3.1 Localization in disordered medium 11

1.3.2 Scaling theory of localization 12

1.3.3 Experiments on 2DEG conductivity 14

1.3.4 Potential fluctuations and the conductivity drop in 2DEG 15

2. Optical properties of 2DEG 19

2.1 Introduction 20

2.2 PL of a high density 2DEG 22

2.2.1 Determination of the energy spectrum 22

2.2.2 Momentum conservation 24

2.2.3 The Fermi edge singularity 26

2.3 PL of a low density electron gas 27

2.3.1 The neutral exciton 28

2.3.2 Negatively charged exciton 29

3. Near field spectroscopy 31

3.1 Near field theory 31

3.1.1 Resolution limits of conventional optics 31

3.1.2 Braking the resolution limit by detecting evanescent waves: near-field approach 32

3.2 Semiconductors near-field spectroscopy experiments 36

3.2.1 Experimental configuration 36

3.2.2 Narrow peaks and inhomogeneous broadening 37

3.2.3 Spectroscopy of a single quantum confined device 39

4. Near field technology 41

4.1 Introduction 41

4.2 Tip performance and fabrication 42

4.3 Tip-sample distance regulation method 46

4.3.1 Electrical sensing method 46

5. Description of the microscope structure 53

5.1 Microscope overview 54

5.2 Microscope construction 54

5.3 Optical system description 57

5.3.1 NF signal as a function of tip height 58

5.3.2 Resolution measurements 59

5.3.3 Optical noise using tip illumination mode 61

6. Near field measurements of a gated 2DEG 64

6.1 Sample structure 64

6.2 Near field imaging of electron distribution using the X^- line 64

6.2.1 Non-uniformity of the local spectra 65

6.2.2 PL intensity fluctuations as a function of gate voltage 71

6.2.3 Fourier analysis of the fluctuations 73

6.2.4 Fluctuation amplitude 74

6.2.5 Modeling the localized system 76

6.2.6 Imaging the donors density 79

6.2.7 Fluctuation in the dense 2DEG phase 80

6.2.8 Temporal intensity fluctuations 81

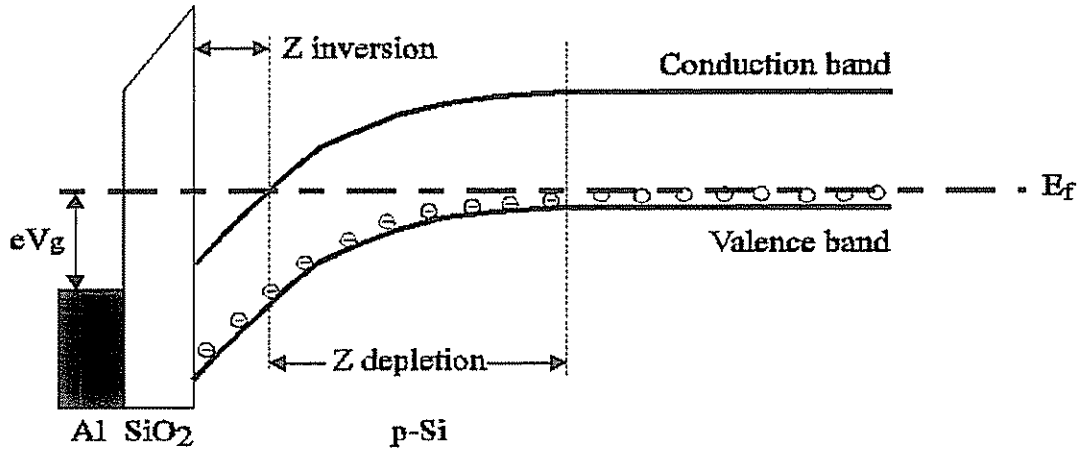
7. Summary 86

1 The two dimensional electron gas

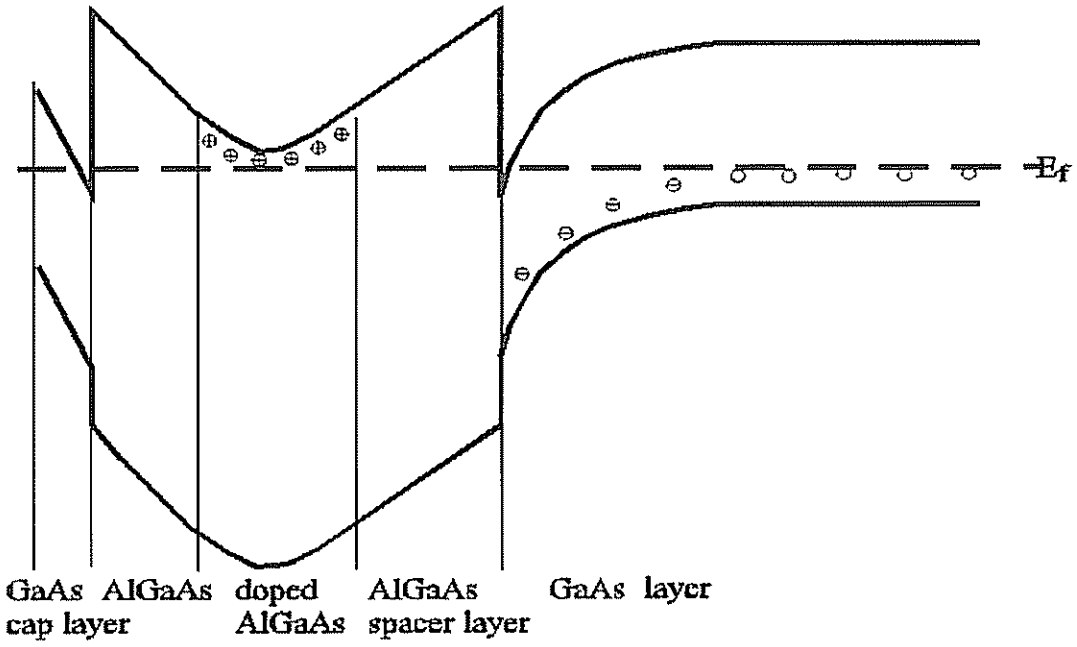
1.1 Realization

A two dimensional electron gas (2DEG) is formed between two different band gap material such as semiconductor and an insulator or two semiconductors. The most commonly used 2DEG based device is Si-SiO_2 MOSFET (Metal Oxide Semiconductor Field Effect Transistor) that was first realized and reported by Kahng and Atalla in 1960 [1]. This device consists of a p-doped Si substrate, a SiO_2 dielectric layer that is grown on top of it, and a metal gate on top of the SiO_2 layer. By applying a positive voltage to the gate relative to the substrate the Si bands bend down and a triangular quantum well is formed between the SiO_2 and the Si conduction band (see Fig. 1a). Electrons from the Si acceptors level flow into the well, creating a 2DEG in it. By varying the gate voltage the electrons density in the well can be controlled [2].

In 1978 Dingle and collaborators [3] introduced a new type of 2DEG, formed by molecular beam epitaxy (MBE) grown AlGaAs/GaAs heterostructure. A typical scheme of the heterostructure is shown in Fig. 1b.



(a)



(b)

Fig. 1: Energy band diagram of (a) Si MOSFET and (b) AlGaAs/GaAs based 2DEG

Since the GaAs conduction band edge lies lower in energy than the donor level in the AlGaAs, electrons from the donors move to the GaAs region, leaving behind ionized donors. The electric field that develops between the positive ionized donors and the electrons in the GaAs bends the energy bands, and a triangular quantum well is formed at the AlGaAs-GaAs interface. The electrons at the GaAs region occupy the energy lowest level in this well creating a 2DEG. This process is stopped when the Fermi level is continuous throughout the heterostructure.

1.2 2DEG Characteristic

This heterostructure based 2DEG has a unique property, which distinguishes it from other 2DEG structures. The doping is introduced only in the high band-gap material (AlGaAs) and is separated from the 2DEG region by an undoped spacer layer. This type of doping is called modulation doping.

Modulation doping has a strong impact on the electron mobility, μ, in the 2DEG. In Si MOSFET 2DEG the main scattering sources, which limit the mobility are crystal discontinuity at the Si-SiO_2 interface and trapped impurities in the SiO_2 layer. Theses scattering sources limit the low temperature mobility to a few 10^4 $cm^2/Vsec$. In modulation doped MBE grown AlGaAs/GaAs the interface between the two semiconductors is atomically smooth and the lattice constant differs only by 0.3% [4], thus interface scattering is strongly suppressed. The main source of scattering is the positive ions at the doping layer, thus by increasing the spacer layer width the scattering rate can be reduced substantially. Typical low temperature mobility is a few 10^6 $cm^2/Vsec$, with the highest reported mobility being 14.4×10^6 $cm^2/Vsec$ [5]

The mobility translates into a mean free path ℓ, according to $\ell=(2\pi n)^{1/2}\mu\hbar/e$. Thus, for typical Si-MOSFET parameters (density of 10^{11} cm^{-2} and mobility of 10^4 $cm^2/Vsec$) ℓ=500Å. In contrast, the mean free path in a high mobility AlGaAs/GaAs device ($\mu=10^7 cm^2/Vsec$) exceeds 50 μm.

1.3 Localization in 2DEG

1.3.1 Localization in disordered medium

The concept of localization was first introduced by Anderson, [6] who showed that in a disordered system electrons could be localized: the electron probability amplitude is maximum at some localization center and vanishes far away from it.

Anderson's model contains mobile entities (electrons in our case), which are distributed at random energies within a range of ±W. Two nearest neighbors electrons at sites i and $i+1$ are connected via an overlap energy integral: $V = \int \phi_{i+1}^{*} T \phi_i d\vec{r}$, where T is an interaction between two neighboring electrons. The randomness of the system is manifested by the electrons energy distribution W. Anderson shows that if the ratio $\frac{W}{V}$ is above a critical value, which has to be calculated numerically, than the electron's wave function is localized.

When $\frac{W}{V}$ is smaller than this critical value, electrons states in the central region of the conduction band are extended. However, the disorder is manifested by localization of electrons near the band edge. Energy E_c, which is called the mobility edge, differentiates between the extended and the localized states [7]. The definition of the mobility edge is that in the limit of zero temperature the conductivity, σ, behaves as:

$$\left.\begin{array}{ll} \sigma > o & E < E_c \\ \sigma = 0 & E < E_c \end{array}\right\}.$$

At finite temperatures the conductivity in the localized regime occurs by thermally activated electrons. The activation can be to an energy above the mobility edge, or to another localized sites (hopping). The conductivity in this regime depends on the

temperature as $\sigma \propto \exp(-\Delta E/k_B T)^p$, where ΔE is the activation energy and the value of p depends on the exact conduction process and the dimensionality. A massive experimental effort provides a strong support for the theory of localization. A review of this subject can be found in Mott and Davies book [8].

1.3.2 Scaling Theory of Localization

One of the most influencing papers on the present understanding of electron localization was written by Abrahams *et al.* and introduced the scaling theory of localization [9]. The underlying idea was to predict the behavior of the conductance of a system as a function of its size and dimension. A review on the scaling theory of localization was written by Lee and Ramakrishnan [10] and by Imry [11].

The conductivity of the system, σ, is the conductance per unit volume. The conductance G, is the measured quantity and it depends on the system size and on σ. However, it is common to refer to the dimensionless conductance, g rather than the conductance itself G:

$$g = G/(e^2/h) \tag{1}$$

Evaluating g as a function of the sample size, L, there are two asymptotic limits, depending on the degree of disorder in the sample. If the scatterers concentration is small, hence the mean free path l is large as compared to the Fermi wavelength $\lambda_f = 2\pi/k_f$, the electrons are subjected to a relatively smooth random potential. In these conditions the conductivity can be expressed as the Drude conductivity ($\sigma = \frac{e^2 n \tau}{m}$). Substituting the 2DEG density $n = k_f^2/2\pi$ and the relaxation time $\tau = l/v_f = lm/\hbar k_f$ one gets: $\sigma = \frac{e^2}{h} k_f l$.

Since $k_f l$ does not depend on L, then for L much larger than l, g does not depend on L and Ohm's low is recovered:

$$g(L) \propto L^{d-2} = const >> 1 \qquad (2)$$

(for 2 dimensional case).

In the other asymptotic limit, states in the vicinity of the Fermi energy are localized, and transport can occur by electron's hopping from an occupied state to an unoccupied state of nearly the same energy. The relevant length scale in this regime is the localization length, ξ, which is a typical scale of the electron wave function extension. The probability for such hopping between two states at a distance r from each other is proportional to $\exp(-r/\xi)$. Thus, the conductance for a sample with size L much larger than ξ, decreases as:

$$g(L) \propto \exp(-L/\xi) << 1 \qquad (3)$$

which is very different from the Ohmic behavior found in the other limit.

The two asymptotic behaviors of g can be obtained using a single scaling function, β, defined as:

$$\beta \equiv \frac{d(\ln g)}{d(\ln L)} = \frac{dg}{dL}\frac{L}{g} \qquad (4)$$

β behaves asymptotically as:

$$\beta(g) = \begin{cases} d-2 & g \to \infty \\ const + \ln(g) & g << 1 \end{cases} \qquad (5)$$

These behaviors are independent of L and of the details of the system.

An important consequence of this result is that β is always negative in two dimensions. If the system is weakly disordered, with a large conductance, g_0, at some length, L_0, then by enlarging L the system flows along the scaling curve, such that g decreases until the strong localization regime, where $\beta \sim \ln(g)$, is reached. **Thus, there are no truly extended**

states in a two-dimensional system, and there always exists some localization length. It should be noted that this statement is valid when interactions between the electrons are neglected.

The localization length, ξ is defined as the length, where the scaling behavior of g changes from Ohmic [Eq.(2)] to localized [Eq.(3)]. It is found that

$$\xi \cong L_\phi \exp(k_f l) \tag{6}$$

where L_ϕ is the phase coherence length. For typical values of a 2DEG in AlGaAs/GaAs: $n=10^{11}$ cm^{-2}, $\mu=4\text{x}10^6$ cm^2/(V sec), one get $k_f=\sqrt{2\pi n}\cong 7\text{x}10^5$ cm^{-1}, and $\ell=k_f\mu\hbar/e\cong 2\text{x}10^{-3}$ cm. Thus, the exponent in Eq. (6) is 140. If we take $L_\phi \sim l$ we get a number which is much larger than the size of the universe ($\sim 15\text{x}10^9$ light years).

The dependence of g on the temperature, T, could be obtained by using $L_\phi \sim T^{-p/2}$, where p is a positive exponent. Thus, in the weak localization regime g should decrease logarithmically with T.

1.3.3 Experiments on 2DEG conductivity

Considerable experimental efforts were done to verify the scaling theory of localization. The systems studied were Si-MOSFET, In_2O_{3-x} [12], Pd films [13] and a 2DEG in AlGaAs/GaAs heterostructure [14]. These experiments distinguish between two distinct regimes, characterized by a different conductivity versus temperature dependence. In the high 2DEG density regime the conductivity decreases slowly as the temperature decreases. Below a certain 2DEG density this behavior changes to a stronger dependence: the conductivity decreases sharply as the sample temperature decreases. These two distinct behaviors are commonly attributed to a cross over from the weak localization regime at a high density, to a strong localization regime at low density. An estimation of the localization length in the two regimes is done in [14] and the continuity is verified.

Recent experiments on Si-MOSFET [15], on Si/SiGe [16] and on p-doped

GaAs/AlGaAs heterostructure [17] give however, different results: a metallic behavior, where the conductance **increases** as the temperature is lowered, is seen at a relatively high carrier density. These new results seem to contradict the scaling theory of localization and are not completely understood. A possible explanation is that in these materials the carrier-carrier interaction is relatively strong, thus neglecting these interaction is no longer justified. Indeed it was argued [18] that when carrier-carrier interaction is strong $\beta(g)$ becomes positive as g goes to infinity, thus, a metal-insulator (M-I) transition should exist in this type of samples.

1.3.4 Potential fluctuations and the conductivity drop in 2DEG

A crossover from the weakly to the strongly localized regime could also be observed when the mobility is measured at a constant low temperature for different densities. This kind of measurements were done on a 2DEG GaAs/AlGaAs system [19],[20]. The mobility is a measure for the random potential and it does not depend directly on the 2DEG density. Indeed, a relatively small decrease in the mobility is observed as the carrier concentration is decreased at a high carrier density. However, below a certain density, which depends on the sample parameters, the mobility starts to decrease faster and becomes strongly dependent on the 2DEG density.

This change in the mobility behavior of AlGaAs/GaAs 2DEG is explained as follows. The donors at the doped layer are distributed randomly during the growth process. Hence when ionized they induce a random potential at the 2DEG layer. A 2DEG with a high density screens these potential fluctuations, and its properties are almost insensitive to density changes. As the 2DEG density is lowered the screening becomes less effective, and the potential fluctuations grow, limiting the 2DEG mobility. The reduction in the electron mobility causes a further decrease of the screening ability of the 2DEG, and a run-away process develops. Thus, at a critical density, a small change in the 2DEG density is accompanied by a large change in the 2DEG mobility.

A quantitative estimation of the potential fluctuations at high and low 2DEG density was

done by Efros[21]. Following his derivation let us assume for simplicity that the donors layer is infinitely narrow (δ doping), and that the donors are randomly distributed and uncorrelated in the plane. The average donor density is denoted by N and the local deviation from N by N(r). [N(r) could be negative or positive].

$$\begin{aligned} \langle N(r) \rangle &= 0 \\ \langle N(r_1) N(r_2) \rangle &= N\delta(r_1 - r_2) \end{aligned} \tag{7}$$

The $\langle\ \rangle$ symbol means averaging over many donor configurations. Let us first assume that there is no screening. The random potential $F_b(r)$ in the 2DEG layer that is created by these fluctuations can be written in the form:

$$F_b(r) = \frac{e^2}{4\pi\varepsilon\varepsilon_0} \iint \frac{N(r')d^2r'}{\sqrt{(r-r')^2 + s^2}} \tag{8}$$

where s is the spacer width, ε is the GaAs dielectric constant (ε≈13) and ε_0 is the dielectric constant of the vacuum. $F_b(r)$ can be expressed as an inverse Fourier transform:

$$F_b(r) = \frac{e^2}{2\varepsilon\varepsilon_0} \iint d^2q N(q) \frac{e^{-qs}}{q} e^{-iqr} \tag{9}$$

Eq.(9) is an inverse Fourier transform of the transform of the convolution described in (8). N(q) is a Fourier transform of N(r), thus $\langle N(q) \rangle = 0$ and therefore also $\langle F_b(r) \rangle = 0$. In order to calculate $\langle F_b^2(r) \rangle$ we should calculate the Fourier transform of (7):

$$\left\langle C(q_1)C(q_2)\right\rangle = \int d^2r_1 d^2r_2 N\delta(r_1 - r_2)\exp(-i(\vec{q}_1\vec{r}_1 + \vec{q}_2\vec{r}_2)) =$$
$$= \int N d^2r_1 \exp(-i(\vec{q}_1 + \vec{q}_2)\vec{r}_1) = \frac{N}{(2\pi)^2}\delta(\vec{q}_1 + \vec{q}_2). \tag{10}$$

Substituting (10) into (9) yields:

$$\left\langle F_h^{\,2}(r)\right\rangle = \frac{1}{(2\pi)^2}\left(\frac{e^2}{2\varepsilon\varepsilon_0}\right)^2 N \int d^2q_1 d^2q_2 \delta(\vec{q}_1 + \vec{q}_2)\frac{\exp(-s(q_1+q_2))}{q_1 q_2}\exp(-i\vec{r}(\vec{q}_1 + \vec{q}_2))$$
$$\underset{\vec{q}_1=-\vec{q}_2}{=} \frac{1}{(2\pi)^2}\left(\frac{e^2}{2\varepsilon\varepsilon_0}\right)^2 N \int 2\pi q dq \frac{e^{-qs}}{q^2} = \frac{1}{2\pi}\left(\frac{e^2}{2\varepsilon\varepsilon_0}\right)^2 N \int_{q_{\min}}^{\infty} dq \frac{e^{-2qs}}{q} \tag{11}$$

$q_{min} \approx 1/L$ is the lowest possible q in the sample (L being the sample size). For $q>1/2s$ the integrand becomes negligible. For $1/L<q<1/2s$ the numerator in the integrand is roughly constant. Hence:

$$\sqrt{\left\langle F_h^2\right\rangle} = W\sqrt{\ln\left(\frac{L}{2s}\right)} \tag{12}$$

where $W = \frac{e^2}{2\varepsilon\varepsilon_0}\sqrt{\frac{N}{2\pi}}$.

For a typical value $N=10^{12}\text{cm}^{-2}$ W is $\approx$30meV, which is very large.

In order to account for screening of the fluctuation by the electrons of the 2DEG one should replace the q in the denominator of (9) and (10) by $q+q_s$, where q_s is the inverse screening length of the 2DEG[22],

$$q_s = \frac{e^2}{2\varepsilon\varepsilon_0}\frac{m}{\pi\hbar^2}(1 - \exp(-(\hbar^2\pi N/k_B T m))). \tag{13}$$

m is the effective mass and k_B is the Boltzmann constant. For cryogenic temperatures and a typical density range ($N>10^{10}$cm^{-2}) q_s do not depend on N and is~(5 nm)$^{-1}$ in GaAs. By substituting $q+q_s$ instead of q in (11) one obtains:

$$\left\langle F^2(r)\right\rangle = \frac{1}{2\pi}\left(\frac{e^2}{2\varepsilon\varepsilon_0}\right)^2 N \int_{q\min}^{\infty} \frac{e^{-2qs}}{(q+q_s)^2} q dq \tag{14}$$

Performing the integration yields:

$$\sqrt{F^2} = W\frac{1}{(2q_s s)}. \tag{15}$$

$q_s s$ is typically ~10, thus in the presence of screening the potential fluctuations are substantially smaller.

It can be seen from this calculation that both in (14) and (11), spatial frequencies that are higher than ~1/2s are exponentially small. Therefore, the fastest spatial frequency of these fluctuations is 1/2s. Moreover, the contribution of spatial frequencies below 1/2s is of the same order of magnitude (this is the reason to the slow logarithmic divergence of small q in (12)[21]).

These potential fluctuations were modeled numerically [23]. In that work screening by the metal gate is also included. It was found that the typical wavelength of the fluctuation is a few times larger then the spacer width, which agrees with the result that the potential fluctuations originate from fluctuation in the donors density rather than from a single donor. The amplitude of the potential fluctuations reduce from 18meV to 2.6meV as the electron density increases, which shows that even at high electron density the fluctuation are not completely screened.

2 Optical properties of 2DEG

In this part the optical properties of the 2DEG will be reviewed. The structure that is analyzed and measured has a small difference as compared with the structure shown in Fig. 1. An AlGaAs layer is grown prior to the GaAs layer such that a quantum well (QW) is formed in the conduction and in the valance bands. The width of this well is typically 300Å or less, and its main purpose is to prevent the photo-excited holes from diffusing into the GaAs substrate. Schematic diagram of the structure with QW is shown in Fig. 2.

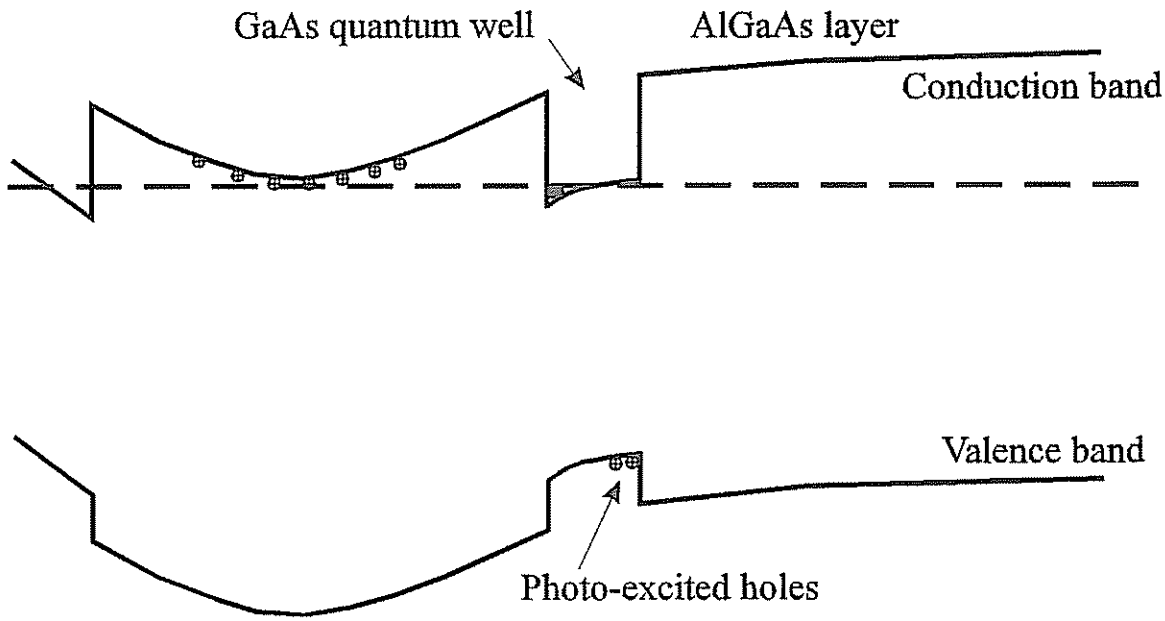


Fig. 2: Energy band diagram of 2DEG in quantum well.

2.1 Introduction

In a semiconductor an electron can be excited from the valance band to the conduction band, leaving a hole in the valance band. This electron–hole pair can recombine spontaneously (with no driving force), emitting a photon at a wavelength that corresponds to the energy that is lost in the process. This spontaneous light emission process is called luminescence. Different luminescence processes may be distinguished by the source of excitation. Here a photoluminescence (PL) is discussed, were the excitation is done by absorbing a photon. In a typical PL experiment the excitation light is at energy higher then the energy gap. The photo-excited electrons and holes decay by non-radiative processes to the lowest possible energy at each band, from which they recombine by emitting a photon. Consequently, PL experiments emphasize the lowest energy transition.

The excited semiconductor has an excess population of electrons and holes, hence it can't be regarded as in thermal equilibrium. However, since thermalization processes are faster than the recombination time, it is assumed that the electrons in the conduction band and the holes in the valance band are in thermal equilibrium among themselves and their energy distribution is described by the Fermi function.

The PL of 2DEG changes substantially as the 2DEG density is varied. To illustrate this we present in Fig. 3 a series of PL spectra, which are measured at different 2DEG densities. Figure 3a shows the spectrum of a high density 2DEG. The spectrum is broad and not symmetric: the low energy side is much more intense than the high energy side. At smaller density the spectrum becomes narrower (Fig. 3b) and more symmetric. The overall PL intensity of the two spectra is roughly conserved, however, in the low density spectrum the intensity is divided over a smaller spectral range, thus, the emission per energy interval is more intense (note the different scale in each figure). As the density decreases further (Fig. 3c) the line shape becomes symmetric. This line evolves into a sharp peak, (Fig. 3d) and then an additional peak appears at a higher energy (Fig. 3e). When decreasing the density further the high energy peak becomes dominant, while the intensity of the low energy peak decreases (Fig. 3f).

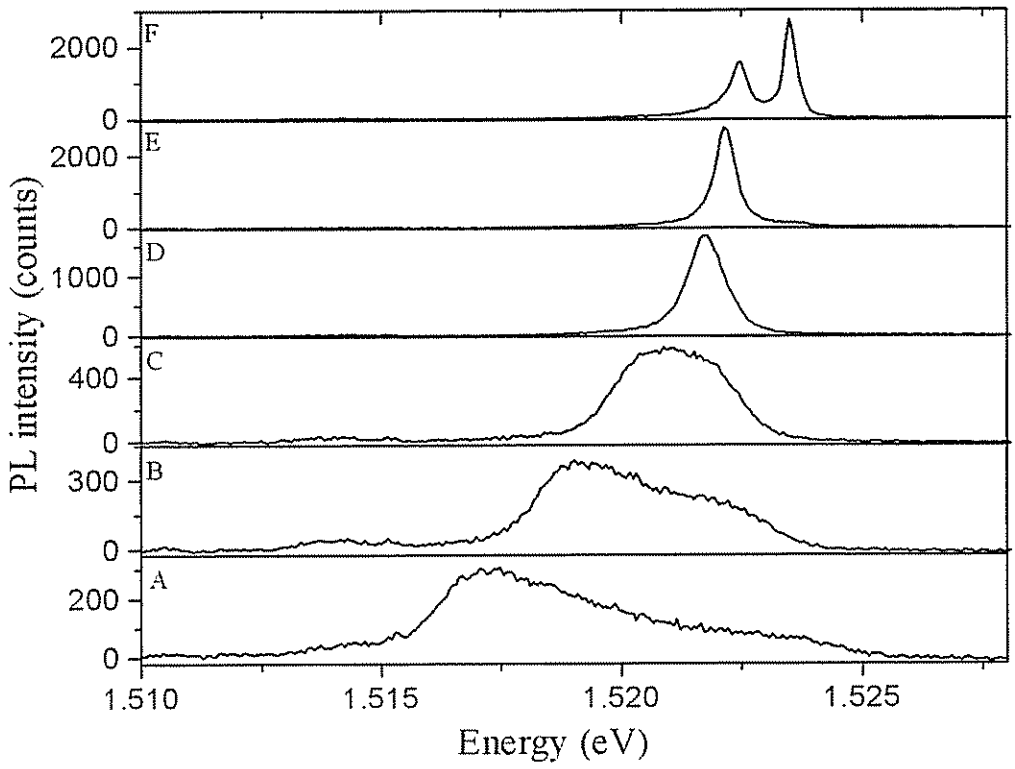


Fig. 3: The evolution of the PL spectrum from high density 2DEG (A) to depleted 2DEG (F). The measurements were collected on a gated 2DEG sample described at chapter 7. The gate voltage range was between 0 to –1 Volts.

It was recently discovered [24] [25] that the high energy peak is due to recombination of an exciton, which is a bound state of electron and a hole. The low energy peak was found to originate from a negatively charged exciton (X^-), a bound state of two electrons and a hole. In the following two sections the evolution of the spectrum from high density to excitons is explained.

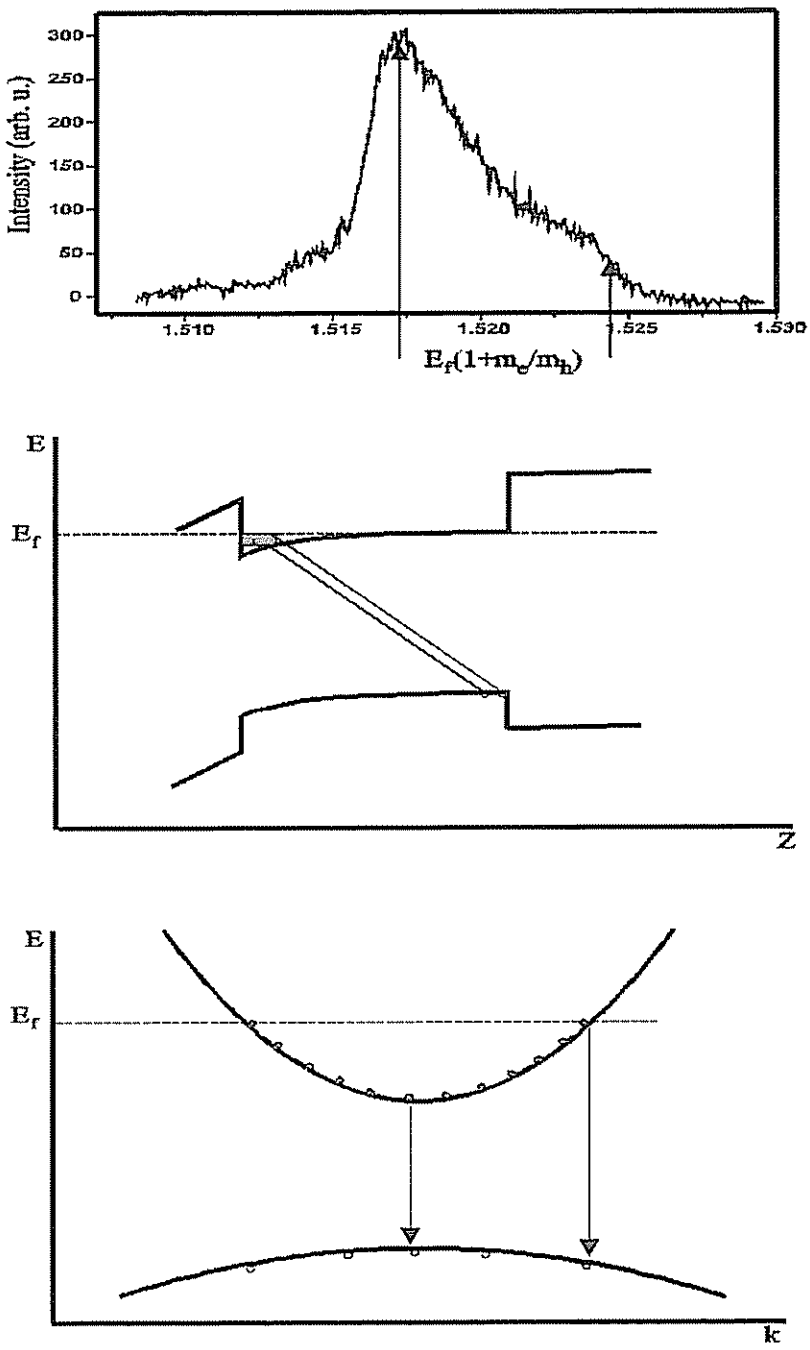


Fig. 4: Recombination processes of 2DEG

2.2.2 Momentum conservation

An important issue concerning the PL from a 2DEG is momentum conservation. The density of the hole gas, which is created by the excitation light, is typically much smaller then that of the density of the electron gas, which mainly consists of native electrons supplied by the donors. This density difference corresponds to momentum difference between the highest energy electrons and holes. The momentum difference is estimated as follows. The density of the photo-excited holes is:

$$n_h = \frac{I_L\left(1-e^{-\alpha z}\right)}{\hbar\omega_L}\tau_s . \qquad (16)$$

I_L is the excitation laser intensity, α is the absorption coefficient, z is the width of the absorbing layer, $\hbar\omega_L$ is the energy of the excitation photons, and τ_s is the spontaneous lifetime. Taking I_L=1W/cm^{-2} at a photon energy of ~1.96 eV (HeNe laser), $\alpha \cong 1\mu m^{-1}$, z= 200Å (the QW width in the experiment) and τ_s as 10^{-10} sec, one gets n_h= 6 10^6 cm^{-2}. This density corresponds to $k_f = \sqrt{2\pi n_h}$ =6.1x10^3 cm^{-1}. For a 2DEG density of 10^{11} cm^{-2} k_f=7.9x10^5 cm^{-1}. The momentum difference is therefore $\Delta p = \hbar\left(k_f - k_h\right) = 8\text{x}10^{-27}$ (kg m)/sec.

The emitted photon (at ~810 nm) has a momentum of $p = \hbar k = \frac{h}{\lambda} = 8\text{x}10^{-28}$ (kg m)/sec, and thus recombination between the highest energy electrons and holes can not conserve momentum. As is seen from Fig. 3a, the high energy part of the spectrum is indeed less intense than the low part, however, there is steel a considerable amount of recombination that seems to violate momentum conservation.

This phenomenon may be interpreted as follows. Assuming that the holes are distributed according to a Boltzman statistics, the observed width of the PL spectrum is E_f $(1+m_e/m_h)$, where m_e and m_h are the electron and the hole masses. Figure 5 shows a fit to

the PL intensity $I(\varepsilon)$ using Boltzman distribution for the holes and a Fermi-Dirak distribution for the electrons:

$$I(\varepsilon) \propto \exp\left(-\frac{\varepsilon_h}{k_B T}\right) \frac{1}{\exp\left[(\varepsilon_e - \varepsilon_f)/k_B T\right] + 1} \tag{17}$$

$\varepsilon = \varepsilon_e + \varepsilon_h + E_g$ is the PL energy, and k conservation implies: $\varepsilon_h = \varepsilon_e \frac{m_e}{m_h}$. This model was suggested in [29], and was tested for various temperatures. Good agreement between the model and the experimental results was obtained. The free parameters in this model are the mass ratio, the temperature and the Fermi energy. We used this model to fit the data in Fig. 3. In fitting the data T=5K and a mass ratio of 10 were used. As it is seen, there is a good agreement with the PL line shape By fixing these parameters, the Fermi energy can be evaluated for the various PL line shapes of Fig. 3 and yield the different densities (see Fig. 3).

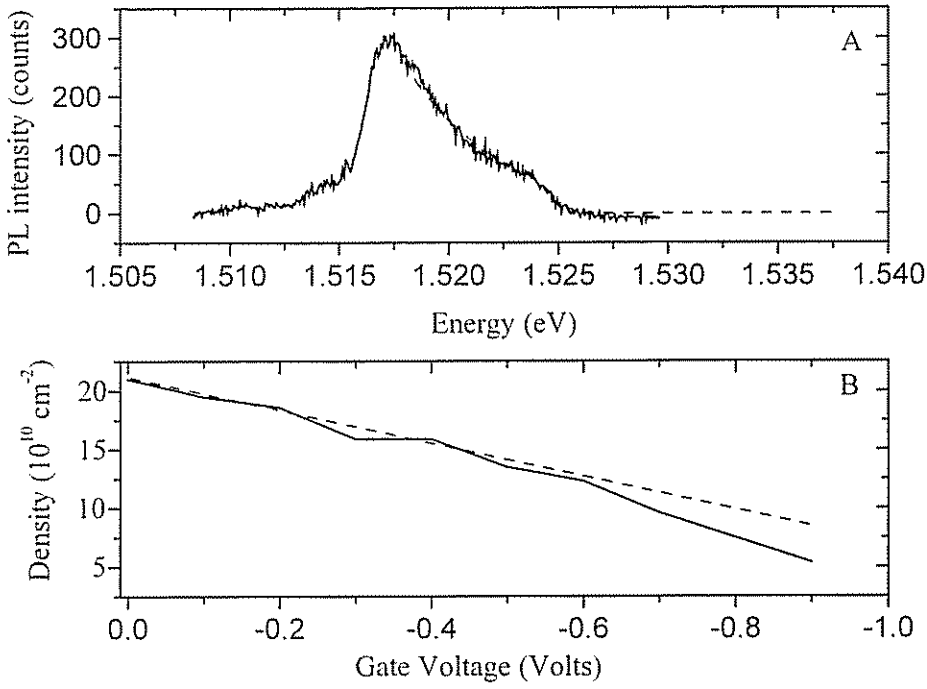


Fig. 5. A: Experimental (solid) and theoretical (dashed) of 2DEG PL. B: 2DEG density as function of gate voltage (solid) and linear fit (dashed).

In Fig. 5b the densities, which are obtained from the fits, are shown as a function of the gate voltage. As it is seen, there is a linear dependence at wide range of densities, as expected from electrostatic considerations. Deviation from the expected linear behavior is observed at low densities. At even lower gate voltages, the width of the PL peak is almost constant (Fig. 3e).

Thus, the above model explained well the PL line shape at the high-density regime. At lower densities it seems that this simple explanation fails.

2.2.3 **The Fermi edge singularity**.

The Fermi edge singularity is an enhancement of the optical transition

strength at energies approaching the Fermi level. This singularity was first observed in the X-ray spectrum of metals and explained by Mahan [30,31]. It was first seen in semiconductors in a InGaAs/InP doped QW [32]. The underlying physics of this phenomenon is the following. The Fermi sea reacts to the appearance of a hole by redistribution of the electrons around the Fermi energy, creating a high-density in the vicinity of the hole. In the limit of an infinite hole mass, recombination can occur at any k vector and in particular near k_f without violating k conservation. In reality, the effect is seen in (relatively) high roughness samples, where the holes are localized in a small region, hence they are almost infinitely heavy. It was argued that this picture changes gradually at low density to a bound state dressed with a collective excitation of the Fermi sea [27]. It is commonly assumed that the sharp peak, which appears in the intermediate density regime, represents such a state [34][37].

2.3 PL of a Low density electron gas

At a low density the single symmetric PL peak evolves into two excitonic peaks. One of them is a direct continuation of the previous single peak, while the other emerges at an energy, which is about 1meV higher. As the density is reduced further, the intensity of the low energy peak decreases and the intensity of the high-energy peak increases. The high-energy peak was found to be due to a recombination of a neutral exciton (X), a bound complex of electron and hole. The low energy peak is found to be due to a recombination of negatively charged exciton (X^-), which is a bound complex of two electrons and a hole [24][25].

The appearance of excitons in the presence of an electron gas is related to the 2D nature of the system. Unlike in three dimensions, where screening plays an important role in reducing the electron-hole interaction, in two dimensions this effect is less important. Excitonic interactions are strongly influenced by phase space filling [27]. This is demonstrated by the observation that in 2DEG the light and heavy hole excitons disappear together, while in a two dimensional hole gas the heavy hole exciton disappears

first. It can be shown [27] that the critical density for exciton appearance is $N_c \approx 1/\pi a_0^2$ where a_0 is the Bohr radius.

In the following a description of the X and X^- will be given, starting with the well documented neutral exciton and continuing with the recently found charged exciton.

2.3.1 The neutral exciton

The exciton is formed due to the electric attraction between the positive hole and the negative electron. Its size and binding energy are calculated by solving the Wannier equation:

$$-\left[\frac{\hbar^2\nabla_r^2}{2m_r}+V(\vec{r})\right]\varphi(\vec{r})=E_v\varphi(\vec{r}).$$

where m_r is the reduced electron hole mass: $\frac{1}{m_r}=\frac{1}{m_e}+\frac{1}{m_h}$. This equation becomes similar to the hydrogen atom equation by inserting the attractive Coulomb potential instead of V(r). The Wannier equation is limited to the range where V(r) changes very slowly compare to the lattice period. The solution should justify this assumption [22].

In 2D the energy levels of a hydrogen atom are quantized as:

$$E_n=-E_0\frac{1}{(n+1/2)^2},\quad n=0,1, \qquad \ldots (18)$$

where

$$E_0=\frac{e^4 m_r}{(4\pi\varepsilon\varepsilon_0)^2 2\hbar^2}. \qquad (19)$$

is the Rydberg energy, and

$$a_0 = \frac{4\pi\hbar^2 \varepsilon\varepsilon_0}{e^2 m_r} \tag{20}$$

is the exciton Bohr radius. In GaAs it gives: $a_0 \cong 100$Å and $E_0 \cong 5$meV. Thus the binding energy of the first energy level is $4E_0 \cong 20$meV. In reality, the exciton is not a true two-dimensional object, because of the finite width of the QW, and the typical values for the binding energy are 7-9meV [33].

2.3.2 Negatively charged exciton

The negatively charged exciton (X^-) was found only recently in CdZnTe/CdTe [34] and in AlGaAs/GaAs QW [24][25]. It was predicted [35] as the semiconductor analog to H^-, a negatively charged Hydrogen ion. The X^- is formed from two electrons and a hole, when a photo-excited electron-hole pair binds to a native electron in the conduction band. Hence, in order to observe the X^- it is necessary to have electrons in the conduction band. The binding energy of an electron to an exciton was predicted in [35] to be 0.03 E_0 (where E_0 is the Rydberg energy) in three dimensions. However, it was pointed out [36] that in two dimensions the binding energy is increased by factor of ~10. The resulting binding energy E_b in GaAs is ~1meV. Measurements of the binding energy dependence on the QW width and a general review on X^- can be found in [37]. The size of the X^- was estimated [37] to be twice as large as the neutral exciton.

The recombination process of the X^- is described in Fig. 6. While one of the electrons in X^- recombines with the hole, the other electron jumps back to the conduction band. In order to conserve energy the resulting luminescence peak appears at energy which is E_b below the X peak.

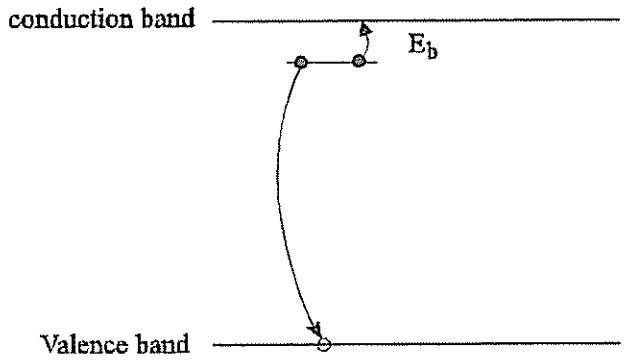


Fig. 6: Recombination process of negatively charged exciton

As mentioned before, a necessary condition for X^- to appear is an excess population of electrons in the QW. These electrons supply the additional electron needed to form the X^-, therefore X^- recombination is an indication for an electron that exist in the QW. Indeed, as the gate voltage becomes more negative there are less electrons in the QW, and thus the X^- intensity is reduced. However, in order for the electron-hole pair to bind to a native electron it is also necessary that this electron will be the **only** electron in the neighborhood of the excited pair. An additional electron near the hole with the same probability amplitude as the other two electrons, will unbind the X^- complex. This indicates that in order to form X^- the electrons wave function can't be extended, with an equal probability along the QW plain, but rather consist of different wave functions, with maximum probability amplitude at distinct locations. In another words, in order for the X^- to form, the electrons should be **localized**.

An attempt to measure the correlation between the drop of the conductivity and the appearance of the X^- peak was done in [24]. These measurements indeed found a strong correlation. In this work we conduct a **local** PL measurement of the X^- intensity, using a unique measurement technique that will be describe below. Strong fluctuations in the X^- intensity are observed when measuring the PL at different locations along the sample. This indicates that the X^- is indeed a local feature rather then an extended quantity. An extensive discussion of the measurement results is done in Chapter 6.

3 Near Field Spectroscopy

3.1 Near Field Theory

3.1.1 Resolution limits of conventional optics

Let us begin with a brief description of the classical imaging theory and its resolution limit. Consider a general imaging system, and presume that it has an entrance and an exit aperture. In the far field approximation it is assumed that the passage of light between the entrance and exit planes can be described in terms of geometric optics. Thus, the entrance and exit apertures are effectively the geometrical projection of each other [38].

Let us first consider a line source, located at the object plane. The light intensity of a line source at x=x_0 is described by : $f(x) = \delta(x - x_0)$. The angular spectrum of such a source is:

$$f(k) = \int_{-\infty}^{\infty} \delta(x - x_0) e^{ikx} dx = e^{ikx_0} . \tag{21}$$

When this object is imaged with an imaging system, whose objective accepts an aperture angle of 2θ, only spatial frequencies $|k_x|$ in the range $[0, k_{\max} = \omega n \sin(\theta)/c]$ are admitted by the apparatus, where ω is the light frequency and n is the refractive index of the medium between the object and the entrance pupil. The image that can be constructed, at the image plane, is given by:

$$f(x) = \int_{-k_{\max}}^{k_{\max}} e^{ikx_0} e^{-ikx} dk = \frac{2 \sin\left[\omega n \sin(\theta)(x - x_0)/c\right]}{x - x_0}. \tag{22}$$

Although the source is infinitely narrow, it is seen at the image plane as a light line with a finite width. The first zero of the sinc function is at

$$x - x_0 = \frac{\pi c}{\omega n \sin(\theta)} = \frac{\lambda}{2n \sin(\theta)}. \quad (23)$$

When having two line sources, the resolution given by the Rayleigh criterion (the minimum of one source is at the maximum of the second source) is:

$$\Delta x = \frac{\lambda}{2n \sin(\theta)}.$$

In the case of a point source and a circular aperture, the image will be an Airy disc. The resolution, which is then given by the width of the Airy disc function, is[39]:

$$\Delta x = \frac{0.61\lambda}{n \sin(\theta)}.$$

To go beyond this resolution limit, the imaging device should be able to collect higher spatial frequencies than $k_x = \omega n / c$. This is done by using Near Field Scanning Optical Microscope (NSOM)

3.1.2 Braking the Resolution Limit by Detecting Evanescent Waves: Near-Field Approach

As we saw above, it is commonly accepted that the resolution limit of an imaging system is of the order of the light wavelength. This is true if the optical aperture is positioned at least a few wavelengths from the object, and the aperture size is larger than the wavelength. These conditions are fulfilled so often in practice that they are usually ignored.

It is easy to understand the origin of these conditions using an analysis based on Fourier optics [40]. Consider a monochromatic optical field, E(x,z), propagating in the z direction, and its angular spectrum, $E(k_x,z)$. They are related as:

$$E(x,z) = \int_{-\infty}^{\infty} E(k_x,z)e^{-ik_x x}dk_x \,. \tag{24}$$

Since E(x,z) must satisfy the wave equation, the spectral amplitudes are known to satisfy the propagation relation :

$$E(k_x,z) = E(k_x,0)e^{ik_z z}\,, \tag{25}$$

and since

$$k^2 = k_x^2 + k_z^2 = \left(\omega/c\right)^2,$$

we get that

$$E(k_x,z) = E(k_x,0)e^{i\sqrt{k^2-k_x^2}z}, \tag{26}$$

where I assumed n=1.

If the field has spectral amplitudes, $E(k_x,z)$, which are different from zero for $k_x>\omega/c$, they will be evanescent, decreasing exponentially with increasing z. The low frequency components ($k_x<\omega/c$) do not damp out but propagate as plane waves, traveling at angles $\theta=\sin^{-1}(k_x\lambda/2\pi)$. The high frequency components ($k_x>\omega/c$) have significant contribution only when $z<<\lambda$, so to get resolution better than λ the aperture of the optical system must be much closer than λ to the object.

It is not enough to place the aperture close to the object. The imaging system should have also the ability to convert those evanescent waves into propagating waves, which will be seen far from the object. It follows from the theory of diffraction that a small object, illuminated with a propagating wave, generates diffracted evanescent waves. By applying the reciprocity theorem we find that a small object located in an evanescent field converts part of this field into propagating waves [41][42]. Therefore, in order to detect the

evanescent waves, that contain information about the small details of the object, one has

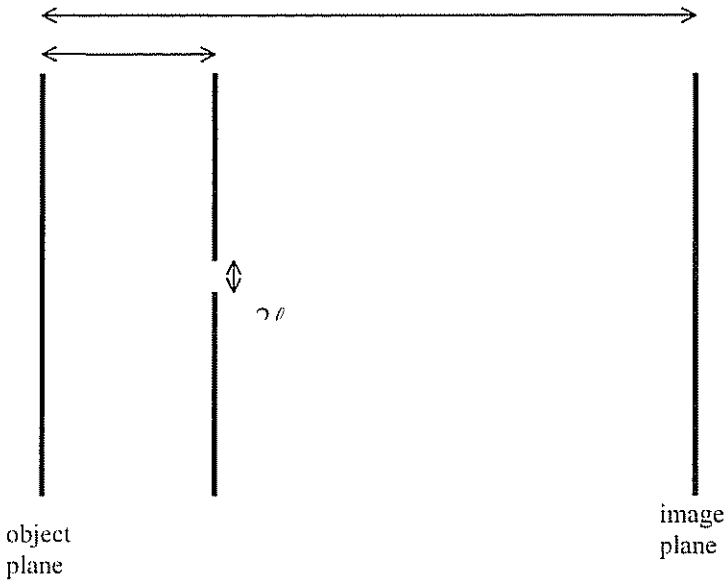


Fig. 7: Scheme of the near field detection method

to use a diffraction center of subwavelength size (the tip), and locate it at a subwavelength distance from the surface to be explored.

To study the resolution limit let us consider the device presented in Fig. 7. The tip is of width 2ℓ, and is placed at a distance ε from the object plane. We are interested at the field in a plane located at a distance z from the object. Let us denote the field at the object plane as $E_1(x,z=0)$ and its angular spectrum as $E_1(k_x,z=0)$. After propagating a distance ε, the field that is incident on the tip aperture will have the spectrum:

$$E_1(k_x, z=\varepsilon) = E_1(k_x, z=0)e^{i(k^2-k_x^2)^{1/2}\varepsilon} \tag{27}$$

The Fourier transform of the tip, centered at x=X, is

$$E_t(k_x) = 2\frac{\sin(k_x\ell)}{k_x}e^{ik_xX} \tag{28}$$

so just behind the tip the spectrum of the field will be:

$$E_2(k_x, z=\varepsilon) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk_x' E_1(k_x', z=\varepsilon) E_t(k_x - k_x') \tag{29}$$

As I showed before, at the far field, only components with $k_x<\omega/c$ will be seen, so the field at the image plane is given by:

$$E_2(x,z) = \frac{1}{(2\pi)^2}\int_{-\omega/c}^{\omega/c} dk_x e^{-ik_x x} e^{i(k^2-k_x^2)^{1/2}(z-\varepsilon)} E_2(k_x, z=\varepsilon), \tag{30}$$

and by substituting E_2,E_1 and E_t we obtain:

$$\begin{aligned} E_2(x,z) = {} & \frac{1}{(2\pi)^2}\int_{-\omega/c}^{\omega/c} dk_x e^{-ik_x x} e^{i(k^2-k_x^2)^{1/2}(z-\varepsilon)} \int_{-\infty}^{\infty} dk_x' E_1(k_x', z=0) e^{i(k^2-k_x'^2)^{1/2}\varepsilon} \\ & \times 2\frac{\sin\left(\left(k_x - k_x'\right)\ell\right)}{k_x - k_x'} e^{i(k_x - k_x')X}. \end{aligned} \tag{31}$$

We notice that, when the width of the tip in Fig. 7 becomes large ($\ell \to \infty$), we have:

$$\lim_{\ell\to\infty} \frac{2\sin\left(\left(k_x - k_x'\right)\ell\right)}{k_x - k_x'} = 2\pi\delta(k_x - k_x'), \tag{32}$$

and Eq. (31) tends towards the far field result :

$$\lim_{\ell\to\infty} E_2(x,z) = \frac{1}{2\pi}\int_{-\omega/c}^{\omega/c} dk_x e^{-ik_x x} e^{i(k^2-k_x^2)^{1/2} z} E_1(k_x, z=0). \tag{33}$$

To compare Eqs. (32) and (33) let us consider a single spatial frequency source:

$$E_1(k_x, z=0) = E_0\delta(k_x - K).$$

By substituting it into Eq. (33) we obtain:

$$\lim_{l\to\infty} E_2(x,z) = \frac{E_0}{2\pi}\begin{cases} e^{-iKx}e^{i(k^2-K^2)^{1/2}z} & K < \omega/c \\ 0 & K > \omega/c \end{cases}$$

This shows that in the far field limit all the spatial frequencies that are greater than ω/c are lost. By similarly introducing this light source into Eq. (31) we obtain:

$$E_2(x,z) = \frac{E_0}{(2\pi)^2} e^{i(k^2-K^2)^{1/2}\varepsilon} \int_{-\omega/c}^{\omega/c} dk_x e^{-ik_x x} e^{i(k^2-k_x^2)^{1/2}(z-\varepsilon)} 2\frac{\sin\left((k_x-K)\ell\right)}{k_x-K} e^{i(k_x-K)X}. \quad (34)$$

It can be seen that $E_2(x,z)$ is non zero, even for spatial frequencies $K>\omega/c$, so that these spatial frequencies contribute to the image, and information on the subwavelength details of the object are now received at the far field. Eq. (34) brings to light two characteristics of subwavelength detection:

1) Smaller tip sizes allow higher spatial frequencies to be detected. This corresponds to the fact that smaller tip sizes produce larger widths of the function $\mathrm{sinc}[(k_x-K)\ell]$. Generally speaking the tip can see anything that is larger than itself.

2) Smaller tip-sample separations allow higher spatial frequencies to be detected. This corresponds to the factor $e^{i(k^2-K^2)^{1/2}\varepsilon}$, which depends on the value of the spatial frequency K. The greater the value of K, the faster the amplitude of that component decay in the z direction.

3.2 Near-field spectroscopy experiments on semiconductors

3.2.1 Experimental configuration

As it is seen from the theoretical explanation, near-field (NF) experiments involve scanning the sample with a small aperture tip while the tip is in close proximity (typically ~100Å) to the sample surface. In reality the tip is made from an optical fiber that is tapered at its end and coated by metal, leaving only a small aperture open. The tip

can serve as a local light source or as a collector to the light emitted by the sample. In either case an additional far-field device should be introduced to the system. This device serves as an excitation source, if the PL light is collected by the tip or as a light collector, if the PL is excited through the tip.

One of the main difficulties in a semiconductor NF measurement is the low optical signal. This is due to the low throughput of the tip and to the small sample area that the tip interacts with. Also, compare to transmission or reflection measurements, the PL process is relatively inefficient. Therefore the excitation and collection configuration should be arranged such that the optical signal will be optimal.

Another non-trivial part in the experimental system is the mechanical part. The mechanics of the system should provide the means to bring the tip from far away to almost contact with the sample surface and stabilize it during the scan. The fact that the measurement takes place at cryogenic environment ($T \leq 4.2K^{o}$), further complicats the issue. The collected optical signal is dispersed by a spectrometer and detected by a high efficiency, low noise detector.

In the following two sections, a brief review of important NF semiconductor spectroscopy experiments done so far is presented. In the first section NF spectroscopy measurements of quantum wells are described. In these experiments the difference in the optical signal between separate locations arises from internal inhomogenuity of the sample. In the second section, measurements done on patterned samples such as a quantum wire are illustrated. In these experiment the main role of the NF measurement is to isolate a single quantum confined lateral structure.

3.2.2 Narrow peaks and inhomogeneous broadening

One of the first experiments in semiconductor NF spectroscopy was done by Hess and his collaborators [43]. In this work, the local photoluminescence (PL) of a narrow single AlGaAs/GaAs quantum well was measured. The far-field luminescence shows a broad emission line, which characterizes narrow quantum wells. The width of the line is

commonly attributed to an inhomogeneous broadening originating from well width fluctuations [44]. The near-field luminescence shows a collection of sharp peaks, located within the energy range of the far-field line. The width of these peaks is limited by the resolution of the spectrometer that was 0.07 meV. The number and the spectral location of these peaks change as the tip moves on the sample surface. The conclusion of the authors was that at such samples the luminescence originates from many small local emission sites that emit at different energies. These are created by fluctuations in the well width and form effective quantum dots. The far-field spectrum is an integration of all sites in the illuminated area, which yields the broad line.

The excitation method used was to illuminate the sample through the tip and collect the QW PL emission by a small f-number lens. This operation method had a major drawback: the resolution is decreased due to diffusion of the photo-excited carriers. Hence, the resolution is limited to $\sqrt{DT} \sim 1\mu m$, where $D \sim 3$ cm^2 sec^{-1} is the diffusion coefficient of holes in GaAs and $T \sim 10^{-9}$ sec is the recombination time.

The lack of spatial resolution didn't pose a significant limit on that measurement since the density of the luminescence sites were such that they could be spectrally distinguished in an area of $1\mu m^2$. The authors estimated the density to be 30-300 μm^{-2}. They found that this density is independent of the well width. These peaks were seen also in a high spatial resolution but diffraction limited measurement [45].

A near-field measurement, similar to that of Hess, was done on a ZnSe/CdSe quantum well [46]. This measurement distinguishes between two types of NF behavior, depending on the growth conditions. In one type of samples sharp peaks originating from zero dimensional confinement appear as the tip approaches the sample. These peaks merge into a broader line as the sample temperature is increased, due to a thermal activation of excitons out of the quantum dot (this effect was directly seen in a high spatial resolution PL measurement of a quantum wire [47]). In the second type of samples there was no significant difference between the far and the near field luminescence. This is explained by well width fluctuations that are on a smaller length scale then the exciton

diameter. These fluctuations weakly localize the excitons, but the density of states remains two dimensional.

3.2.3 Spectroscopy of a single quantum confined device

Considerable efforts has been devoted to produce quantum wires (QWR) and quantum dots, in which electrons are confined to one or zero dimension. These devices have promising technological applications and are at forefront of basic scientific research.

Regular optical measurements of these devices integrate over an array of such devices. In order to reveal the properties of a single device (and to investigate the a uniformity of quantum wire) a high resolution measurement should be performed.

The first attempt to scan a QWR was done by Grober et al. [48]. This measurement was also the first cold temperature NSOM measurement. The sample contained an array of AlGaAs/GaAs QWR, fabricated in the cleaved edge overgrowth technique [49]. The illumination technique that was used was to excite the sample through a lens that focuses the light on the sample. The tip was located between the lens and the sample. It was found that in this configuration the tip masks the area under it and the PL signal was small. In order to perform the measurement the tip was lifted to 1000Å above the sample, which suppressed the spatial resolution. This measurement was not capable of measuring a single QWR.

Spectroscopy of a single QWR was done later [50]. The wire was produced using the same method described in [49]. Both the excitation and collection were done through the NSOM tip. This method, although providing a good spatial resolution, has a drawback: it creates a high local density of photo-excited electrons and holes. (This is a very non-desired situation in 2DEG measurements. Furthermore we have tested this method and found that it generates additional noise, which will be discussed later.)

Richter *et al* measured a different type of QWR [51], fabricated by MBE growth on a patterned 311A GaAs substrate [52]. In this work, different sub-bands of the QWR

were resolved using PL excitation technique. The excitation was done using a far-field source. In this method a relatively large area (~1mm^2) of the sample is illuminated and the PL is collected by the tip, hence the resolution is determined by the tip diameter. This technique provides a uniform excitation of the sample and maintains a high spatial resolution. Its disadvantage is the weak signal that can be collected.

The emission spectrum of quantum dots (QD) was also measured [53] using NF technique. The QD fabrication process is described in [54]. A spatially resolved PL from the QD and from the surrounding region was measured. The QD size was estimated by scanning the area and detecting one of the QD recombination channels to be 300 nm.

Time and polarization resolved NF measurement was done on a patterned ZnSe/ZnCdSe QW in a magnetic field [55]. The polarization of the PL was measured as the tip advances towards the sample. It was found that the degree of polarization decreases sharply as the tip enters the NF zone. The authors attribute this to the fact that an evanescent field does not carry angular momentum.

4 Near field technology

4.1 Introduction

The invention of the scanning tunneling microscope (STM) in 1982 [56] opened the way for the development of various kinds of scanning probe microscopes (SPM), which the most familiar of them is the atomic force microscope (AFM) [57]. These microscopes sense the short range interaction between a probe (tip) and the sample. This interaction gives rise to a signal, which can be a tunneling current in an STM and attracting force in AFM. When the tip is scanned across the investigated sample at a closed proximity to its surface, it maps the strength of this interaction. The resolution of these microscopes is determine by the lateral scale over which the interaction occurs. This scale is determined by the size of the tip and by its height above the sample.

All SPM's have some common key features, which enable the imaging:

1. The scanning is done using a piezoelectric based scanner that translates an electric potential difference into a physical movement. The sensitivity of these type of scanners ranges between 10 to 1000 Å/V and depends on the size, geometry and the type of piezoelectric material. The most common geometry, which is used, is a tube [58] coated at its outer diameter by 4 electrodes parallel to the tube axis and one electrode at its inner diameter. By applying a voltage difference to opposite electrodes, the tube bends and gives rise to a lateral motion. Applying a voltage difference between the outer electrodes to the inner one, the tube extracts or contracts. The power supply to the piezo-scanner should have low noise and supply a voltage up to a few hundreds Volts.

2. In order to prevent the tip from crashing against the sample surface, a distance regulation mechanism is utilized. It is done by detecting the tip-sample interaction, the strength of which depends on the distance between them. The detected signal is incorporate into a feedback loop such that the tip-sample distance is adjusted in order

to keep this signal constant. In this way the tip scans the sample at a constant height, and follows changes in the sample topography.

3. The scanning is conducted at tip-sample distance of 10-100Å. Therefore prior to the scan, the tip should approach the sample from a distance of a few hundreds microns to the scanning distance. This mechanism is called coarse approach, and it enables to move the tip over a long distance with very small steps.

4. Since the tip-sample distance is very small, the scanning is very sensitive to a mechanical noise. The source of this noise could be acoustic or a mechanical vibration of the lab floor. To minimize this noise the microscope should be acoustically isolated and a few stages of vibration isolation should be incorporated. In addition, the microscope should be designed as rigid as possible, such that the microscope vibrates as a unit without relative motion of its separate components.

5. Since many of the interesting physical phenomena happen to occur at cryogenic temperatures, many of the SPM's were designed for this temperature range. In these microscopes thermal contraction should be taken into account. Also coarse manipulation of the sample and coarse approach is done from outside the cryostat. The scanning range in a low temperature SPM is smaller due to change in the piezoelectric constant with the temperature [59].

In addition to these points that are common to all SPM, there are some unique features that characterize NSOM. These points will be discussed below.

4.2 Tip performance and fabrication

The most important single component in an NSOM system is the tip that collects or emits light close to the sample, and delivers it to the detection system. The tip determines the resolution of the measurement and the intensity of the optical signal. Because of its importance to the operation of the system a large effort is made to improve its

performance and to reduce its diameter. In this work commercial tips manufactured by Nanonics have been used. However, some effort was devoted to developing a capability to produce tips using the etching method. This project is still under development.

A scanning electron microscope picture of an uncoated tip is shown at Fig. 8

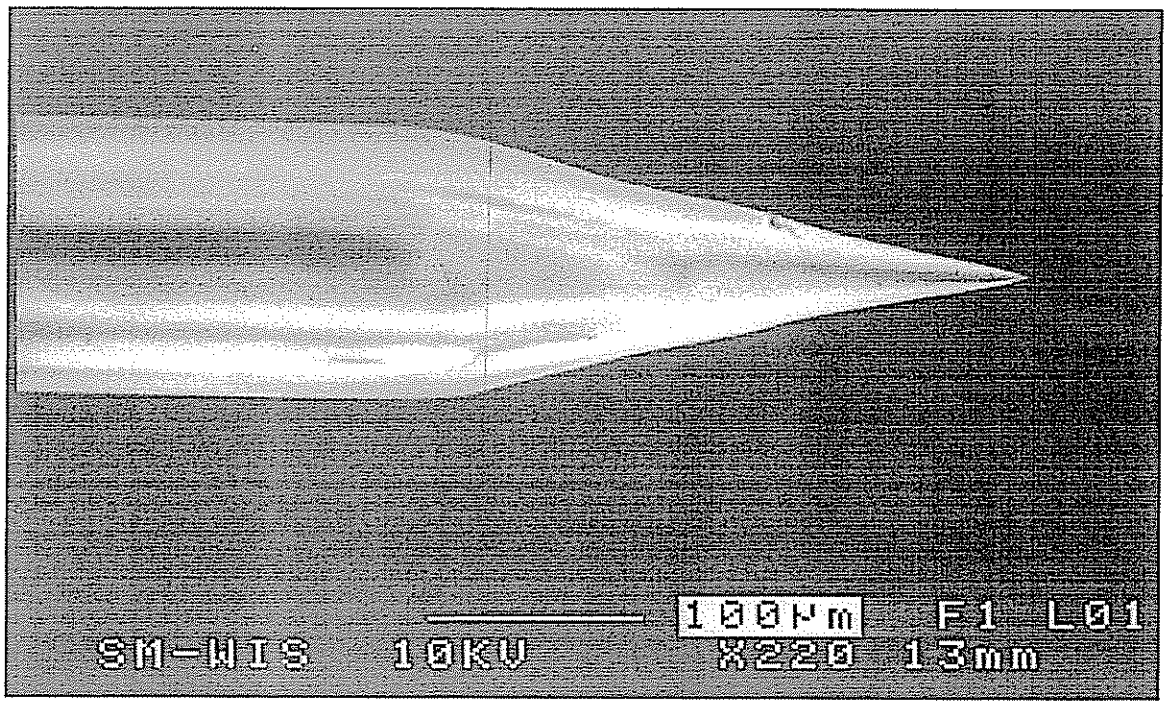


Fig. 8: SEM image of tip produced by the etching technique. The tip is not coated with metal.

The first tips were fabricated from a glass micropipette [60] or from a quartz crystal rod [61]. These tips were inefficient in transmitting light, as they reflected and absorbed most of the optical signal. Because of absorption the tips tend to warm and were limited in the optical power that can be coupled into them. Also the coupling of light into the tip was ineffective and it was an additional source of loss.

The current tip fabrication technology is largely based on a method published in 1991 by Betzig *et al.* [62]. In this method an optical fiber is used to deliver the light from a laser to the sample in illumination mode, or from the sample to the detecting system, in collection mode. The coupling of light to the fiber is easy, and it transmits light with very high

efficiency (with negligible loss in the visible and in the near IR). At its end the fiber is tapered until a small tip is formed at the end. The tapering area of the fiber is coated with Al, leaving only the aperture open. This basic structure, of tapered tip at the end of an optical fiber, is largely adopted by most of the NF community. The main effort since than is to try to improve the tip performance, meaning less attenuation at smaller diameters.

As it is described in Chapter 3, the evanescent light decays rapidly with the distance from the sample surface. Therefor, in order to minimize the losses in the tip it should broaden fast, such that the part were the light is evanescent will be as short as possible. For this reason there is a tendency to maximize the tip cone angle. However, it was found that there is no direct and simple relation between the cone angle and the attenuation. There is an additional important factor that influences the tip performance, the ratio between the core and the clad diameters at the tip. As the light penetrates into the tip, it propagates in the glass area of the tip, clad and core. The modes that propagate in the clad become leaky or are absorbed in the coating, thus it is important that the core to clad ratio is as large as possible.

Two main methods are used for tip fabrication: etching and pulling. In the pulling method [63] the fiber is heated and pulled. The heated area becomes thin until it brakes and a sharp tip is formed. The pulling can be done in two steps and various parameters, such as the heating rate and pulling velocity, influence the final shape of the tip. It is usually done in a commercial pipette puller. The advantage of the pulling method is that it creates a smooth surface that is easy to coat with metal. The disadvantage is that the tip cone angle is small at small tip diameters and that the core to clad ratio is constant along the tapering, hence at the tip end there is almost no core. The current performance of tips manufactured by the pulling technique is an attenuation factor of $10^{4\sim5}$ for a tip diameter of 100 nm at a wavelength of 514 nm.

The etching is done by immersing the fiber into an acid such as HF [64,65,66]. The HF surface is covered with some oil. As the etching progresses the length of the fiber, which is immersed, decreases and the fiber radius shrinks. The process terminates when the

fiber is completely out of the HF. The advantage of this method is that the tip cone angle is large and all the glass at the and of the fiber is of core material, thus light that penetrates into the tip is coupled efficiency to the core. The disadvantage is that the tip surface is rough and it is difficult to coat it with metal. The published tip performance for the etching technique is attenuation factor of 10^3 for tip diameter of 70 nm at a wavelength of 633 nm [67].

Table 1. shows typical parameters of a few commercial tips characterized by us using the pulling method.

TIP DIAMETER (nm)	CONE ANGLE	CORE DIAMETER (μm)	CORE TO CLAD RATIO	ATTENUATION AT λ=817 nm
490	24^o	8	0.064	10^4
380	30^o	8	0.064	10^4
550	31^o	8	0.064	$5\ 10^3$
270	34	8	0.064	200
200	22^o	50	0.4	100
400	34^o	50	0.4	50

Table 1: tip performance.

4.3 Tip-sample distance regulation method

As describe above, the tip to sample distance regulation is based on some interaction between them, which scales with the tip height above the sample. The method, which is used for distance regulation in NSOM, makes use of shear a force interaction between the sample and the tip. The tip is vibrated parallel to the sample surface, and as it approaches the surface the shear force interaction reduces the amplitude of the vibration. This reduction of amplitude can be detected and serves to regulate the tip-sample distance. The shear force (called also lateral force) scheme was first suggested for use in an AFM, [68] and was adopted by the NSOM community.

Many techniques have been proposed for sensing the tip vibration, most of tham make used of optical sensing. The early ones used the transmitted NSOM light, [69] or illuminated the tip from the side with two beams, and detected the phase difference between them [70].

A very popular optical sensing method, where a single beam is focused on the tip from the side, perpendicularly to the vibration direction, was proposed in [71]. The resulting diffraction pattern is detected by a position sensitive detector (PSD). The tip vibrations modulate the signal intensity, giving rise to an oscillating electrical signal at the output of the PSD, whose amplitude is proportional to the tip vibration amplitude. The major drawback of this method is the addition of optical excitation, which could interfere with the measurement. For low temperature measurement, like the one in this work, there is an additional difficulty of aligning this focused beam onto the tip [55].

4.3.1 Electrical sensing method

A few non-optical methods were proposed for sensing the tip-sample interaction. These methods are based on detection of a piezoelectric signal generated by the tip vibration. This signal can be measured on the piezo-tube itself [72, 73], or more efficiently on a

piezoelectric tuning fork, which is attached on one side to the piezo-tube and on the other to the tip [74].

We have tested some of the optical and electrical sensing methods and chose the tuning fork with a large quality factor method [75]. As shown in Fig. 9, the optical fiber, with its aluminized tapered tip, is glued along the side of one of the prongs of a quartz crystal tuning fork. Such tuning forks are commercially available for operation at a frequency of 32768 Hz. In order to excite the mechanical resonance of the fork, it is rigidly mounted on a ceramic piezoelectric tube serving as a dither piezo. The tuning fork and the tip are vibrated parallel to the sample surface at the resonance frequency of the joint system (tip and fork). The bending of the prongs generates an oscillating piezoelectric potential proportional to the tip oscillation amplitude. The prong holding the tip is a cantilever of length L and rectangular cross-section wt. It bends periodically along the X direction (see Fig. 9)

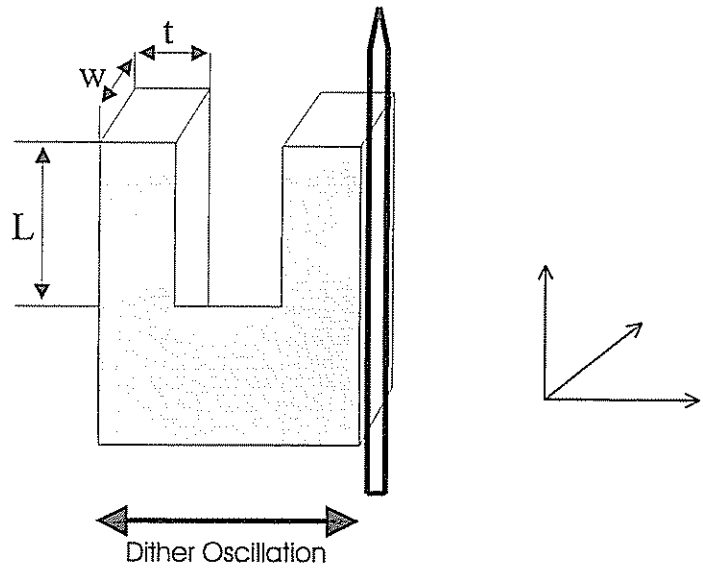


Fig. 9: Diagram of a crystal quartz tuning fork with a fiber tip glued along one of its prongs. The dither vibrates the whole device.

The parameters needed to determine the dynamics of the prong are its dimensions L=4 mm, w=0.4 mm, t=0.6 mm, its Young modulus E=7.87×10^{10} N/m^2, and its density 2650 kg/m^3. The time dependence of the tip position can be described by an effective harmonic oscillator equation of motion driven at a frequency w:

$$m\ddot{x} + m\gamma\dot{x} + kx = Fe^{iwt}. \tag{35}$$

where m=0.2427(Ltw) is an effective mass, corresponding to about ¼ of the mass of one of the prongs. F is the amplitude of the driving force, tuned by adjusting the voltage applied on the dither piezo tube. $k=Ewt^3/(4L^3)$ is the static compliance of one prong, and $m\gamma\dot{x}$ is a phenomenological viscous force, which is the sum of all drag forces acting on the tip. The solution for the equation of motion [Eq. (35)] is:

$$x = \frac{F}{m(\omega_0^2 - \omega^2 + i\gamma\omega)} e^{i\omega t}.$$

The amplitude of this solution is given by the Lorenzian:

$$A = \frac{F}{m\left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega_0^2\right]^{1/2}}, \tag{36}$$

where I assumed $\omega \approx \omega_0$. This assumption is justified because $\gamma << \omega_0$, so the Lorenzian full width at half maximum, $\Delta\omega = \sqrt{3}\gamma$, is small compared to ω_0, and we are interested only at frequencies near ω_0.

On resonance [$\omega=\omega_0=(k/m)^{1/2}=1.0150t(E/\rho)^{1/2}/L^2$] the tip amplitude, found from Eq. (36), is:

$$A_0 = \frac{F}{m\gamma\omega_0} = \frac{\sqrt{3}FQ}{k} \tag{37}$$

where the quality factor is defined as $Q=\omega_0/\Delta\omega$. A typical tip amplitude is less than 1 nm,

so it doesn't damage the resolution, which is ~100 nm. The drag force experienced by the fork, which is equal to the driving force, is given by:

$$F_D = m\gamma\dot{x} = m\gamma i\omega_0 x\left(\omega = \omega_0\right) = Fe^{i\omega_0 t}.$$

It is important to apply a low drag force on the tip to prevent damage to it. To calculate the amplitude of the drag force we use Eq. (37) to get:

$$\left|F_D\right| = F = \frac{kA_0}{\sqrt{3}Q}.$$

A_0 is determined from the electric signal of the fork, according to the calibration of 27μV/nm at room temperature and about one tenth of it at 4 K. Q is found from the resonance curve (Fig. 10). the resulting typical drag force, acting on the tip in our system, is a few tens of nN. We reduce the drag force by reducing the driving force to a minimum value, which still gives a measurable piezoelectric signal.

A typical resonance of the piezoelectric signal amplitude is shown in Fig. 10. The solid line is measured with a free fork at a room temperature, the dashed line is measured with a fork that a tip is attached to it at room temperature, and the dotted line is the same as the dashed line but at a low temperature (4 K). The free fork amplitude is divided by 4, and the low temperature amplitude is multiplied by 3 for viewing convenience. One can see that there is a significant reduction of the quality factor after attaching the tip to one of the prongs of the tuning fork, One can also see the change in amplitude and frequency of the resonance when cooling the NSOM.

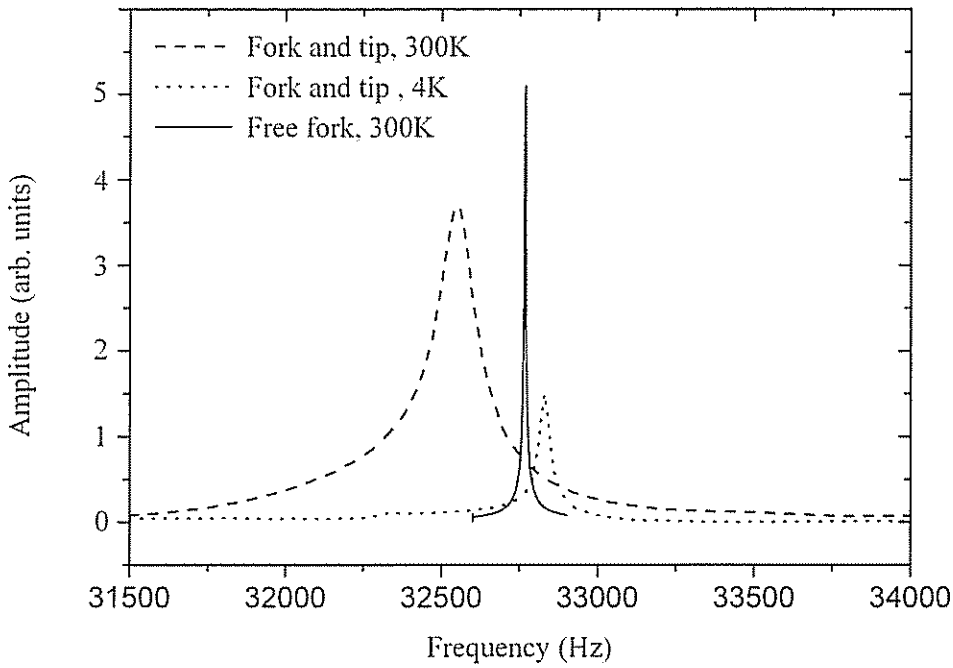


Fig. 10: Resonance curves of the tuning fork. The dither piezoelectric tube is driven with a signal of 100 mV. At 4K the modulation amplitude is reduced by a factor of ~10.

When used for feedback, the fork is driven at resonance. The oscillating signal, picked up between the two contacts of the tuning fork and measured using a lock-in amplifier synchronized with the dither frequency, is monitored by the feedback system as the tip approaches the sample. A reduction of the amplitude of that signal is measured as the tip is within ~20 nm of the sample surface. This change in the signal is shown in Fig. 11

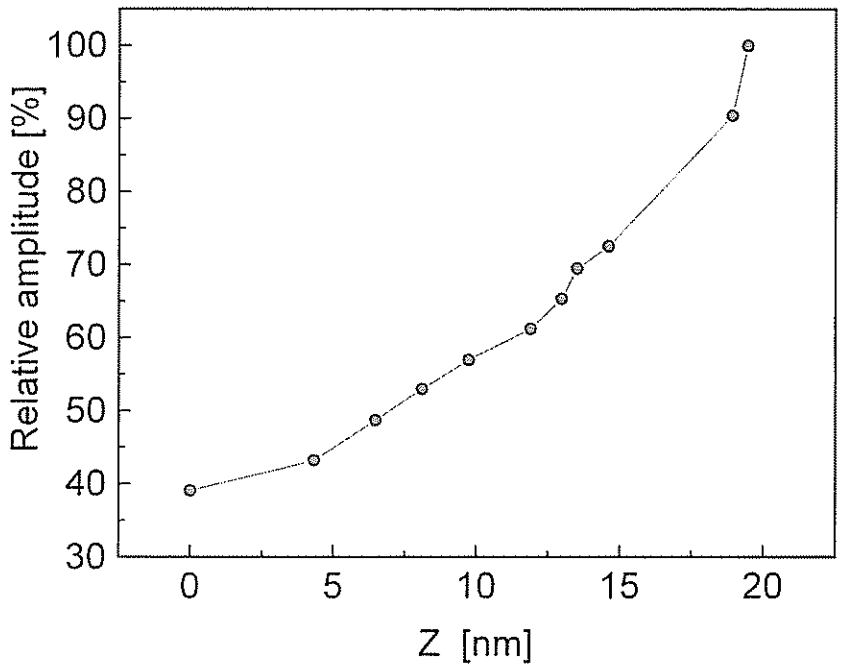


Fig. 11: Approach curve of the normalized signal amplitude as a function of the tip-sample distance. The measurement has been made at temperature of 4 K.

Such piezoelectric signals can be used in conjunction with an electronic feedback loop for keeping the tip at a constant distance from the sample surface (~ 5-20 nm). If this signal is smaller/higher than the set point of the feedback system, then the piezoelectric tube, to which the sample is attached, shrinks/stretches until the signal is equal to the set point and the tip is brought to the correct distance from the sample. This distance is determined by the set point.

The signal of the piezoelectric tube can be used for surface topography imaging, as demonstrated in Fig. 12. The structure was fabricated by etching of GaAs mesa lines. The depth of the etching is 300 nm, the width of the lines is 6 μm and the period is 10 μm. The structure was scanned, perpendicular to the mesa lines (in the Y direction), in steps of 100 nm.

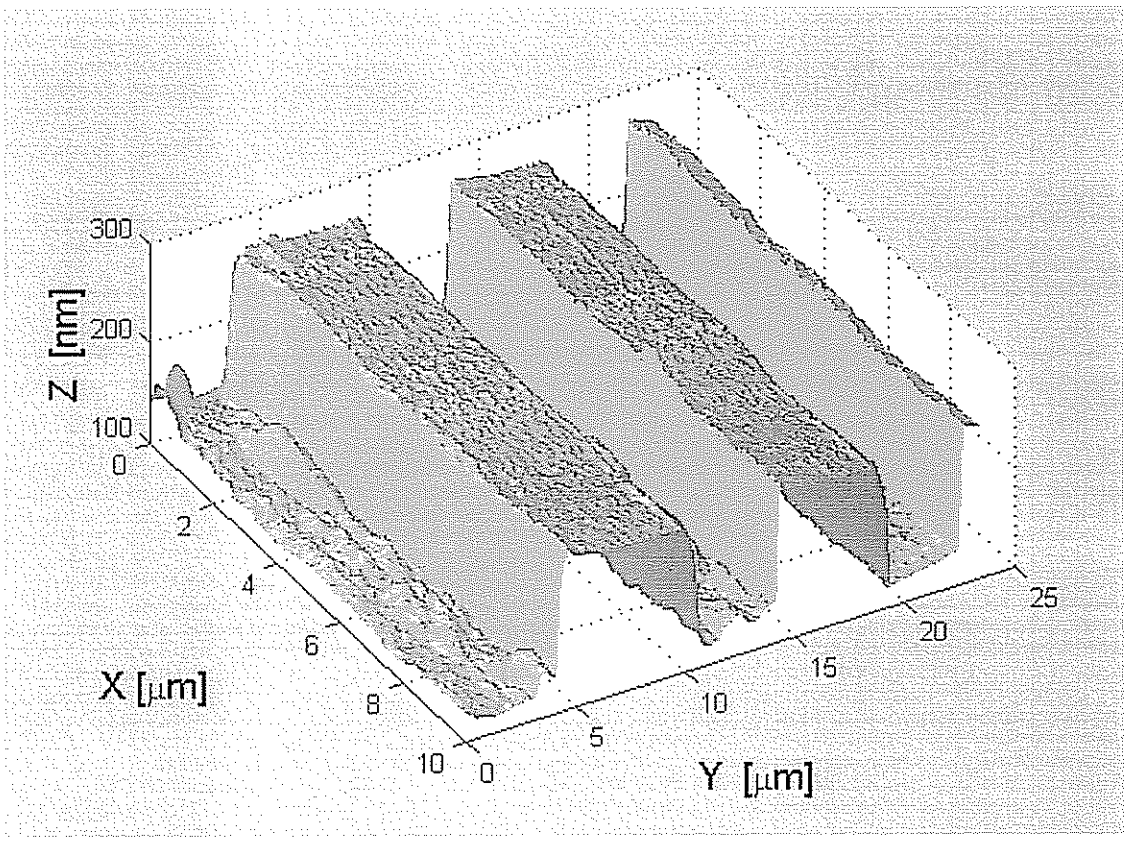


Fig. 12: A scan of etched mesa lines. The width of the lines is 6 μm and the period is 10μm. The surface was scanned perpendicular to the lines direction (in the Y direction). The step size is 100 nm.

5 Description of the microscope structure

Three version of the microscope were developed during the work. The first one was used for feasibility tests at room temperature. Shear force interaction, scan range, tip-sample approach mechanism, as well as convenience of operation were studied. The second microscope type, for cryogenic temperatures, was based on the first model. It was designed to operate in a relatively small (17mm diameter) sample tube of an optical cryostat. The tip-sample distance regulation was based on an optical technique. [71] Although some PL measurement where conducted with this version of the microscope, it has been found difficult to operate. Recent developments in tip-sample distance regulation using a piezo tube or a quartz tuning fork (see above), made it possible to use a non-optical sensing, and hence insert the microscope into a cryostat, which does not have optical windows (excitation and collection are done with optical fibers). These new techniques and the need to increase the scanned area and to operate in a more reliable fashion had led us to design and build a new version of the microscope. The key features, which characterize our microscope, are:

- It operates in a storage Dewar with a 44mm bore 6 T solenoidal magnet.

- It has a very stable mechanical structure so scanning is conducted while the Dewar is standing on the laboratory floor.

- The cool-down procedure is simple: there is a large (0.7 mm) Z range, which enables positioning the tip at a safe distance from the sample at the beginning of the cooling process.

- It has a 2 mm coarse X-Y movement of the sample relative to the tip, with ~1 μm resolution.

5.1 Microscope overview

The NSOM is operated inside a single-walled, exchange gas-cooled cryostat with 36 mm ID which fits inside the storage Dewar. The body of the microscope was made extremely rigid by machining it out of solid stainless steel. This and the low horizontal resonance frequency (~6 Hz) of the body (mass ~1 kg) on its long stainless steel support tube provides isolation from horizontal room vibrations. The tip-sample distance regulation is shear force detection using a commercial 33 kHz tuning fork. [75]

Inertial sliders (also called stick-slip sliders) [76,77] and inchworm-type positioners are commonly used for coarse X, Y, and Z motion in scanning microscopes. Their advantage over mechanical positioners is that they are completely electrically activated so that the problem of converting mechanical motion of a control rod along the Z-axis of a cryostat to precise X-Y motion is avoided. However, the careful balance between impulse and friction required by inertial scanners, and the exact tolerances essential to inchworm-type devices, are parameters that change with temperature. We have chosen to use classical mechanical means, whose implementation is more straightforward, reliable and simple to operate. There is no need for nanometer X-Y translation when the scan size is ~10 μm.

Except for dithering, the only motion by the tip is coarse Z. Tip-sample distance control and scanning are done by the sample piezo. Coarse X-Y motion is done by translating the sample piezo.

5.2 Microscope construction

The cryogenic end of the microscope is shown in Fig. 13. It is of modular design consisting of a top flange and four sub-assemblies that are easily separable for changing the tip or sample. The body is machined from type 303 stainless steel. Assembly I contain the X-Y table (Ia), which holds the 2" long by 1/4" diameter sample piezo (Ib). Onto the top of this piezo is cemented a 4-contact G-10 socket, into which the sample holder is

plugged. The table is essentially square, with two adjacent sides having a 2:1 slope [78].

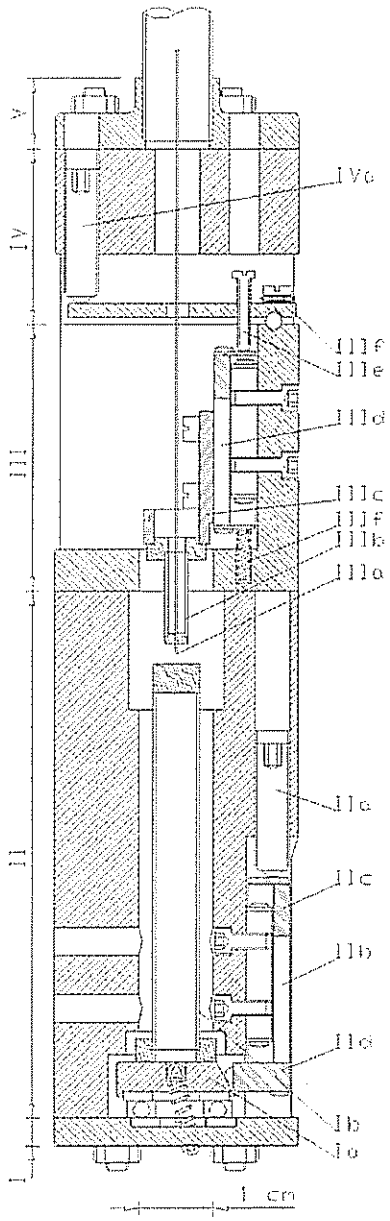


Fig. 13: A schematic drewing of the NSOM.

The table rests on a stainless steel thrust bearing. The eight 3/32" stainless steel balls in The table is pulled against the thrust bearing by a spring that is attached to the center of the table, passes through the center of the bearing, and is attached to the base. The point of attachment to the base is more than 1 mm off center so that the horizontal component of the tension keeps the sloping sides of the table firmly pressed against the balls in the X and Y drives (see Assembly II) over its entire ±1 mm range about the center.

Assembly II is the drive unit for the X-Y table. Its function is to convert the rotary motion of the X and Y 3/16"-100 hex adjustment screws (IIa), spaced 90° apart and driven by rods from outside the cryostat, to linear motion in the X and Y directions. They do this by pushing on linear ball slides (IIb). Two modifications were made to the slides. The end caps on the top ends were replaced with extensions (IIc) so that the adjustment screws would not hit the base. The end caps on the bottom end were replaced with aluminum blocks (IId). Into the block on the X-axis were epoxied two WC balls separated by several millimeters parallel to the Y-axis. The block on the Y-axis has a single ball. The three-point contact between the balls and the sloping sides of the table determine the orientation of the table. The sloping sides of the table also keep the slides (IIb) firmly pressed against the drive screws (IIa). MoS_2 dry lubricant is rubbed onto all moving parts to decrease friction and prevent galling.

Assembly III contains the tip (IIIa) and the coarse Z-motion stage. The tip is glued to the tuning fork (not shown) which itself is glued to a 4-segment 2mm O.D. x10mm piezo (IIIb). The bracket (IIIc) holding the piezo is screwed to the carriage of a ball slide (IIId) which is pushed by an offset adjustment screw (IIIe) against a return spring (IIIf). The offset adjustment compensates for variations in the distance in which tips extend past the tuning fork. A lever (IIIf) pivots on two WC balls and provides a reduction of 6.5:1 for the coarse-Z drive screw (see Assembly IV).

Assembly IV is the Z drive. The coarse-Z drive screw (IVa) is a 3/16"-100 hex adjustment screw. The top flange (V) is soldered to the thin-wall stainless steel support tube. Not shown in Fig. 13 is a small plate in a recess on the side of Assembly III that can

be tilted in the manner of a standard mirror mount. The fiber for side illumination (see below) is cemented inside a capillary tube which is soldered to this plate and which is bent so that the fiber is aimed at the tip. The X and Y drive screws are turned by hand. Their 254 μm pitch and the 2:1 slope on the X-Y table moves the table 127 μm/rev. Outside the cryostat, a 6:1 reduction gear and 100-turn counter give a resolution of better than 0.2 μm over the ±1 mm range. The resolution is adequate to allow overlap between scanned areas since the horizontal range of the sample piezo at 4.2 K is 11 μm (at 130 V). The 254 μm pitch coarse-Z drive screw is rotated from outside the cryostat by a standard optical laboratory stepper motor with 1000 steps/rev. Thus, the theoretical Z resolution, taking into account the 6.5:1 reduction provided by lever IIIf, is 39 nm/step. The Z range of the sample piezo at 4.2 K is 1.5μm (for 130 V).

All homemade parts of the microscope were made from non-magnetic material. However, non-magnetic commercial components did not always exist. In particular, the thrust washer in the X-Y table, the balls and rails of the ball-slides, and the ball of the hex adjustment screws are made from hardened 400-series stainless steel, which is magnetic. Even the WC balls are magnetic. These ferromagnetic materials limit the use of the microscope to < 4 T.

5.3 Optical system description

The excitation source for the PL measurements was either a Ti-sapphire or a He-Ne laser. The collected PL was measured by a 0.5m spectrometer and detected by a thermoelectrically cooled, back-illuminated CCD detector with matrix of 1100x330 pixels, 24μx24μ pixel size. The gain of the CCD is 4, namely, four incoming photons give one count. The system spectral resolution determined by the spectrometer and the pixel width of the CCD is 0.08 nm (each exposure covers a spectral range of 41.8 nm in the relevant spectral regime of 810 nm). The data were taken using a tip with a diameter of ~250 nm and transmission at 817 nm of 0.005. The diameter was measured with a scanning electron microscope. The transmission factor was measured both by

coupling light to the fiber and measuring the output power through the tip, and by illuminating the tip and measuring the light intensity from the other end of the fiber. Both methods gave roughly the same result.

An essential point in the operation of the microscope is related to the way the photo-excitation and the PL collection were performed. Operation in the so-called illumination mode, where the sample is photo-excited through the tip and the PL is collected by a broad lens, is not appropriate [43]. The resolution in this mode is limited by diffusion of the photo-excited carriers (typically ~1μm in GaAs samples). The alternative method of operation is the "collection-mode" where the tip is used to collect the emitted PL. In applying this method there are two possible means of excitation. One popular technique is to use the tip itself for excitation.[50] This is a simple and straight forward technique, but it creates background optical noise which will be discussed below. The second technique is to use a separate single-mode fiber oriented such that the light is nearly parallel to the surface. This form of excitation may create some shadowing of the illumination by the tip, but gives uniform excitation and a background-free PL signal. In this work we have used the collection mode throughout our measurements. In the following we describe some test measurements that characterize the system performance in this operation mode.

5.3.1 NF signal as a function of tip height

An indication of a NF measurement is the increase in the collected signal as the tip approaches the near field zone (another indication is the loss of polarization at this zone [79]). In Fig. 14 the PL intensity versus the tip-sample separation is shown. The PL is taken at collection mode, while the excitation is done by a separate single-mode fiber. It is seen that the PL intensity increases continuously as the tip approaches the sample. However at the last micron, the signal rises by a factor of ~2. This factor is tip dependent: the smallest the tip, the larger is the increase, a factor of ~6 is reported in [60] for tip diameter of 100nm. The rise is a clear indication of entering the near-field zone.

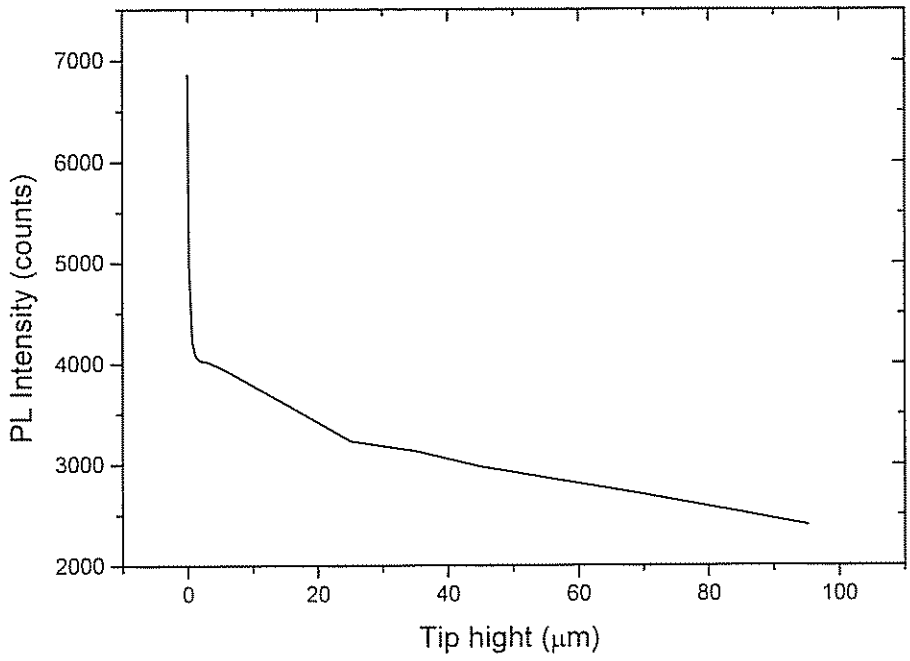


Fig. 14: PL intensity vs. tip height above the sample

5.3.2 Resolution measurements

In order to check the spatial resolution of the optical system and to compare the optical and shear-force signals, a special sample has been prepared, consisting of seven GaAs quantum wells, 10 nm each, separated by 20 nm AlGaAs barriers. 30 nm thick Pd/Au strips with 1,5 μm width and 3.5 μm pitch were evaporated on top of the sample surface. The measurement is done using the tip illumination mode, with 0.6 mW at a wavelength of 730 nm coupled to the fiber, the integration time is 1 sec. The tip is scanned over the sample, collecting the PL emission from the quantum wells.

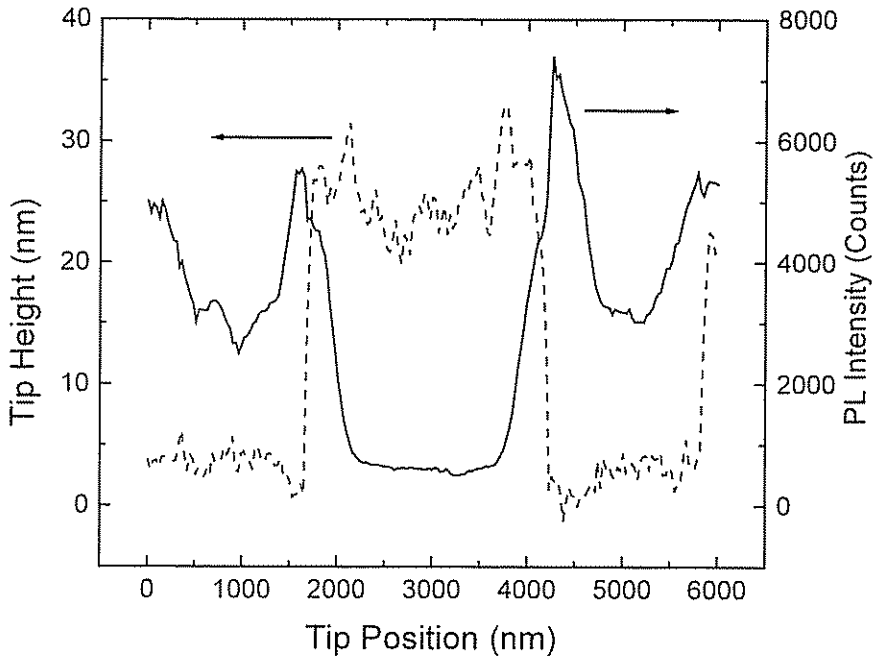


Fig. 15: Integrated PL (solid) and shear-force (dashed) of a test sample

Fig. 15 shows the total counts under the PL peak (solid) and the topography of the sample (dashed) versus the tip position. The resolution, which is defined as the distance between the points at which the intensity rises from 10% to 90% of its maximum, is 250 nm. This is in a good agreement with the SEM measurements of the tip aperture diameter.

It is important to realize a significant factor determining the spatial resolution when measuring a signal from a QW buried below the surface. The PL is emitted at a wavelength around λ from a layer which is ~100nm below the surface. PL emission carrying spatial frequencies smaller than $n\backslash\lambda$, where n is the refraction index, can propagate in the medium to the surface. There they can be detected by the tip, which has to be in the near-field region of the sample surface. PL emission carrying higher spatial frequencies decays almost completely over the distance to the sample surface [40] In GaAs ($n \cong 3.5$) and for λ~815nm this spatial frequency cut-off is 230^{-1} nm^{-1}. Thus there is

not much to gain by using tip with diameter smaller than 250nm.

5.3.3 Optical noise using tip illumination mode

We have studied the noise that is created excitation through the tip. We have found that this noise has two sources: PL emission of the fiber core (doped with GeO_2 in our fibers) and spontaneous Raman scattering. When illuminating the fiber with the He-Ne laser (632.8 nm), the noise is dominated by the PL emission.

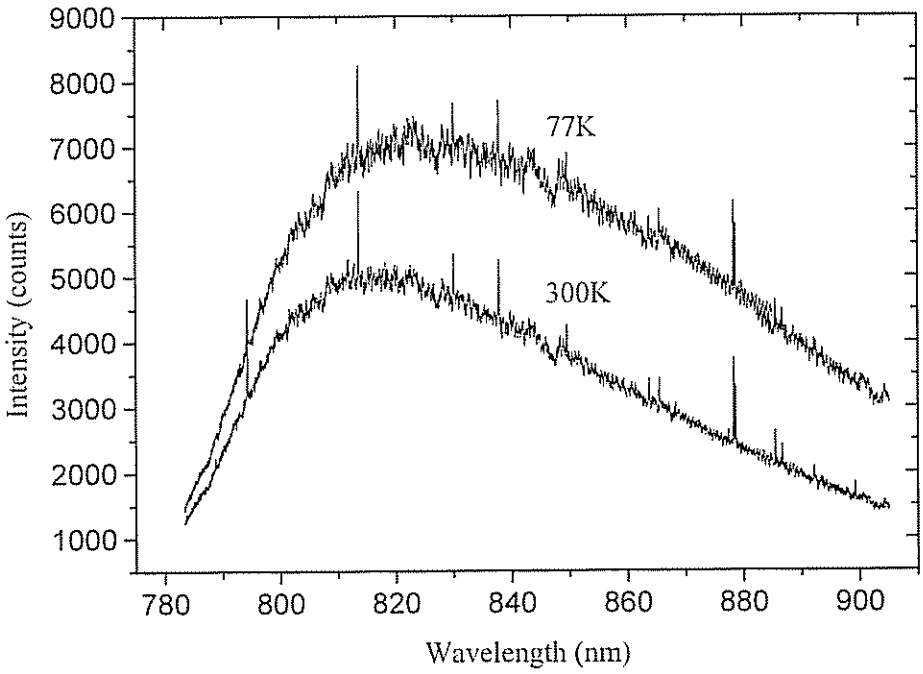


Fig. 16: PL noise at two 300K and 77K.

Fig. 16 shows spectra of the background noise at fiber temperatures of 300 and 77K. These spectra were taken while illuminating the fiber by HeNe laser. The power coupled into the fiber was 0.6 mW, the integration time 10 sec, and the fiber length 4.5 m. When the fiber is cooled down to 77K, the noise increases by ~1.5. It has been found that the

noise intensity scales linearly with excitation intensity and with the fiber length. At a typical wavelength of 815 nm, the PL noise is ~300 counts/(mW m sec).

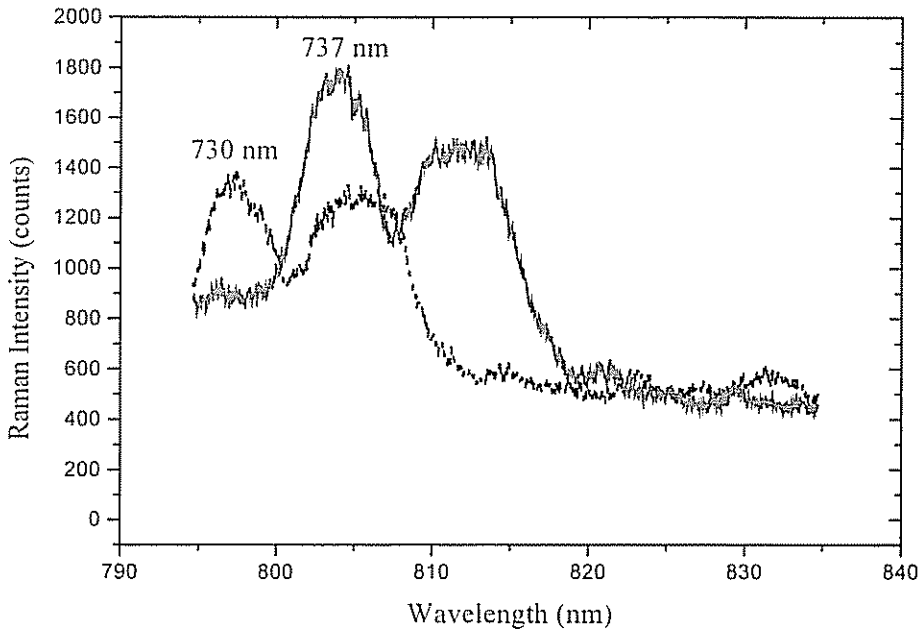


Fig. 17: Intensity of Raman scattering noise at two illumination wavelength.

As the fiber is illuminated at longer wavelengths, spontaneous Raman scattering dominates the noise. Fig. 17 shows the Raman background noise spectrum at illumination wavelengths of 730nm and 737nm. The illumination power is 0.3 mW, the integration time 10 sec and the fiber length 3m. It seen that the noise spectrum is shifted from the exciting laser by a fixed spacing: as the laser wavelength is changed from 730 to 737nm the noise spectrum shifts by exactly 7nm. This is the expected behavior of Raman scattering. The Raman noise also scales linearly with the fiber length and with the illumination intensity, but does not depend on the fiber temperature. At 815 nm the Raman noise is ~70 counts/(mW m sec). We have found that the Raman noise is unpolarized, in contrast with the noise reported in Ref [50]. The noise source in

[50] is different, and probably related to the metal coating of their tip.

The spontaneous Raman intensity, I_s, is related to the laser intensity, I_L, by [80]

$$I_s = \frac{nD\omega_s}{c\omega_L} zI_L , \tag{38}$$

where n is the index of refraction of the fiber, D is the Raman coefficient, c is the speed of light, I_s and I_L are the Raman and the laser angular frequencies, and z is the fiber length. D was calculated from Eq. ((38) and found to be independent of I_L and z. We found $D \approx 10^{-6}$ sec^{-1}, which is in good agreement with a previous measurement [81].

When exciting via a far field source, with a power density of 100mW/cm^2, the measured PL signal level for a 20 nm width GaAs single quantum well was ~40 counts/sec at the exciton peak maximum. The integrated signal under the exciton peak is 600 counts/sec. In this excitation mode the noise is dominated by photon shot noise.

6 Near Field measurements of a Gated 2DEG

It was shown in Chapter 3, that the X^- complex is formed from a photo-excited electron hole pair and a native electron. Therefore, X^- recombination is an indication for the presence of an electron at the 2DEG layer. In this research work we utilized this property of the X^- together with the high spatial resolution of the low temperature NSOM to image the electrons distribution in the 2DEG plain.

6.1 Sample structure

The sample measured in this work consists of a 20 nm GaAs quantum well followed by 37.5 nm $Al_{0.37}Ga_{0.63}As$ spacer layer and a 10 nm layer $Al_{0.37}Ga_{0.63}As$ doped with silicon at concentration of $3.5x10^{18}cm^{-3}$. The structure is capped by a 20 nm undoped $Al_{0.37}Ga_{0.63}As$ and 10 nm GaAs. A 2x2 mm^2 mesa was etched, and ohmic contacts were alloyed into the 2DEG layer. A 4 nm Pd/Au semi-transparent gate was evaporated on top. The continuity and uniformity of the gate was verified by SEM imaging. The 2DEG concentration was $4x10^{11}cm^{-2}$, and the mobility $1.3x10^6cm^2/V$ sec, both measured at 4.2 K. Upon excitation with the laser the electron density is reduced by an amount that depends on the laser intensity. Consequently, the gate voltage needed to deplete the sample is intensity dependent.

6.2 Near field imaging of electron distribution using the X^- line

In Fig. 18 a series of PL spectra of a gated 2DEG is shown. These spectra ware taken with the tip at the NF zone at different gate voltages. By comparing these spectra to the far-field spectra shown in Fig. 3, it is evident that there is no dramatic difference between the local PL and the large area far-field emission of the 2DEG. In both figures we see clearly the evolution from a broad spectrum, of free electrons, to an excitonic spectrum, characterized by two peaks. The high and low energy peaks are due to a

recombination of a neutral (X) and negatively charged (X^-), respectively.

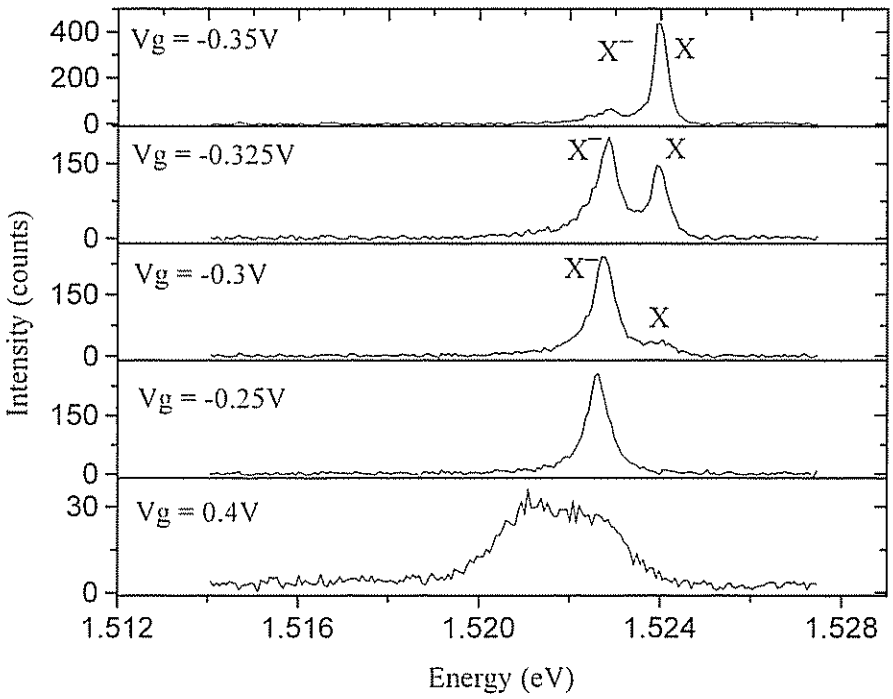


Fig. 18: Near –Field PL evolution of a gated 2DEG

This work is mainly concentrated in the gate voltage regime where the spectrum becomes excitonic. The underlying idea is to scan the sample with the NSOM tip and collect the local PL emission at a constant gate voltage. Since the X^- is a local object representing a localized electron, a change in the local X^- PL intensity should be observed as the tip move from regions with high electrons concentration to lower one.

6.2.1 Non-uniformity of the local spectra

Fig. 19 shows four near-field spectra measured at different locations at the same gate voltage, V=-0.135 Volts.

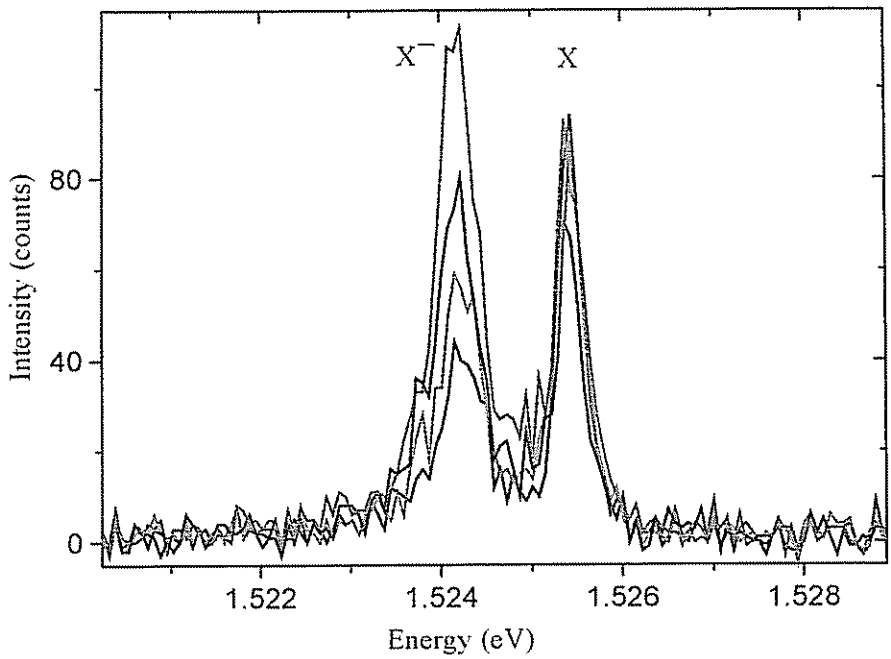


Fig. 19: Four PL spectra taken at different locations at the same gate Voltage.

At this gate voltage the spectrum is excitonic: the high and low energy peaks are the X and X^-, respectively. It can be seen from the figure that the near-field spectrum is different from one location to another: while the height of the X peak is almost similar in the four spectra, the height of the X^- peak varies substantially. Since the X^- is an indicator for the presence of electrons, these variations in its intensity show that the electrons are *non-uniformly distributed.*

This behavior of the X and the X^- peaks is verified by measuring the sample PL with the

NSOM tip along 11 μm long line, with 100 nm steps between subsequent measurements. The results are shown in Fig. 20.

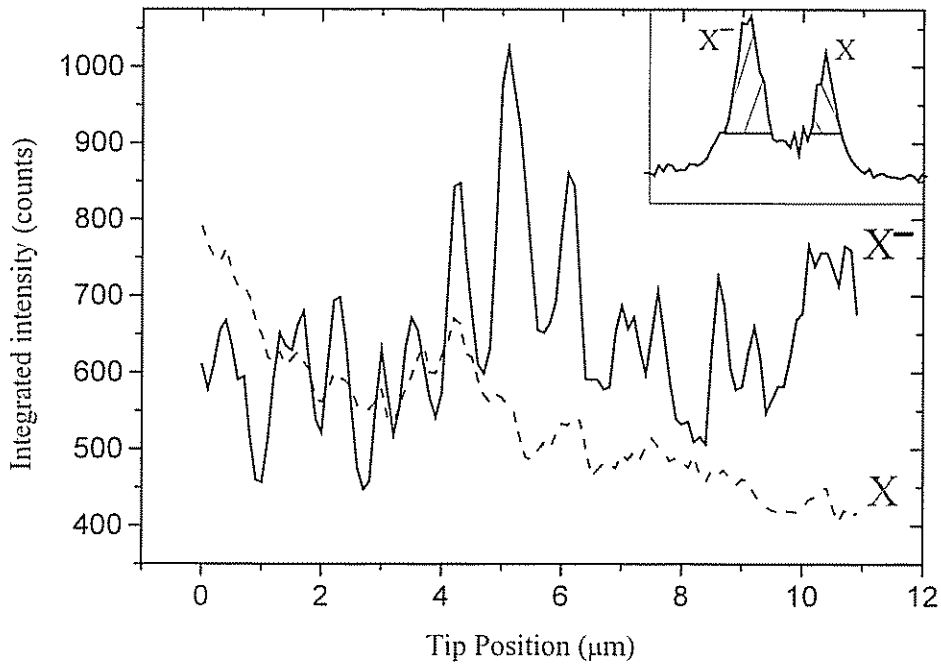


Fig. 20: Integrated intensity of the X and X^- peaks along an 11 μm line scan.

To construct the figure from the measurement, the intensity under the X and X^- peaks of each local PL spectrum were integrated. The integrated intensity is shown as a function of the tip position. The inset of Fig. 20 shows the integrated areas of the X and X^- peaks on a typical spectrum that was taken in this scan.

It is evident that there are large fluctuations in the X^- intensity. These fluctuations occur throughout the scanned region and on any length scale, (see below) down to the resolution limit. The X peak exhibits much smaller fluctuations. It should be

emphasized that these fluctuations are stable over time: repeating the measurement over and over again reproduces the line scan.

As the X^- is a negatively charged object it can be localized by the random electrostatic potential caused by the non-uniform distribution of the donors. The amplitude of these electrical potential fluctuations can be as high as ~10-20 meV, as calculated in Chapter 2. The neutral exciton on the other hand, is electrically neutral and therefore it is only weakly influenced by the electric field, and has a relatively high diffusion length. This explains the small variation in the X intensity relative to the X^-.

To further prove the local nature of the X^- intensity fluctuations we performed repetitive measurements of the PL spectrum, first with the tip at the NF region (~10nm above the sample surface), and than with the tip at a height of 1μm above the sample. The intensities of the X and the X^- peaks were integrated as in Fig. 20. The result is shown at Fig. 21. It is seen that at a height of 1μm, the fast spatial frequencies disappear. Only spatial frequencies smaller than a (few micron)$^{-1}$ survive. This is also seen in the X line.

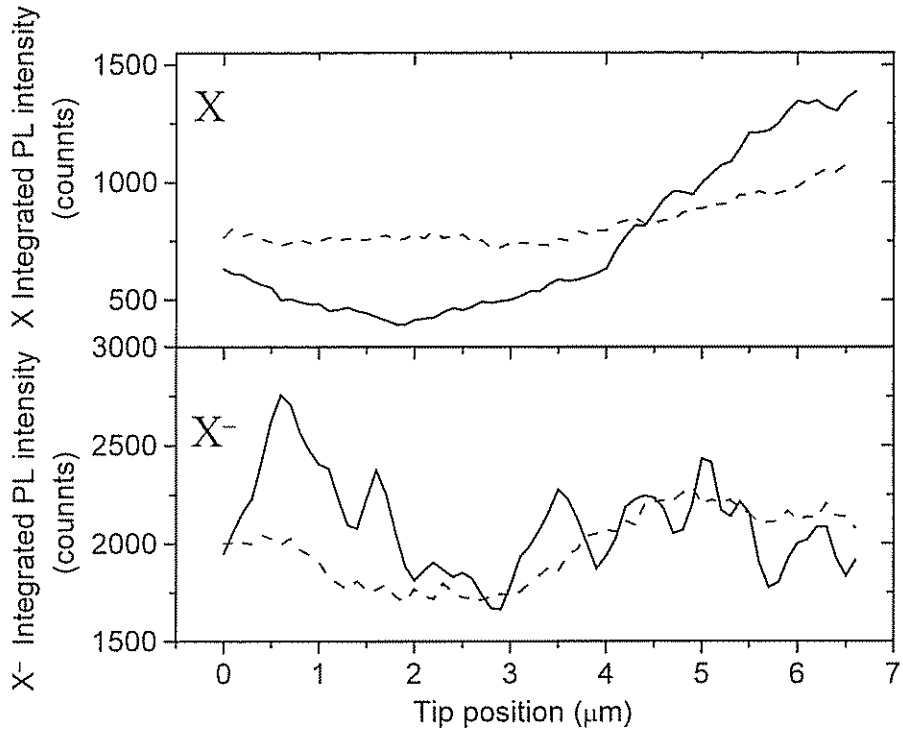


Fig. 21: Intensity of the X^- and X peaks along 7μm scan. The solid and dashed curves are measured at a tip height of 10 nm and 1μm.

Thus, the far-field spectrum in the gate voltage regime where the spectrum is excitonic is an integration of many local spectra that are different from each other. The far-field intensity of the X and X^- in the spectrum reflects the average electron density at the area, from which the PL is collected.

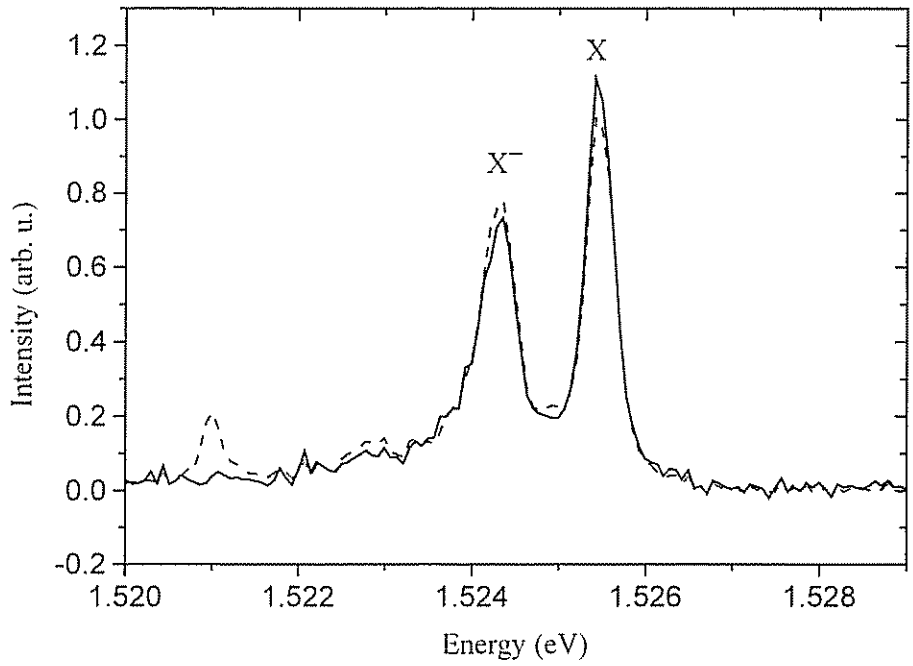


Fig. 22. Solid curve: A sum of 121 spectra collected by scanning an area of 1x1 μm^2. Dashed curve: Far-field spectrum taken ~1 μm above that area.

Fig. 22 demonstrates this statement. It is constructed by the following procedure: First a square of 1x1 μm^2 is scanned and the PL spectra from all the points are summed up (Fig. 22 solid line). Then the tip is withdrawn to a distance of 1 μm from the sample and the PL is measured above the center of this square (Fig. 22 dashed line). It can be seen that the two curves nearly coincide.

6.2.2 PL intensity fluctuations as a function of gate voltage

Fig. 23 shows the evolution of the integrated X^- PL intensity along 11μm scan as a function of gate voltage.

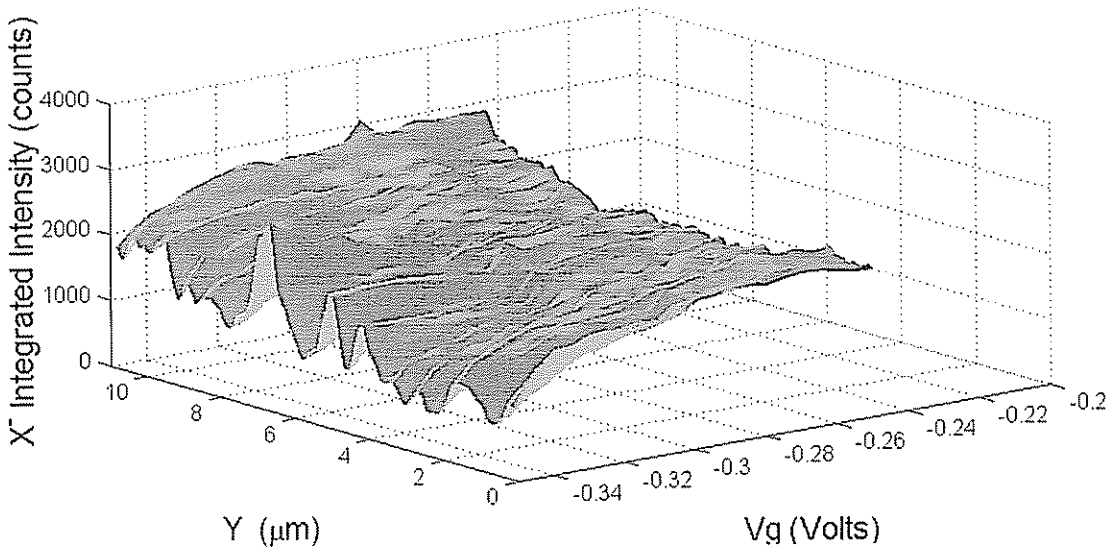


Fig. 23: Integrated intensity of the X^- peak along 11μm line at different gate Voltages.

At small negative gate voltages, the intensity along the line is relatively smooth. As the gate voltage becomes more negative the X^- intensity exhibits large spatial fluctuations. It can be seen that these fluctuation evolve adiabatically with the gate voltage: as the gate voltage becomes more negative local maxima become higher and local minima- lower

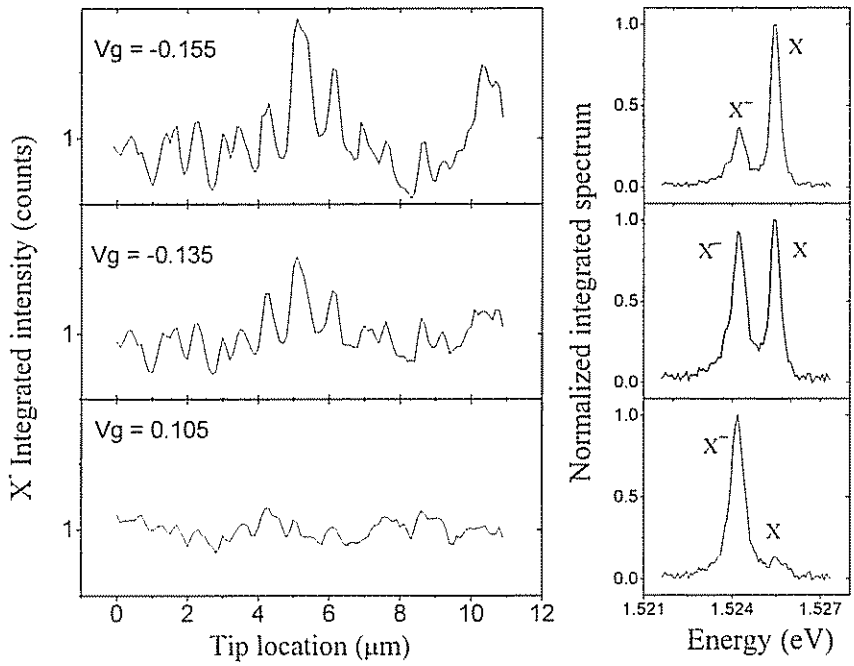


Fig. 24, left: X^- integrated intensity along 11μm scan at three gate voltages. Right: average spectrum of these lines.

Figure 25 shows three line-scans, each taken at a different gate voltage: Vg=0.105 V, -0.135 V and -0.155 V. (Fig. 24 is measured at a higher intensity than Fig. 23, hence the gate voltage are different). These voltages correspond to different relative intensities of the X and X^- peaks in the 2DEG spectrum. The average spectrum along each line-scan is presented at the right side of Fig. 24. It can be seen that at Vg=0.105V the spectrum is dominated by the X^-, at Vg=-0.135V the X and X^- peaks are nearly equal, and at Vg=-0.155 the spectrum is dominated by the X peak. The line scans are normalize by dividing the value at each point along the line by the average value of that curve. Comparing these line scans, one can clearly see that the relative fluctuation amplitude increases with the gate voltage. In the lower curve, where the gate voltage is positive, the ratio between the maximum and the minimum is 1.4, while in the upper curve

this ratio is more than three. The positions of the maxima and minima, however, are almost fixed in space, and are almost unaffected by the gate voltage.

6.2.3 Fourier analysis of the fluctuations

Frequency components of a line scan can be calculated by fast Fourier transform (FFT). This procedure is limited by the number of data point along the scan and the scan length (111 points along 11μm). An additional factor that poses a limit on the FFT is the finite diameter of the tip, which sets a high frequency cut-off. In the following we present the wavelengths of the FFT in units of μm.

An FFT for the three lines that are presented in Fig. 24, is shown in Fig. 25. It is evident that as the 2DEG is depleted (more negative gate voltage) the intensity of the entire Fourier components (long and short) increases (note the different scale). This change in amplitude will be further discussed in the next section. However, it is also seen that the relative intensity of the short wavelength components decreases compare to the longer ones. The most noticeable peak at the dense 2DEG (Vg=0.105) is at 3.3 μm, and is shifted to 5.5 μm at the dilute 2DEG (Vg=-0.135 and Vg=-0.155). This behavior is a result of the fact that long wavelength components are the most easiest to screen, hence, as the 2DEG is depleted the long wavelength components, which were screened by the dense 2DEG, becomes more pronounced.

The shortest component at all spectra that is steel pronounced is at 0.9μm, which is ~3 times the tip diameter. Below this cut-off wavelength the amplitude decreases abruptly. This abrupt jump in the amplitude is due to the tip diameter that averages out spatial components smaller than 3 tip diameters. It should be noticed, however, that Fourier components exist at all possible wavelengths down to the FFT resolution.

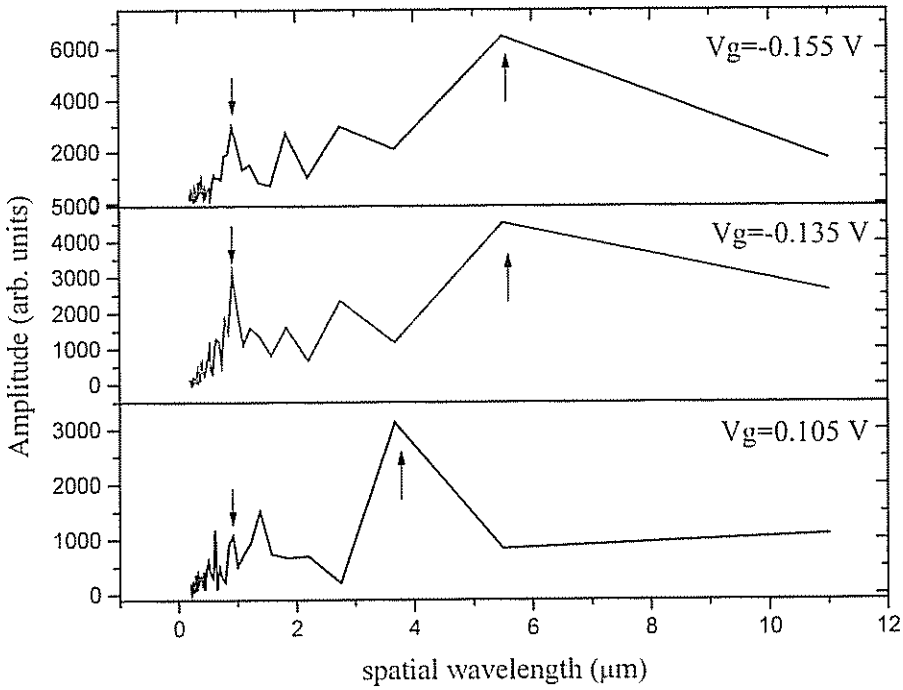


Fig. 25: FFT of the three scans presented in Fig. 24.

6.2.4 Fluctuation amplitude

Fig. 26 summarizes the dependence of the fluctuation amplitude on gate voltage (note that this measurement is taken at a weaker illumination intensity then the measurements of Fig. 23 and Fig. 24, hence the gate voltages are more negative). The graph describes the dependence of the normalized standard deviation of the X^- fluctuations amplitude on gate voltages. The two spectra are the typical spectra at the gate voltage points that are denoted by the arrows. It can be clearly seen that the fluctuations amplitude is constant over a large range of gate voltages and starts to increase at the voltage, at which the spectrum becomes excitonic. This correlation between the appearance of the excitonic spectrum and the rise of the fluctuations amplitude is very significant.

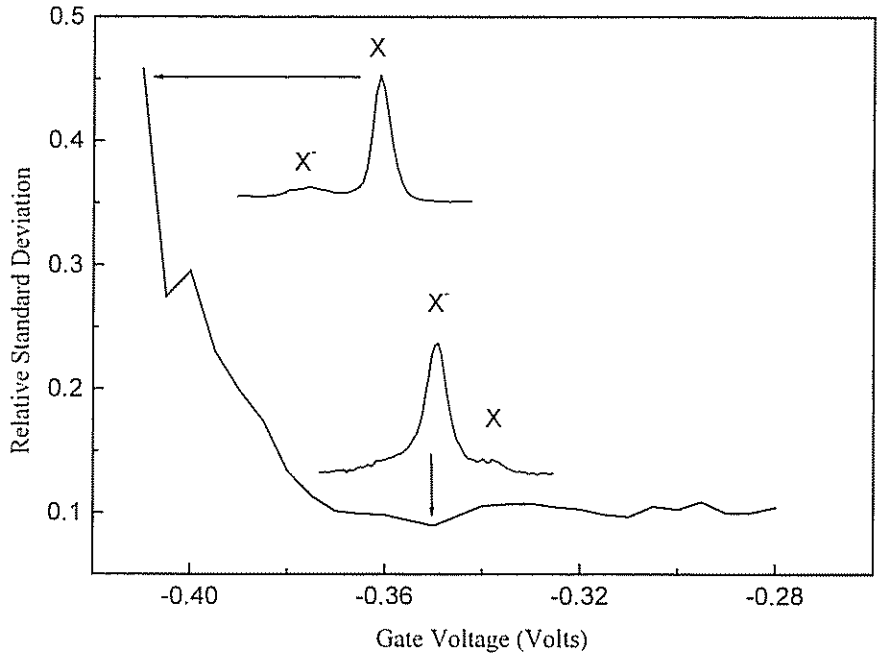


Fig. 26: Relative standard deviation vs. gate voltage.

The excitons appearance is an indication for a change in the screening properties of the 2DEG: it becomes ineffective in screening the electron-hole interaction. It follows that the 2DEG is also ineffective in screening the random potential induced by the ionized donors in the doped AlGaAs layer, and the fluctuations in this potential grow. These donors are randomly distributed, with an average distance between them of a few nm. They give rise to a random electrostatic potential in the 2DEG plane, with the smallest spatial period being the spacer width (37.5 nm in our sample).[21] The high spatial frequencies of the observed fluctuations and the fixed positions of the minima and maxima are consistent with this explanation.

It is evident from Fig. 26 that fluctuations in the PL intensity are observed throughout the gate voltage range. At the range where the PL is 2DEG-like, namely - non-excitonic and shifted to lower energies, the electrons are delocalized. The weak fluctuations in this range are independent of the gate voltage, and are due to localization of the photo-excited

holes in the donors potential. The observed increase in the fluctuations amplitude as the spectrum becomes excitonic *is it the onset of the electrons localization.*

Since the spatial resolution of the measurement is ~250 nm, it is impossible to collect light from a single localization site, which has a typical size of 40 nm, the spacer width. Thus, at each tip location the tip collects light from several sites. The changes in the X^- intensity from one tip location to another originate from a non-uniform occupation of these localization sites.

6.2.5 Modeling the localized system

To model the localized system let us assume a simplified model, where the potential fluctuation consists of a periodic potential in the plane, in which the electrons are randomly distributed. In such a model the tip samples a sub-region, in which there are N sites with a probability p to be occupied by electrons. Fig. 27 describes such a model with $p=0.3$. The circle in the middle describes the tip with diameter of 5 localization sites, thus $N\cong 25$

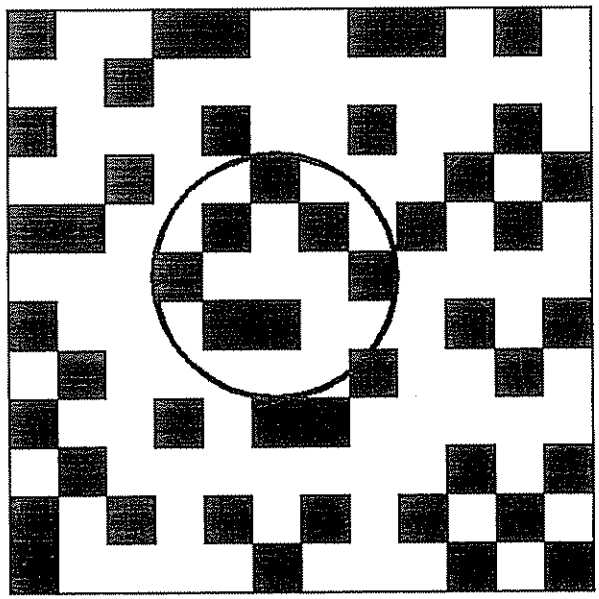


Fig. 27: Schematic diagram of the simplified model describes in the text.

The average number of electrons occupying the area under the tip is $n_{av} = pN$, and the corresponding standard deviation is $\sigma_n = \sqrt{pN(1-p)}$. The average PL intensity of the X^- line that is measured by the tip can be written as $I_{av} = \alpha n_{av}$, where α is a proportionality factor. Hence, the maximum intensity, I_{max}, can be written as: $I_{max} = \alpha N$. The standard deviation, σ in the X^- intensity is proportional with the same proportionality factor to σ_n. Thus α can be expresed in terms of measureble quantities I_{av}, I_{max}, and σ:

$$\alpha^2 = \frac{\sigma^2}{\sigma_n^2} = \frac{\sigma^2}{\frac{I_{av}}{\alpha}(1 - \frac{I_{av}}{I_{\max}})}$$

$$\alpha = \frac{\sigma^2}{I_{\max}\left(1 - \frac{I_{av}}{I_{\max}}\right)}$$

By substituting α, N can be calculated to be:

$$N = \frac{I_{\max}}{\alpha} = \frac{I_{\max} I_{av}(1 - I_{av}/I_{\max})}{\sigma^2} .$$

Since I_{av} and σ depend on the gate voltage, the consistency of this model can be checked by calculating N for each gate voltage. Fig. 28 shows N for the measurement in Fig. 26. It can be seen that indeed a relatively small scatter around an average value of $N \cong 30$ is obtained. Since the tip diameter is 250 nm this implies an average fluctuation size of 40 nm. This is in a very good agreement with the expected size, the spacer width (37.5 nm).

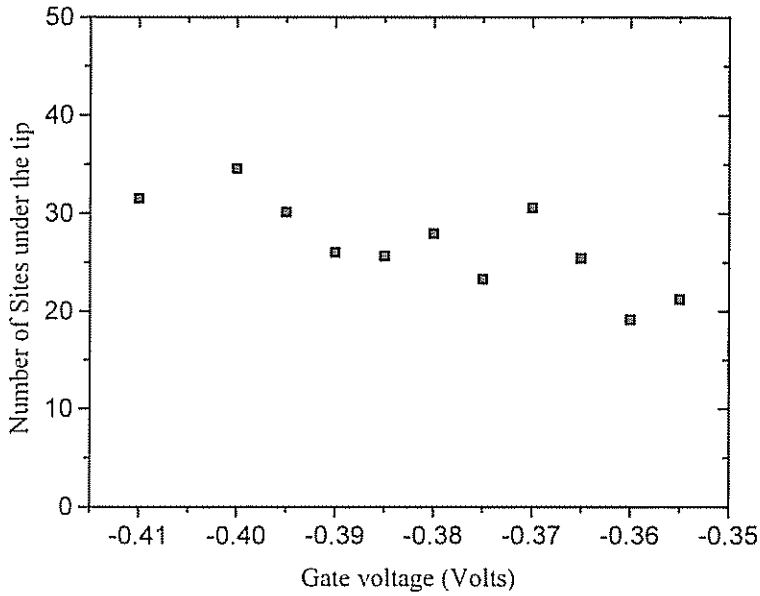


Fig. 28: Number of sites under the tip as a function of the gate Voltage.

Fig. 29 shows the distribution of the X^- intensities along a scanned area of $11x11\mu m^2$. The dashed curve describes a Gaussian fit to this distribution (the inset shows typical spectra of the measurement). The nice fit is an indication to the randomness of the measurement. This randomness, is consistent with our model and is due to the random nature of the donors distribution during the growth process.

It is important to note that a gate voltage range, where the spectrum is excitonic at one area and 2DEG-like in another area, has not been found. Such a spectrum might represent an isolated puddle of free electrons. This fact indicates that there are no large clusters of free electrons, and the localization is of single electrons.

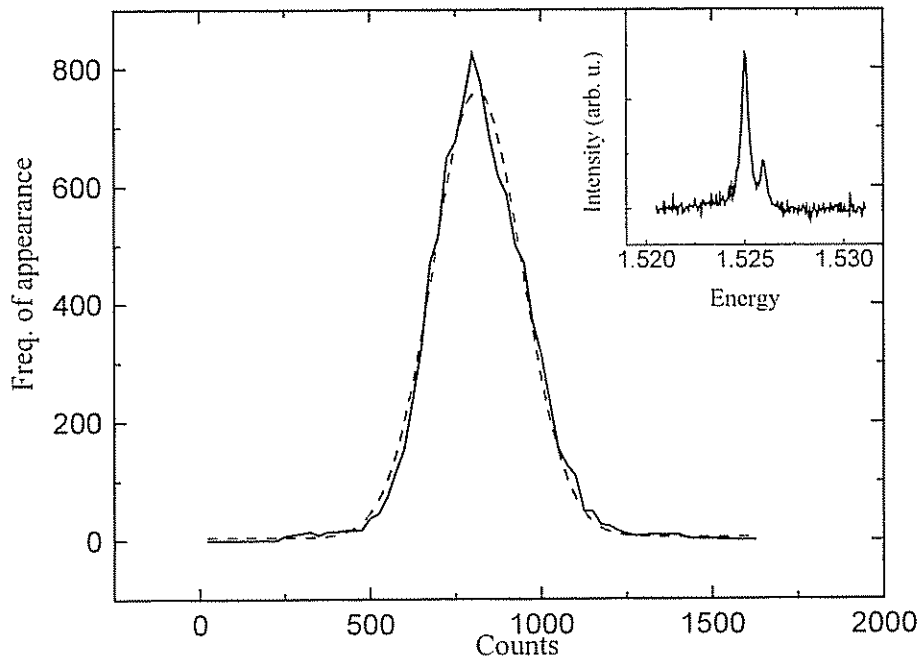


Fig. 29. Solid curve: Intensity distribution of the X^- peak. Dashed curve: Gaussian fit to this distribution. Inset: Integrated spectra of the scanned area.

6.2.6 Imaging the donors density

Since the position of the minimum and maximum in the X^- PL intensity, is nearly independent of the gate voltage, imaging of the donors distribution can be done. The underlying idea is that the localizing potential is due to fluctuation in the donors density, N_d. It was shown in [21] that the number of localized electrons in a given area is proportional to the square root of the local donor density $\sqrt{N_d}$. Thus, the fluctuations in the X^- intensity mirror the fluctuations in the donors density. By fixing the gate voltage at a convenient value and scanning the sample, we obtain a two dimensional map of the X^- intensity, which directly related to the donors distribution. Fig. 30 shows such a map for a region of 6x6 μm^2.

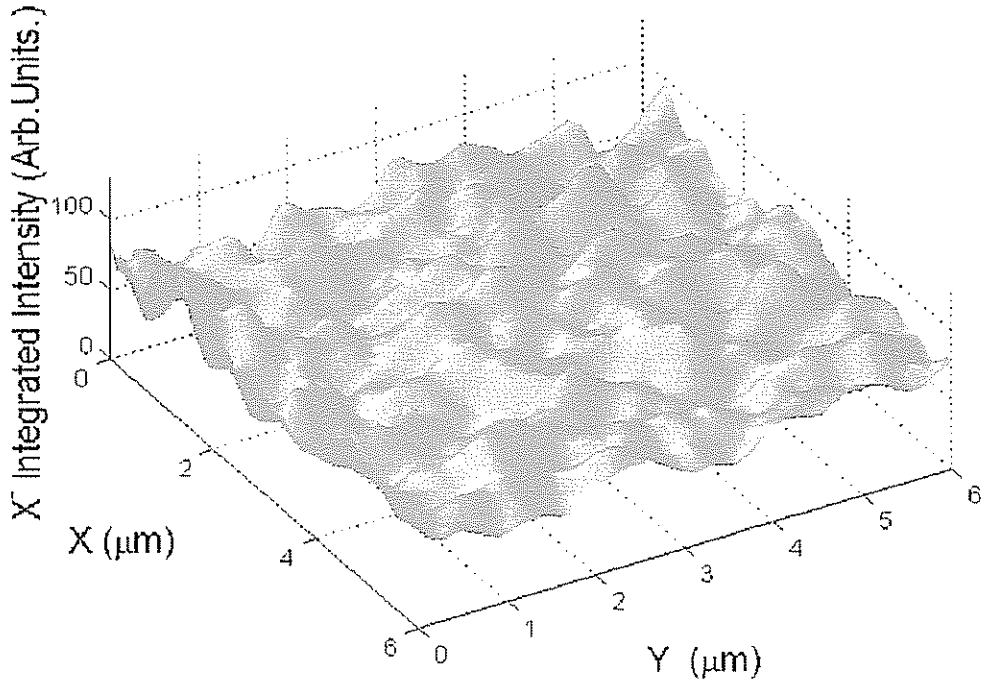


Fig. 30: 6x6 μm² image of X^- energy distribution.

6.2.7 Fluctuation in the dense 2DEG phase

The donors density fluctuations are expected to influence the 2DEG also in its denser phase. The induced potential fluctuation were calculated in Chapter 2 in the presence of screening by the 2DEG electrons. It was shown that the screening reduces substantially these fluctuations. However, a mean fluctuation value of ~1mev should be felt by the 2DEG. These potential fluctuations should give rise to density fluctuations, and hence one may naively expect to observe local variation in the PL line-width: the PL line-width should be broad at a high density region and narrow at a low density region. These variations in line-width are not observed. There are several possible mechanisms,

which can explain this absence of line-width variations. Further investigations with better spatial and spectral resolution are needed to better understand this issue.

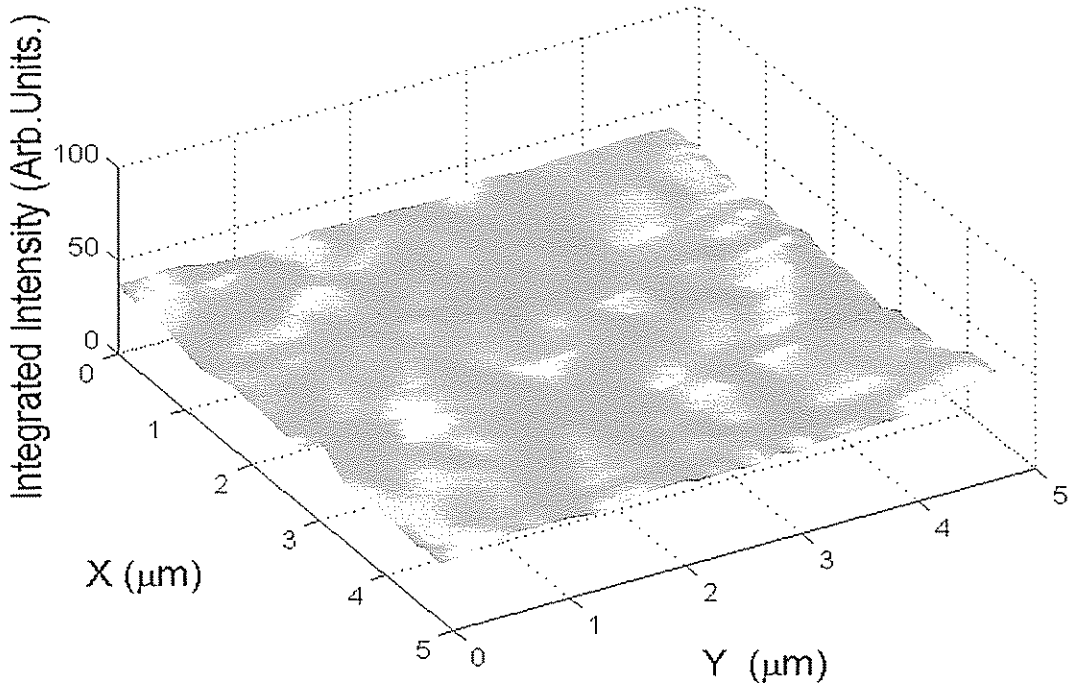


Fig. 31: Surface map of PL intensity from a dense 2DEG.

On the other hand, the intensity of the PL in this regime does fluctuate. These fluctuations are due to local fluctuation in the density of the photo-excited holes that are also influenced by potential landscape. Image of the intensity of the PL peak at the dense 2DEG phase is shown in Fig. 31. The Fermi energy of the 2DEG during the scan were estimated to be ~2-3meV.

6.2.8 Temporal intensity fluctuations

In addition to spatial fluctuations of the X^- and X intensities, it is found that there are

also temporal intensity fluctuations. This phenomenon occurs while exciting the sample with a high enough power (above 50mW/cm^2).

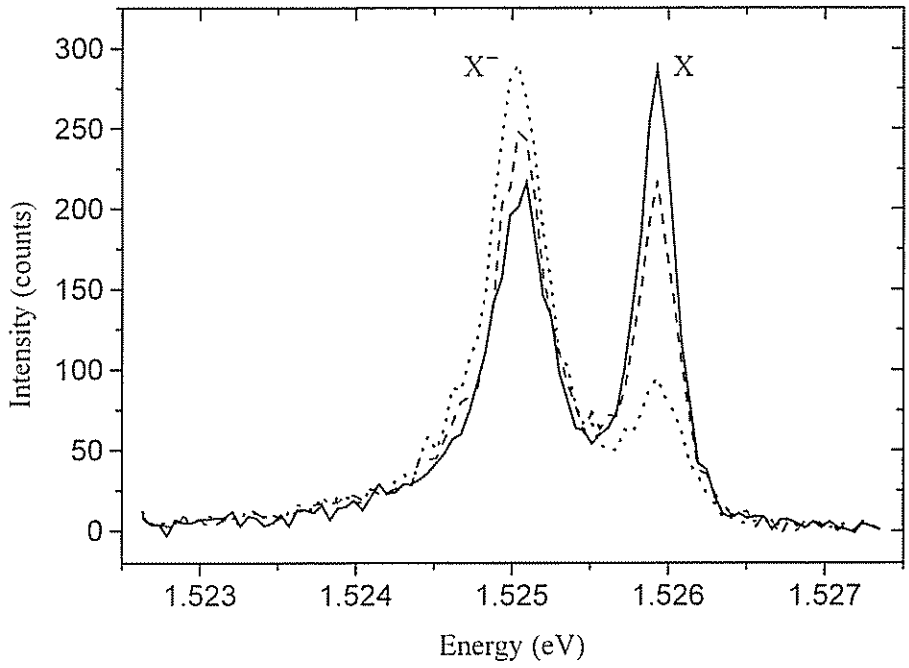


Fig. 32: Three spectra taken at the same tip position at excitation intensity of 60mW/cm^2.

Fig. 32 shows three spectra taken at the same tip position at high excitation level of 60 mW/cm^2. It is evident from the figure that in contrast to the spatial fluctuations, where the changes are much more pronounced in the X^- intensity, the temporal fluctuations are mainly observed in the X intensity.

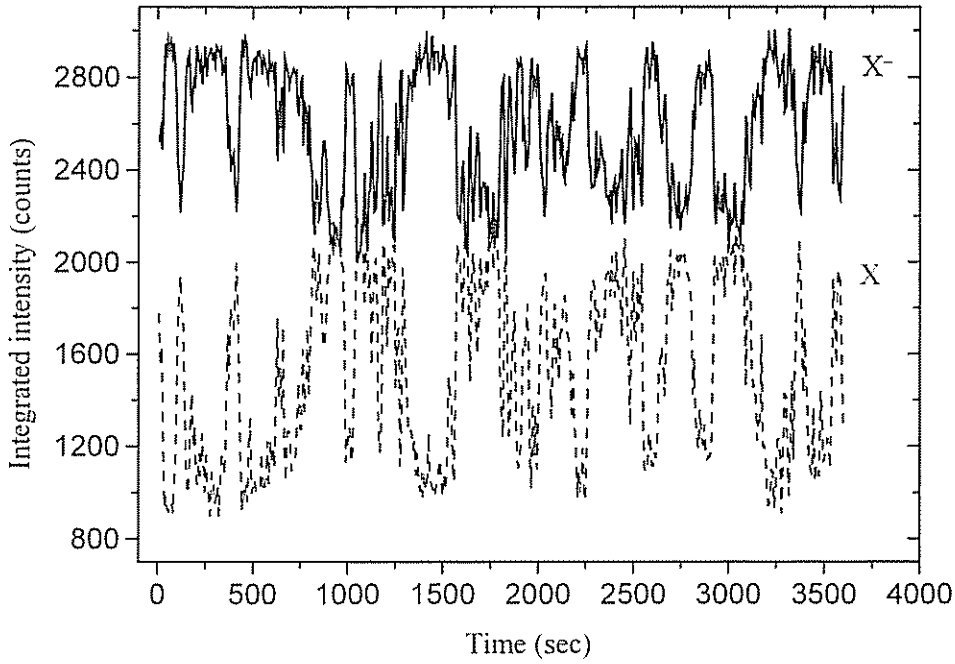


Fig. 33: Integrated intensity of the X^- and X during 1 hour.

Fig. 33 shows the integrated intensities of the X^- and X peaks during 1 hour measurement (480 points). It is seen from the figure that the fluctuations at the two peaks are at anti-phase. An FFT of these traces shows that slow frequencies are more pronounced than the higher ones (this is also seen from the figure).

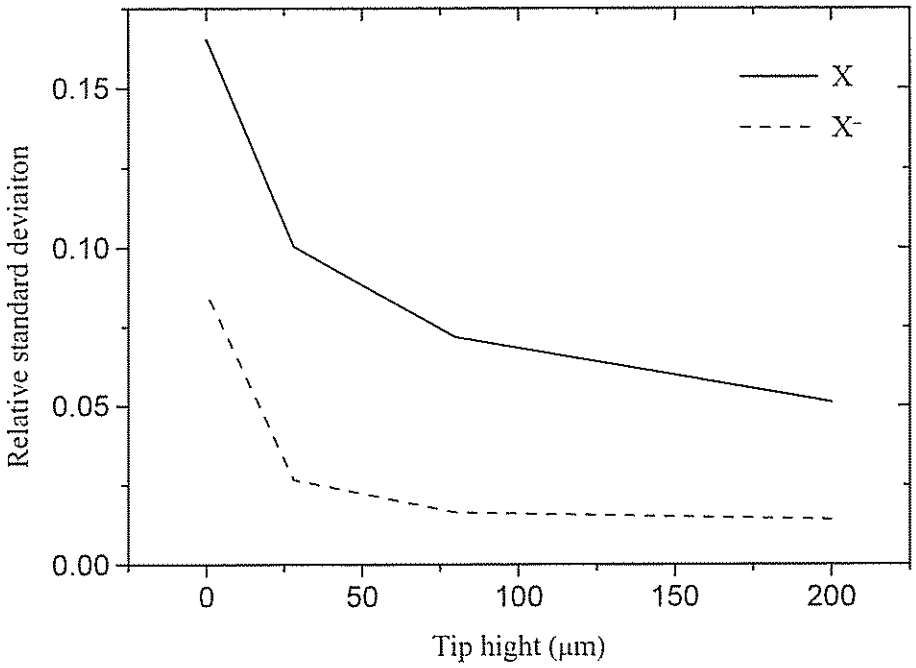


Fig. 34: Relative standard deviation of the X and X^- integrated intensity as a function of the tip height.

Fig. 34 shows the relative intensity fluctuations as a function of tip height above the sample. The relative standard deviation is calculated by dividing the absolute standard deviation by the average intensity. It is seen that the temporal fluctuations decay as the tip is withdrawn from the sample. However, even at a height of ~30μm the fluctuations in the X intensity is steel observed.

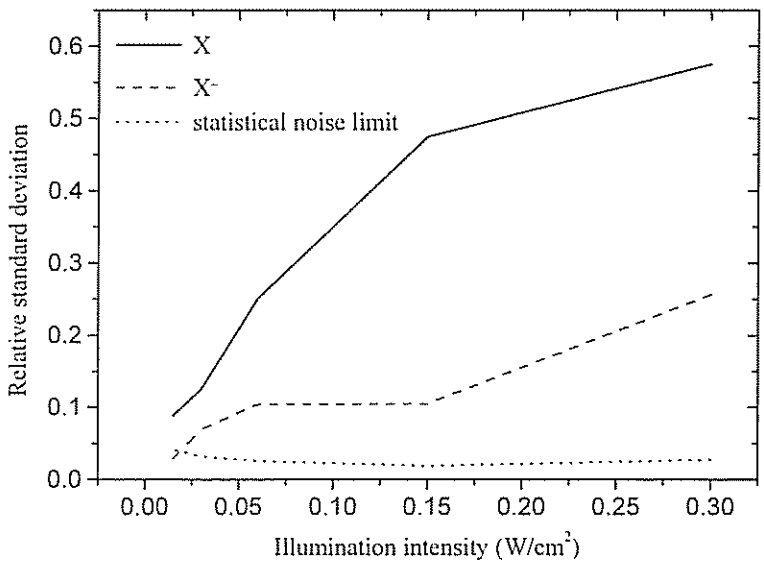


Fig. 35: Relative standard deviation of X and X^- lines.

Fig. 35 shows the relative standard deviation of the temporal intensity fluctuations of the X and X^- peaks as a function of illumination intensity. The dotted line represents the limit of the statistical shot noise, which is the square root of the average. It is seen that at low illumination density the statistical noise becomes the dominant noise source at the X^- line, however, temporal fluctuation in the X intensity persist to the lowest measured illumination intensity.

These temporal fluctuations are probably due to charge movement in the donors layer. Electrons that are activated to the donors layer can move there and change the lateral potential at the 2DEG layer. These changes affect the local population of the X and X^-. Since the neutral exciton has higher oscillator strength then the X^-, it reacts stronger to these changes. These charge migrations occur on large areas with low time scales. This explained the anti-phase behavior of the two excitons since they are both competing on the same number of photo-excited holes.

7 Summary

A low temperature near-field system for PL measurements was developed. The measurements were conducted in a Helium storage Dewar in stable temperature environment of ~4K. The duration of the measurements can be as long as 1 month without heating the microscope. The microscope mechanics enable coarse Z movement of ~1 mm, and coarse X-Y movements of ±1 mm. The scan range is $11x11\mu m^2$. A tuning fork piezo crystal detects the shear force signal during the scan. This signal is fed back to the scanning piezo to maintain a constant height during the scan, hence a surface map of the sample is obtained simultaneously with the optical signal. The operation and stability of the microscope were tested by scanning lines that were formed on a GaAs substrate.

Near-Field PL measurements were conducted at collection mode operation, the illumination was done with a separate single mode fiber. This mode ensures high resolution and uniform excitation of the sample. The spatial resolution of the microscope was checked, it was found that it is determined by the tip diameter. The typical collected PL signal levels were a few hundreds photons/sec.

Photoluminescence measurements of a gated 2DEG were done at various electron densities with a resolution of 250 nm. It was found that at the dense 2DEG limit the width of the PL peak was spatially uniform. Small changes in the intensity were attributed to inhomogeneous distribution of the photo-excited holes. This non-uniform distribution is due to lateral potential fluctuation induced by the remote ionized donors.

As the electron gas depleted below some critical density of a few 10^{10} cm^{-2} the spectrum becomes excitonic. In this density regime it was found that the X^- peak intensity changes across the sample. These intensity fluctuation increases as the 2DEG depleted. The location of the minima and maxima of the fluctuations were found to be fixed and independent of the gate voltage. The intensity of the X peak, was found to be relatively uniform. We explained the intensity changes of the X^- peak by single electron

localization in the lateral potential fluctuations induced by the remote ionized donors. A simple model, based on statistical occupation of these potential fluctuations, was developed to explain our results. The fluctuation of the X^- can also be used to image the donors distribution.

Temporal fluctuations of the X and X^- intensities were found at high excitation intensity. These fluctuations are due to charge movement in the donors layer that chang locally the lateral potential.

[1] D. Kahng and M. M. Atalla, *Silicon-silicon dioxide field induced surface devices*, in IRE-AIEE solid -State Device Res. Conf., (Carnegie Inst. of Technol. Pittsburgh, PA), (1960). D. Kahng, *A Historical Perspective on the Development of MOS Transistors and Related Devices*, IEEE Transactions on Electron Devices, **ED-23**, 655 (1976).

[2] For detailed see S. M. Sze, *Physics of semiconductor devices*, John Wiley and Sons 1981.

[3] R. Dingle, H. Stormer, A. C. Gossard and W. Wiegmann, *Electron Mobilities in Modulation-Doped Semiconductor Heterojunction Superlattices*, Appl. Phys. Lett. **33** 665, (1978).

[4] *Gallium Arsenide*, Edited by J. S. Blakemore, Americen Institute of Physics, (1987).

[5] V. Umansky, R. de-Picciotto and M. Heiblum, *Extremely High-Mobility Two-Dimensional Electron Gas: Evolution of Scattering Mechanism*, Appl. Phys. Lett. **71,** 683 (1997).

[6] P. W. Anderson, *Absence of Diffusion in Certain Random Lattices*, Phys. Rev. **109**, 1492 (1958).

[7] N. F. Mott, *Conduction in Non-Crystalline Materials*, Phil. Mag. **13**, 989 (1966).

[8] N. F. Mott G. A. Davies, *Electronic properties of non-crystalline materials*, 2nd edition, Clarendon press, Oxford (1979).

[9] E. Abrahams, P. W. Anderson, D. C. Licciardello and T. V. Ramakrishnan, *Scaling Theory of Localization: Absence of Quantum Diffusion in Two Dimensions*, Phys. Rev. Lett. **42,** 673 (1979).

[10] P. A. Lee and T. V. Ramakrishnan, *Disordered Electronic Systems*, Review of Mod. Phys. **57**, 287 (1985).

[11] Y. Imry, *Introduction to mesoscopic physics*, Lecture notes.

[12] Z. Ovadyahu and Y. Imry, *On the role of the correlation length near the onset of non-metallic conduction* ,J. Phys. **C16**, L471 (1983).

[13] Y. Liu, B. Nease, K. A. McGeer and A. M. Goldman, *Scaling of the Electrical-Conductivity of Ultrathin Amorphous Palladium Films* ,Europhys. Lett. **19**, 409 (1992).

[14] F. W. Van Keuls, H. Mathur, H. W. Jiang and A. J. Dahm, *Localization Scaling Relation in Two Dimensions: Comparison With Experiment*, Phys. Rev. B **56**, 13263 (1997).

[15] S. V. Kravchenko, G. V. Kravchenko, J. E. Fureaux, V. M. Pudalov and M. D'Iorio, *Possible Metal-Insulator-Transition at B=0 in 2 Dimensions*, Phys. Rev. **B50**, 8039 (1994); S. V. Kravchenko, W. Mason, G. E. Bowker, J. E. Fureaux, V. M. Pudalov and M. D'Iorio, *Scaling of an anomalous metal-insulator transition in a two-dimensional system in silicon at B=0*, Phys. Rev. **B51**, 7038 (1994), S. V. Kravchenko, D. Simonian, M. P. Sarachik, W. Mason, and J. E. Fureaux, *Electric Field scaling at a B=0 Metal-insulator Transition in Two-Dimensions*, Phys. Rev. Lett. **77**, 4938 (1996).

[16] K. Ismail, J. O. Cho, D. Popovic, A. B. Fowler and S. Washborn, preprint cond-mat/9707061 (1997).

[17] Y. Hanin, U. Mairav, D. Shahar, C. C. Li, D. C. Tsui and H. Shtrikman, *The Metallic-Like Conductivity of a Two-dimensional hole system*, Phys. Rev. Lett. **80,** 1288 (1998).

[18] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **84** 168 (1983), *Influence of Coulomb interaction on the properties of disordered metals*, (JETP) **57**, 97 (1983).

[19] C. Jiang, D. C. Tsui and G. Weimann, *Threshold Transport of High-Mobility Two-Dimensional Electron Gas in GaAs/AlGaAs Heterostructures*, Appl. Phys. Lett. **53**, 1533 (1988).

[20] T. Sajoto, Y. W. Suen, L. W. Engel, M. B. Santos and M. Shayegen, *Fractional quantum Hall effect in very low density GaAs/$Al_xGa_{1-x}As$ heterostructures*, Phys. Rev. B **41**, 8449 (1990).

[21] A. L. Efros, F. G. Pikus and V. G. Burnett, *Destiny of States of a Two-Dimensional Electron Gas in Long-Range Random potential*, Phys. Rev. B **47**, 2233 (1993).

[22] H. Haug and S. W. Koch, *Quantum theory of the optical and electronic properties of semiconductors*, World Scientific (1990).

[23] J. A. Nixon and J. H. Davies, *Potential Fluctuations in Heterostructure Devices*, Phys. Rev B **41**, 7929 (1990).

[24] G. Finkelstein, H. Shtrikman and I. Bar-Joseph, *Optical Spectroscopy of a Two-Dimensional Electron Gas Near the Metal-Insulator Transition*, Phys. Rev. Lett. **74**, 976 (1995).

[25] A. J. Shilds, M. Pepper, D. A. Ritchie, M. Y. Simmons and G. A. C. Jones, *Quenching of Excitonic Optical Transitions by Excess Electrons in GaAs Quantum Wells*, Phys. Rev. B **51**, 18049 (1995).

[26] G. Bastard, *Wave mechanics applied to semiconductor heterostructures*, Les editions de physique, Paris (1988).

[27] S. Schmitt-Rink, D. S. Chemla and D. A. B. Miller, *Linear and Nonlinear Properties of Semiconductor Quantum Wells*, Adv. Phys. **38**, 89 (1989).

[28] F. Stern and S. Das Sarma, *Electron energy levels in GaAs-$Ga_{1-x}Al_xAs$ heterojunctions*, Phys. Rev. B **30**, 840 (1984).

[29] R. Kuchler, G. Abstreiter, G. Bohm and G. Weimann, *Photoluminescence Lineshape of Narrow N-Type Modulation-Doped Quantum-Wells*, Semicond. Sci. Technol. **8**, 88 (1993).

[30] G. D. Mahan, *Excitons in Degenerate Semiconductors*, Phys. Rev. **153**, 882 (1967).

[31] G. D. Mahan, *Excitons in Metals: Infinite Hole Mass*, Phys. Rev. **163**, 612 (1967).

[32] M. S. Skolnik, J. M. Rorison, K. J. Nash, D. J. Mowbray, P. R. Tapster, S. J. Bass and A. D. Pitt, *Observation of Many-Body Edge singularity in Quantum-Well Luminescence Spectra*, Phys. Rev. Lett. **58**, 2130 (1987).

[33] R. L. Green, K. K. Bajaj and D. E. Phelps, *Energy levels of Wannier excitons in GaAs-$Ga_{1-x}Al_xAs$ quantum-well structures*, Phys. Rev. B **29**, 1807 (1984).

[34] K. Kheng, C. T. Cox, Y. Merle dAubigune, F. Bassani, K. Saminadayar and S. Tatarenko, *Observation of Negatively Charged Excitons X^- in Semiconductors Quantum Wells*, Phys. Rev Lett **71**, 1752 (1993).

[35] M. A. Lampert, *Mobile and immobile effective-mass-particle complexes in nonmetallic solids*, Phys Rev. Lett. **1**, 450 (1958).

[36] B. Stebe and A. Ainane, *Ground State Energy and Optical Absorption of Excitonic Trions in Two Dimensional Semiconductors*, Superlattices and Microstruct. **23**, 545 (1989).

[37] A. J. Shields M. Papper, D. A. Ritchie and M. Y. Simmons, *Influence of excess electrons and magnetic fields on Mott-Wannier excitons in GaAs quantum wells*, Adv. in Phys. **44**, 47 (1995).

[38] M.A. Paesler, P.J.Moyer, *Near-Field Optics*, Chap 4, Wiley-Interscience press, New York, (1996).

[39] M. Born, and E. Wolf, *Principles of Optics*, Pergamon press, London (1965).

[40] G.A. Massey, *Microscopy and pattern generation with scanned evanescent waves*, Appl. Opt. **23**, 658 (1984).

[41] J. M. Vigoureux and D. Courjon, *Detection of Nonradiative Fields in Light of the Heisenberg Uncertainty Principle and the Rayleigh Criterion*, Appl. Opt. **31**, 3170 (1992).

[42] J. M. Vigoureux, F. Depasse, and C. Girard, *Superresolution of near-field optical microscopy defined from properties of confined electromagnetic waves*, Appl. Opt. **31**, 3036 (1992).

[43] H. F. Hess, E. Betzig, T. D. Harris, L. N. Pfeiffer and K. W. West, *Near-Field Spectroscopy of the Quantum Constituents of a Luminescent System*, Science **264**, 1740 (1994).

[44] C. Weisbuch, R. Dingle, C. Gossard and W. Wiegmann, *Optical characterization of interface disorder in GaAs-$Ga_{1-x}Al_xAs$ multi-quantum well structures*, Solid State Comm. **38**, 709 (1981).

[45] K. Brunner, G. Abstreiter, G. Bohm, G. Trankle and G. Weimann, *Sharp-Line Photoluminescence and two-Photon Absorption of Zero-Dimensional Biexcitons in a GaAs/AlGaAs structure*, Phys. Rev. Lett. **73**, 1138 (1994).

[46] F. Flack, N. Samarath, V. Nikitin, P.A. Crowell, J. Shi, J. Levi and D. D. Awschalom, *Near-field optical spectroscopy of localized excitons in strained CdSe quantum dots*, Phys. Rev. B **54**, R17312 (1996).

[47] J. Hasen, L. Pfeiffer, A. Piczuk, S. He, K. W. West and B. S. Dannis, *Metamorphosis of a quantum wire into quantum dots*, Nature **390,** 54 (1997).

[48] R. D. Grober, T. D. Harris, J. K. Trautman, E. Betzig, W. Wegscheider, L. Pfeiffer and K. W. West, *Optical spectroscopy of a GaAs/AlGaAs quantum wire structure using near-field scanning optical microscopy*, Appl. Phys. Lett. **64**, 1421 (1994).

[49] L. Pfeiffer, K. W. West, H. L. Stormer, J. P. Eisenstain, K. W. Baldwin, D. Gershoni and J. Spector, *Formation of high quality two-dimensional electron gas on cleaved GaAs*, Appl. Phys. Lett. **56,** 1697 (1990).

[50] T. D. Harris, D. Gershoni, R. D. Grober, L. Pfeiffer, K. W. West and N. Chand, *Near-Field Optical spectroscopy of Single Quantum Wire*, Appl. Phys. Lett. **68,** 988 (1996).

[51] A. Richter, G. Behme, M. Suptitz, Ch. Lienau, T. Elsaesser, M. Remsteiner, R. Notzel and K. H. Ploog, *Real Space Transfer and Trapping of Carriers into single GaAs Quantum Wires Studied by Near-Field Optical spectroscopy*, Phys. Rev. Lett. **79,** 2145 (1997).

[52] R. Notzel, M. Ramsteiner, J. Menniger, A. Trampert, H. P. Schonherr, L. Daweritz and K. H. Ploog, *Patterned Growth on High-Index GaAs (N-11) Substrates - Application to Sidewall Quantum Wires*, J. Appl. Phys. **80**, 4108 (1996).

[53] Y. Toda, M. Kourogi, M. Ohtsu, Y. Nagamune and Y. Arakawa, *Spatially and Spectrally resolved imaging of GaAs quantum–dot structures using near-field technique*, Appl. Phys. Lett. **69**, 827 (1996).

[54] Y. Nagamune, S. Tsukamoto, M. Nishioka and Y. Arakawa, *Growth-Process and Mechanism of Nanometer-Scale GaAs Dot-Structures Using MOCVD Selective Growth*, J. Cryst. Growth **126**, 707 (1993).

[55] J. Levy, V Nikitin, J. M. Kikkawa, A, Cohen, N. Samath, R. Garcia and D. D. Awschalom, *Spatiotemporal Near-Field Spin Microscopy in Patterned Heterostructures* Phys. Rev. Lett. **76,** 1948 (1996).

[56] G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, *Surface Studies by Scanning Tunneling microscopy*, Phys. Rev. Lett. **49**, 57 (1982).

[57] G. Binning, C. F. Quate and Ch. Gerber, *Atomic Force Microscope*, Phys. Rev. Lett. **56**, 930 (1986).

[58] G. Binning and D. P. E. Smith, *Single-tube three-dimensional scanner for scanning tunneling microscopy* Rev. Sci. Instrum. **57**, 1688 (1986).

[59] K. G. Vandervoot, R. K. Zasadzinski, G. G. Galicia and G. W. Crabtree, *Full Temperature Calibration from 4-K to 300-K of the Voltage Response of Piezoelectric Tube Scanner PZT-5A for Use in Scanning Tunneling Microscopes*, Rev. Sci. Instrum. **64**, 896 (1993).

[60] A. Harootunian, E. Betzig, M. Isaacson and A. Lewis, *Super-resolution fluorescence near-field optical microscopy*, Appl. Phys. Lett. **49**, 674 (1986).

[61] D. W. Pohl, W. Denk and M. Lanz, Appl. *Optical stethoscopy: Image recording with resolution of $\lambda/20$*, Phys. Lett. **44**, 651 (1984).

[62] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner and R. L. Kostelak, *Braking the Diffraction Barrier: Optical Microscopy on a Nanometric Scale*, Science **251**, 1468 (1991).

[63] G. A. Valaskovic, M. Holton and G. H. Morrison, *Parameter control, characterization, and optimization in the fabrication of optical fiber near-field probes*, Appl. Opt. **34**, 1215 (1995).

[64] T. Saiki, s. Mononobe, M. Ohtsu, N. Saito and J. Kusano, *Tailoring a high-transmission fiber probe for photon scanning tunneling microscope*, Appl. Phys. Lett. **68**, 2612 (1996), see ref. there.

[65] P. Hoffmann, B. Dutoit and R. P. Salathe, *Comparison of mechanically drawn and protection layer chemically etched optical fiber tips*, Ultramicroscopy **61,** 165 (1995).

[66] K. M. Takahashi, *Meniscus Shapes on Small Diameter Fibers*, J. of Colloid and Interface Science **134**, 181 (1990).

[67] S. J. Bukofsky and R. D. Grober, *Video-Rate Near-Field Scanning Optical Microscopy*, Appl. Phys. Lett. **71**, 2749 (1997).

[68] M. A. Taubenblatt, *Lateral force and topography using scanning tunneling microscopy with optical sensing of the tip position*, Appl. Phys. Lett. **54**, 801 (1989).

[69] E. Betzig, P. L. Finn and J. S. Weiner, *Combined shear force near-field scanning optical microscopy*, Appl. Phys. Lett. **60,** 2484 (1992)

[70] R. Toledo-crow, P. C. Yang, Y. Chen and M. Vaes-Iraveni, *Near-field differential optical microscopy with atomic force regulation*, Appl. Phys. Lett. **60**, 2957 (1992).

[71] A. Shchemelinin, M. Rudman, K. Lieberman and A. Lewis, *A simple lateral force sensing technique for near-field micropattern generation*, Rev. Sci. Instrum. **64**, 3538 (1993).

[72] A. Drabenstedt, J. Wrachtrup, and C. von Borczyskowski, *A Distance Regulation Scheme for Scanning Near-Field Optical Microscopy*, Appl. Phys. Lett. **68,** 3497 (1996).

[73] J. Barenz, O. Hollricher, and O. Marti, *An Easy-to-Use Non-Optical Shear-Force Distance Control for Near-Field Optical Microscopes*, Rev. Sci. Instrum. **67**, 1912 (1996).

[74] W. A. Atia and C. C. Davis, *A Phase-Locked Shear-Force Microscope for Distance Regulation in Near-Field Optical Microscopy*, Appl. Phys. Lett. **70**, 405 (1997).

[75] K. Karrai and R. D. Grober, *Piezoelectric tip-sample distance control for near field optical microscopes*, Appl. Phys. Lett. **66,** 1842 (1995).

[76] C. L. Jahncke and H. D. Hallen, *A Versatile Stable Scanning Proximal Probe Microscope*, Rev. Sci. Instrum. **68**, 1759 (1997).

[77] M. J. Gregor, P. G. Blome, J. Schöfer, and R. G. Ulbrich, *Probe-surface interaction in near-field optical microscopy: The nonlinear bending force mechanism*, Appl. Phys. Lett. **68**, 307 (1996).

[78] R. D. Grober, T. D. Harris, J. K. Trautman, and E. Betzig, *Design and implementation of a low temperature near-field scanning optical microscope*, Rev. Sci. Instrum. **65**, 626 (1994).

[79] C. Obermuller, K. Karrai, G. Kolb and G. Abstreiter, *Transmitted radiation through a subwavelength-sized tapered optical fiber tip*, Ultramicroscopy **61**, 171 (1995)

[80] R. W. Boyd, *Nonlinear Optics*, Academic Press Inc. New York, Chap. 9.

[81] R. H. Stolen, *Parametric amplification and Frequency Conversion in Optical Fibers*, Proceeding of the IEEE **68**, 1232 (1980).

Publication related to this work

1. G. Eytan, Y. Yayon, M. Rappaport, H. Shtrikman and I. Bar-Joseph, *Near-Field Spectroscopy of a Gated Electron Gas: Direct Evidence of Localization,* to be published in Phys. Rev. Lett. August (1998).

2. G. Eytan, Y. Yayon, I. Bar-Joseph and M. Rappaport, *A storage Dewar near-field scanning optical microscope*, submitted to Review of Scientific Instruments.

3. G. Eytan, Y. Yayon, M. Rappaport, H. Shtrikman and I. Bar-Joseph, *Near-Field Spectroscopy of a Gated Electron Gas: Direct Evidence of Localization,* to be published in the Proceeding of the 24 International Conference on Semiconductor Physics, world Scientific (1998)

Certificate of originality

This thesis describes my original work. Parts of the work were performed in cooperation with the assistance of other researchers and research students in the Physics department of the Weizmann Institute as follows:

Developments of the latest version of the cryogenic NSOM were done together with Dr. Michael Rappaport and the MSc. student Yossi Yayon.

Measurements of the optical noise in fibers were taken together with Yossi Yayon.

תקציר

בשנים האחרונות הייתה התפתחות גדולה בתחום מיקרוסקופי הסריקה, מספר מיקרוסקופים הומצאו ביניהם מיקרוסקופ השדה הקרוב. מיקרוסקופ זה מאפשר אינטראקציה אופטית עם השכבה הנמדדת ברזולוציה מרחבית הקטנה בהרבה מאורך הגל של האור שבו משתמשים. במקביל להתפתחות זו חלה התקדמות עצומה בשיטות גידול גבישים חצי מוליכים. התקדמות זו מאפשרת יצירה של מערכות בעלות מספר ממדים הקטן משלוש כדוגמת גז אלקטרונים דו-ממדי. בשני העשורים האחרונים נעשו מחקרים רבים בתחום זה.

בעבודת מחקר זו אנו משלבים בין שני תחומים טכנולוגיים ומחקריים אלו, ומשתמשים במיקרוסקופ שדה קרוב כדי למדוד פוטולומיניסנציה של גז אלקטרונים דו-ממדי ברזולוציה מרחבית גבוהה.

שני הפרקים הראשונים בעבודה זו מציגים את הרקע הרלוונטי בנושא גז אלקטרונים דו ממדי וספקטרוסקופיה שלו. פרק 3 מסביר מספר נקודות תאורטיות וניסיוניות הקשורות למיקרוסקופית שדה קרוב.

מאמץ רב הושקע במהלך העבודה על מנת לבנות מערכת מיקרוסקופ שדה קרוב לטמפרטורות נמוכות ולאפיין מערכת זו. עבודות הפיתוח של אמצעי החישה של השטח על ידי הטיפ ושל אפיון ביצועי הטיפ מתוארים בפרק 4. תאור המבנה המכני של המיקרוסקופ והשיקולים שהנחו אותנו בבניית מבנה כזה מתוארים בפרק 5. כמו כן מכיל פרק זה שורת מדידות שנעשו לאפיון פרמטרים שונים של מערכת המדידה.

לבסוף, בפרק 6, אנו מביאים שורת מדידות שדה קרוב של פוטולומיניסנציה של גז אלקטרונים דו-ממדי בצפיפות משתנה. ממדידות אלו עולה כי כאשר מדללים את גז האלקטרונים מתחת לצפיפות קריטית מסוימת עוצמת הלומינסנציה נעשית אי-אחידה מרחבית. אי-אחידות זו נגרמת כתוצאה מהתמקמות האלקטרונים ביחידות בתוך בורות פוטנציאל, בורות אלו נוצרים בגלל התפלגות מרחבית אי אחידה של האטומים התורמים. אטומים אלו טעונים במטען חיובי ומשרים פוטנציאל אקראי בשכבת האלקטרונים. אנו מדגימים כי ניתן להשתמש בשינויי עוצמת הלומינסנציה כדי למפות את ההתפלגות המרחבית של האטומים התורמים. בנוסף לכך אנו מראים כי כאשר מעוררים את הדוגמא בעוצמה גבוהה נגרמות פלוקטואציות זמניות בעוצמת הלומינסנציה. פלוקטואציות אלו נגרמות כתוצאה מזרמי מטען בתוך שכבת האטומים התורמים.

ספקטרוסקופית שדה-קרוב של גז אלקטרונים דו-ממדי

חיבור לשם קבלת התואר "דוקטור לפילוסופיה"

מאת:

גיא איתן

המחלקה לפיסיקה

מכון ויצמן למדע

מוגש למועצה המדעית של מכון ויצמן למדע

יולי 1998