MASTER

Fabrication and characterization of GaAs/(Al,Ga)As quantum dots

Verschuren, C.A.

Award date:
1995

Fabrication and characterization of GaAs/(Al,Ga)As quantum dots

Coen Verschuren

Verslag van een afstudeerproject verricht bij de Sharp Laboratories of Europe Ltd., Oxford, UK. Dit afstuderen vond plaats onder toezicht van de afdeling Vaste Stof Fysica van de faculteit Technische Natuurkunde, Technische Universiteit Eindhoven.

Report of a graduation project performed at the Sharp Laboratories of Europe Ltd. in Oxford (UK). This project was carried out under supervision of the Solid State Division of the Department of Physics of the Eindhoven University of Technology.

Supervisors in the UK:	Dr. G. Duggan, Dr. T.D. Bestwick, Dr. M. Dawson.
Supervisor in the Netherlands:	Prof.Dr. J.H. Wolter

Summary

Electron beam lithography and electron cyclotron resonance (ECR) plasma etching have been used to fabricate uniform arrays of free-standing quantum dots with dimensions in the range 25-60 nm from GaAs/AlGaAs 5 nm single quantum well material. The 25 nm dots are believed to be the smallest ever produced in this way.

The single quantum well material has been characterized using low temperature (5 K) photoluminescence (PL) and photoluminescence excitation spectroscopy before and after processing into quantum dots. Although the etched surfaces of the dots are not intentionally passivated and remain exposed to air, strong PL is measured from the features. The PL linewidth from the dots remains in the range 8 to 4 meV (FWHM) compared with the 3 meV linewidth from the unprocessed SQW material, indicating high uniformity. The measured PL intensity, when corrected for the area of the quantum well remaining in the dots, increases rapidly as the dot diameter is reduced: The integrated PL intensity of the 25 nm dots is more than 150 times greater than the reference unprocessed SQW excited under the same conditions, which is a remarkable increase.

PLE shows clear free heavy and light hole excitonic features in the quantum dots. They are increasingly blue-shifted with decreasing dot diameter, due to the additional lateral confinement of the carriers. The measured positions of the heavy and light hole excitons shift by less than 1 meV in 60 nm dots, but are shifted by 8 and 20 meV respectively in 25 nm diameter dots. PLE is preferable to PL when determining blue-shifts because, representing predominantly the absorption spectrum, it is less sensitive to local fluctuations in the quantum well width. This study is the clearest observation yet of light and heavy hole excitons in a range of quantum dots of these sizes.

Table of contents

Introduction

The refined growth technique of molecular beam epitaxy (MBE) has allowed the realization of ultrathin, 2-D, semiconductor heterostructures such as quantum wells. Their applications as quantum well lasers and high electron mobility transistors (HEMTs) are a huge success. Since it has been anticipated that further improvements in devices would be gained, it was a natural trend to continue to diminish systems' dimensionality to 1-D (quantum well wire, QWW) and 0-D (quantum dot, QD). Though spectacular effects have already been demonstrated in transport properties, the progress in the field of 1-D and 0-D structures for optical applications is much slower. The necessity of confining both electrons and holes, together with that of large confinement energies for room temperature devices lead to stringent requirements on fabrication processes. No ideal technique has been found so far, and numerous approaches are used to produce the wires and dots. These techniques can be classified in two main groups: the techniques starting with a quantum well, providing confinement in one direction, and adding confinement in one or the two remaining directions, and those where a single growth step produces the wires or dots. Most of them are potentially able to yield structures with lateral dimensions of the order of 10 nm (25 nm MOCVD-grown pyramids [Nagamune94], 30 nm discs through electron beam lithography, reactive ion etching and anodic oxidation [Marzin94]), so that the key issues now concern the mastering of size fluctuations and reducing the surface damage introduced in the fabrication process. These effects tend to mask intrinsic phenomena through severe linewidth broadening and orders of magnitude reduction in luminescence efficiency, making systematic studies difficult. The aim of this project was to fabricate uniform and efficient arrays of quantum dots by a combination of electron beam lithography and low-damage electron cyclotron resonance (ECR) plasma etching. The fabricated dots were assessed by low temperature photoluminescence (PL) and photoluminescence excitation (PLE) spectroscopy.

1. Lower dimensional systems

1.1. Introduction

In order to investigate the fundamentals in fabricating and characterizing these structures, the well-known GaAs/AlGaAs semiconductor system is used.

GaAs, $Al_xGa_{1-x}As$ and AlAs are so-called III-V compound materials: their primitive unit cell consists of one atom from the third column of the Periodic Table (Ga or Al) and one from the fifth (As). $Al_xGa_{1-x}As$ is a random alloy (unit cells consist of either AlAs or GaAs), with x the aluminium concentration.

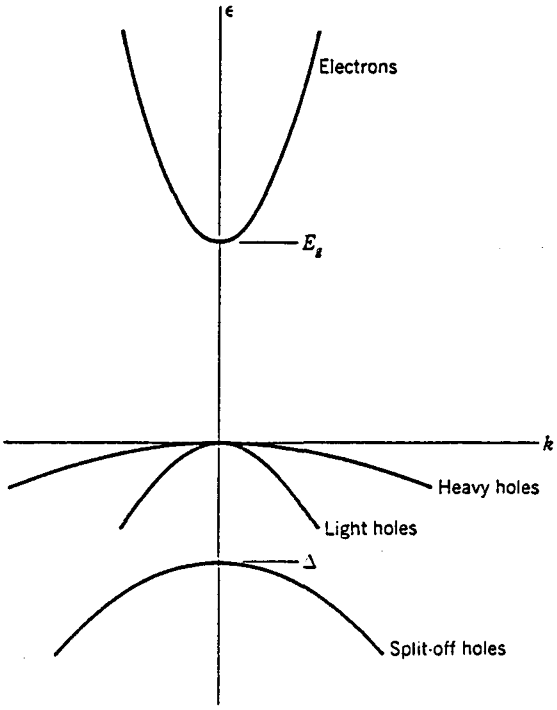


Fig. 1.1. *Simplified view of the band edge structure of a direct gap semiconductor. [Kittel76]*

For GaAs both the conduction and the valence band extrema occur at k=0, and is therefore called a direct gap material: it can emit light through the recombination of an electron and a hole, whereas indirect materials such as Si can not (without violating k-conservation). As is shown in Fig. 1.1, the valence bands are threefold. The upper band is called the heavy hole band, because the corresponding effective mass, which is inversely proportional to the second derivative of E with respect to k, is large. The other band, whose curvature is larger, is therefore referred to as the light hole band. At k=0 these bands are degenerate. The third band is split off by the spin-orbit splitting. [Kittel76]

GaAs has a relatively large energy gap (1.52 eV at 4 K). The addition of aluminium widens the gap (AlAs: 2.36 eV), but causes little disturbance otherwise: GaAs and AlAs have almost the same lattice constants (a=5.65Å and a=5.66Å respectively). Although AlAs is indirect, it is possible to grow lattice-matched, direct-gap heterostructures using the pair of GaAs and $Al_xGa_{1-x}As$, which is why it is so widely used. The aluminium concentration x is usually chosen to be around 0.3 to ensure that the potential barrier in the conduction band is large but that the alloy is still a direct gap material.

1.2. $GaAs/Al_{0.3}Ga_{0.7}As$ quantum wells

The simplest quantum situation to be dealt with consists of a single layer of material (GaAs) embedded between two thick layers of another material ($Al_{0.3}Ga_{0.7}As$), which has a larger bandgap and with the band discontinuities such that both band edges of the smaller gap material are below those of the wide-gap material. See Fig. 1.2.

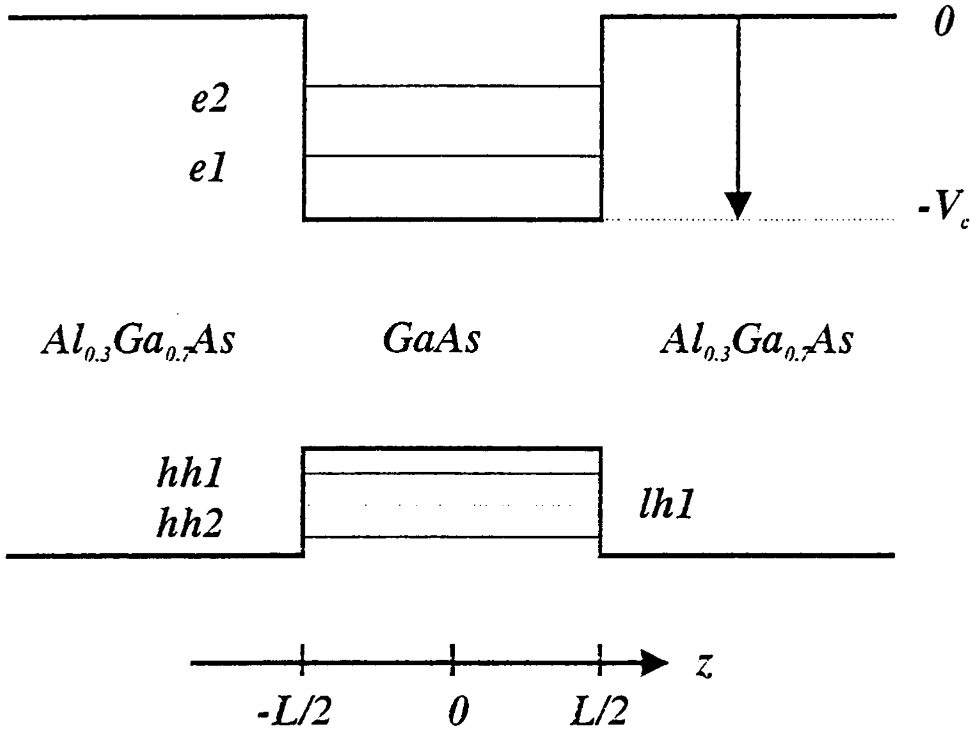


Fig. 1.2. *Electronic bandstructure of a quantum well.*

In this way both electrons and holes are confined in the GaAs, hence the name 'quantum well'. The GaAs layer must be thick enough ($L>20$ Å) so that the layer has the properties of macroscopic GaAs and thin enough to ensure that L is smaller than the de Broglie wavelength. Both materials are assumed to be intrinsic, so that no band bending occurs.

1.2.1. Energy levels

The calculation of the energy levels in the conduction band is rather straightforward, very similar to the simple particle in a box problem. The Schrödinger equation is

$$\left(-\frac{\hbar^2}{2m^*(z)}\frac{\partial^2}{\partial z^2} + V_c(z)\right)\chi_n(z) = \epsilon_n\chi_n(z) \tag{1.1}$$

where $m^*(z)$ is the electron effective mass of the material, $V_c(z)$ represents the energy level of the bottom of the conduction bands, and ε_n is the so-called confinement energy of the carriers.

The continuity conditions at the interfaces are that $\chi_n(z)$ and $[1/m^*(z)][\partial\chi_n(z)/\partial z]$ should be continuous. Using these conditions, equation (1.1) can be solved exactly to yield the wave functions and energies [Weisbuch91]. See also Fig. 1.3. The even solutions are

$$|z|<L/2\ :\ \chi_n(z) = A\cos(k_w z) \tag{1.2}$$

and

$$\epsilon_n = -V_c + \frac{\hbar^2 k_w^2}{2m^*_{GaAs}} \tag{1.3}$$

$$|z|>L/2\ :\ \chi_n(z) = B\exp(-\kappa_b z) \tag{1.4}$$

and

$$\epsilon_n = -\frac{\hbar^2 \kappa_b^2}{2m^*_{AlGaAs}} \tag{1.5}$$

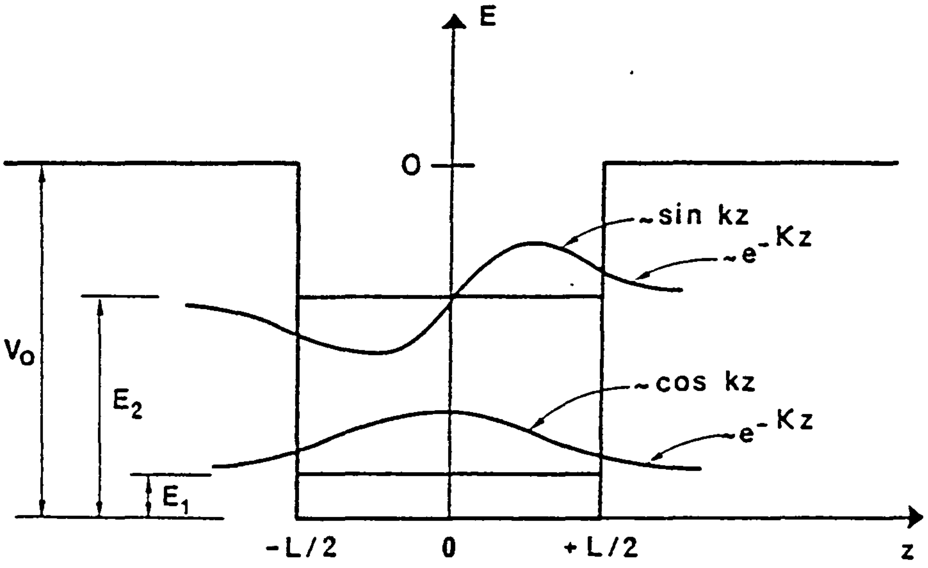


Fig. 1.3. *First two bound energy levels and wave functions in a finite quantum well. [Weisbuch91]*

The continuity conditions yield equations for k_w and κ_b, which can be solved numerically or graphically [Schiff68, Weisbuch91]. The effect of (2-D) confinement is that the energy spectrum in the direction perpendicular to the interface is turned into a series of discrete levels.

In the direction parallel to the interface, no influence from the quantum well is expected. Hence, the total energy of an electron in the lowest confined level (l=1) at the conduction band measured from the bottom of the bulk band is given by equation (1.3). In Fig. 1.4 this energy is plotted as a function of k: the subbands replace the continuous bulk band, and in the direction z perpendicular to the interface the confined states behave as bound states. In the right-hand side of the bottom half of the figure, the density of electron states, ρ, is shown, which is associated with the sub-band structure on the left as a function of energy E. In a bulk (3-D) parabolic band, the density of states goes as $E^{1/2}$, as indicated by the broken line. The 2-D density of states is given by

$$\rho_{2\text{-}D}(E) = \frac{l m^*}{\pi \hbar^2} \qquad (1.6)$$

with l=1,2,3,..., again in the parabolic approximation.

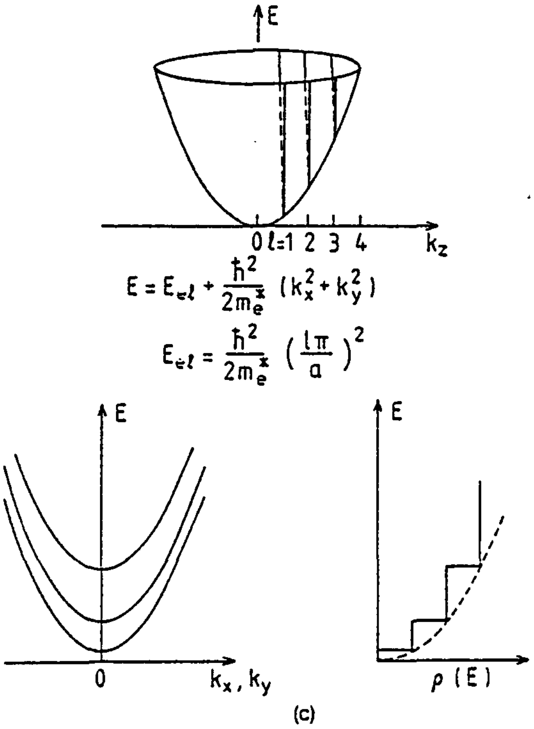


Fig. 1.4. *The construction of the total energy of an electron in the conduction band of a quantum well structure; z is the direction perpendicular to the interface;* ρ *is the density of states (the broken line indicates the form for a bulk crystal). [Jaros90]*

This means that the density of states at the bottom of the conduction band of the quantum well is finite, whereas that for the bulk GaAs is zero: all dynamic phenomena remain finite at low kinetic energies and low temperatures, such as scattering, optical absorption, and gain.

The procedure used above to describe confinement of electrons at the conduction band edge can also be applied to the valence band. In fact, the situation is more complex, see e.g. [Weisbuch91] chapter II. Of course, the barrier height and the effective mass are different. It is customary to measure the energy of the confined levels at the valence band from the top of the valence band. Then the picture of the total energy of the valence subbands is formally the same as that shown in Fig. 1.4. However, as shown in Fig. 1.1, there are two different bands at the top of the valence band, the heavy hole and the light hole band, with different effective masses. Hence, two sets of confined levels derived from the valence band edge are expected, one associated with m_{HH}^* and one with m_{LH}^*. See Fig. 1.5. The conduction band states appear twice simply for convenience. It can also be seen that the degeneracy between the HH and LH bands, present in bulk for k_x, k_y = 0 (see section 1.1) is lifted in the quantum well situation.

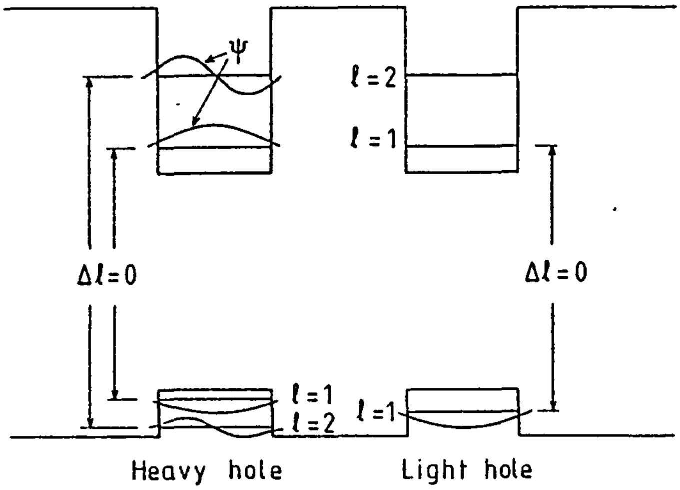


Fig. 1.5. *The band diagram for a quantum well, indicating allowed optical transitions between the confined levels in the conduction band and the confined levels associated with the heavy (left-hand side) and light (right-hand side) hole valence bands; l labels the confined levels in each well as in Fig. 1.4. Only transitions between levels with the same l are possible.*

1.2.2. Optical properties

Light of frequency ω applied to a bulk crystal of GaAs can be absorbed by valence electrons when the photon energy ħω is larger than or equal to the magnitude of the forbidden gap. As is shown above, the effect of confinement changes the energy of conduction and valence states in such a way that the energy needed in order to take an electron from the valence band into the conduction band is larger than the bulk gap of GaAs. In other words, the principal transition across the gap (between the lowest lying states with l=1), which determines the threshold for absorption and emission, will show a blue-shift that depends on the confinement effect.

However, the absorption spectrum is not simply determined by the creation of a free electron and a free hole. The electron and the hole are correlated in their motion (this special 'particle' is called the exciton) in a way that can be described as the simple Coulomb attraction of the electron and the hole. As in the case of the hydrogen atom, this attraction leads to bound states, the lowest of which is one effective Rydberg (Ry^*) below the continuum level and in which the electron and hole are bound to each other within an effective Bohr radius a_B^*. The effective Rydberg and Bohr radii are given by

$$\mathrm{Ry}^* = \frac{2m_{red}e^4}{\hbar^2(8\pi\epsilon)^2} = \frac{m_{red}}{m_0}\frac{1}{\epsilon_r^2}\mathrm{Ry} \tag{1.7}$$

$$a_B^* = \frac{4\pi\epsilon\hbar^2}{m_{red}e^2} = \epsilon_r\frac{m_0}{m_{red}}a_B \tag{1.8}$$

where m_{red} is the reduced mass of electron and hole ($1/m_{red}=1/m_e+1/m_h$); ε_r is the relative permittivity of the semiconductor; Ry=13.6 eV; and a_B=0.529 Å. From $a_B^*\approx100$ Å, one infers that the wave function and energy levels of excitons are quite modified in a quantum well where the thickness is usually of the order of or smaller than the Bohr diameter $2a_B$.

In the limiting exact 2-D case, one should obtain the usual 2-D Rydberg value $\mathrm{Ry}_{2\text{-D}}=4\mathrm{Ry}_{3\text{-D}}$ for the infinitely deep well. The energy levels are then given by

$$\mathrm{Ry}_{n,2\text{-D}} = \mathrm{Ry}_{3\text{-D}}[1/(n-\tfrac{1}{2})^2] \tag{1.9}$$

For finite well thickness, exciton binding energies have been calculated using a variational method. See [Weisbuch91] for an overview and references, and [Haug90] chapters 4 and 10 for details. The exciton binding energies as a function of well thickness increase with decreasing well thickness from 4.2 meV in bulk GaAs to 9 meV in 5 nm wells with $Al_{0.3}Ga_{0.7}As$ barriers. Further reduction of the well thickness results in a strong decrease in binding energy: the exciton can no longer be confined in the well, i.e. the wavefunctions leak into the barrier and the bulk case is found again. Fig. 1.6 shows a comparison of the absorption coefficients due to (a) 3-D or (b) 2-D excitons. The broken lines indicate the density of states.

1.3. GaAs/$Al_{0.3}Ga_{0.7}As$ quantum dots

The quantum dots discussed here are fabricated by etching cylindrical pillars from (nominally) 5 nm single quantum well material, see Fig. 1.7. In the ρ ($\rho^2 = x^2 + y^2$) direction the confinement is complete due to the GaAs/air interface, so there the model of infinite confinement for a cylinder can be used to calculate the energy levels. In the z direction there

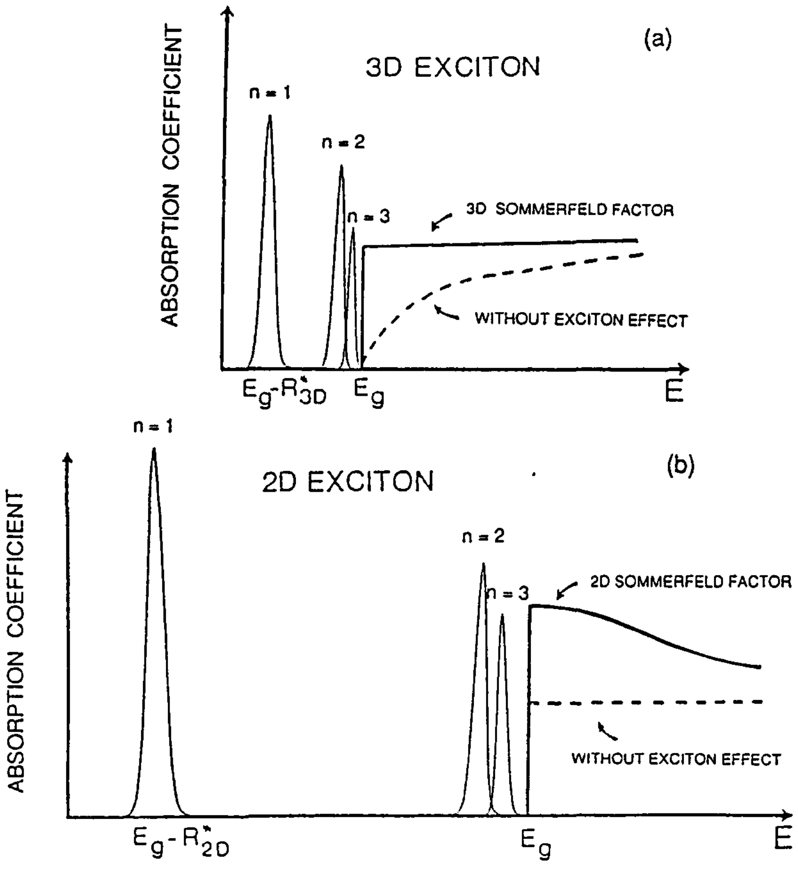


Fig. 1.6. *Comparison of the absorption coefficients due to (a) 3-D or (b) 2-D excitons. [Weisbuch91]*

is wavefunction leakage from the GaAs into the AlGaAs, necessitating using a finite confinement of a quantum well, as discussed in section 1.2.

The total confinement energy is given by

$$\epsilon_{conf}^{total} = \epsilon_{conf}^{z} + \epsilon_{conf}^{\rho} \qquad \textbf{(1.10)}$$

with ε_{conf}^{z} and $\varepsilon_{conf}^{\rho}$ discussed in section 1.2 and below, respectively. The differences between the quantum well and the quantum dot will be due purely to the change in confinement in the ρ direction. Therefore, only $\varepsilon_{conf}^{\rho}$ will be considered here.

The electrons and holes are confined in a cylindrical potential, and the wavefunctions have to vanish at the boundaries ($\rho=R$). Changing to cylindrical coordinates and assuming infinite barriers, the problem can be solved analytically [Banyai87]: from the boundary conditions it follows that the wavefunctions must be Bessel functions (order zero). The first zero of $J_0(x) = 0$ is $\alpha_0 = 2.405$, so that $J_0(\alpha_0\rho/R)$ has its first zero at $\rho=R$. Normalizing the wavefunction yields

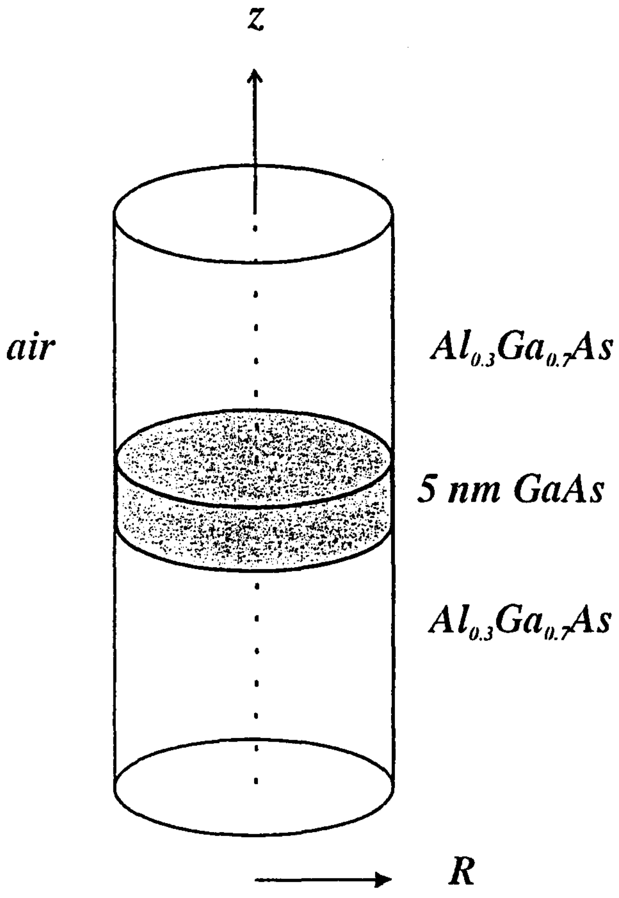


Fig. 1.7. *Schematic diagram of a quantum dot.*

$$\chi(\rho) = \frac{J_0(\alpha_0 \rho / R)}{\sqrt{\pi} R J_1(\alpha_0)} \tag{1.11}$$

The corresponding confinement energy is

$$\epsilon = \frac{\alpha_0^2 \hbar^2}{2m^* R^2} \tag{1.12}$$

where m^* is the effective mass of the electron, light hole or heavy hole. As can be seen from (1.12) the confinement energy increases quadratically with decreasing wire (or dot) radius.

In fact, the properties of excitons confined in isolated quantum dots are determined by the competing effects of the electron-hole attraction and the three-dimensional size quantization. [Bryant91] calculated the ground state energy of a confined heavy-hole exciton in a square GaAs dot (width = 0) by use of a variational approach. The results are qualitatively the same for dots with a small finite thickness and for light-hole excitons. In large dots ($R > 50$ nm) the Coulomb interaction, which scales as $1/R$, dominates and the excitons are bulk-like: the exciton binding energy is enhanced, thereby acting against the blue-shift due to the additional confinement. See also chapter 5 and [Xia89]. For decreasing dot diameters, the confinement effects (1.12) become important, and are dominant for dot sizes comparable to the free-exciton radius, which is 10.9 nm. In much smaller dots, which cannot confine the exciton, the $1/R^2$ scaling will break down.

The density of states for a quantum wire goes as $E^{-1/2}$, whereas that for bulk is $E^{1/2}$ and for 2-D is constant for each sub-band (see section 1.2 and broken lines in Fig. 1.8). In the case of quantum dots the density of states is a delta-function at the discrete energy of each dot. This enormously sharpens resonant behaviour and therefore energy selectivity. Another important property is the lower dispersion of optical properties over k-states. Taking the injection laser as an example (Fig. 1.8), the occupancy of the same number of electrons of 2-D, 1-D and 0-D states above inversion will of course lead to higher gains due to the concentration of electrons and holes over fewer k-states, as long as only one or very few quantized states are within the relevant energy range from the ground state, usually $\sim kT$. Another way of seeing this is to remark that the saturated gain per carrier for fully inverted conduction and valence bands is independent of dimensionality. Therefore, concentrating carriers in a phase-space that is less extended in energy will lead to a larger maximum in the spectral gain curve, although the integrated gain is of course constant.

In conclusion, two effects are expected when the dimensionality is reduced from 2-D to 0-D:

1. a blue-shift in the optical transitions,
2. an increase in luminescence intensity,

both increasing with decreasing dot diameter.

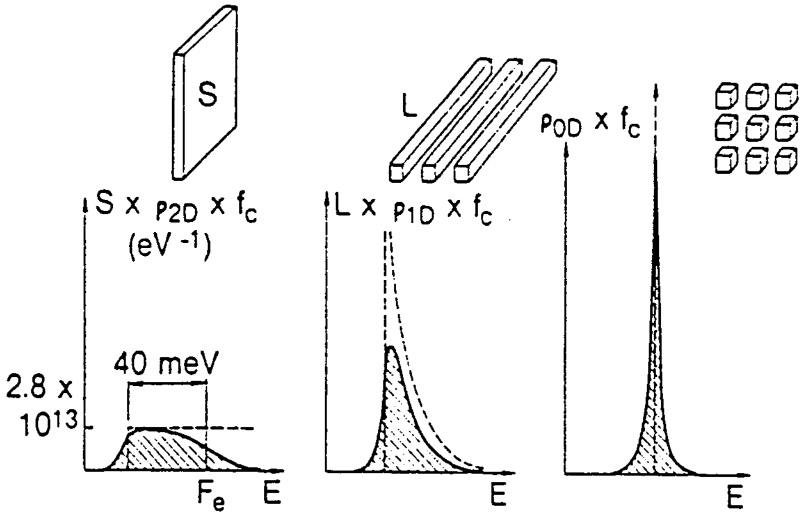


Fig. 1.8. *Schematics of gain curves in 2-D, 1-D, and 0-D structures. Similar numbers of electrons and holes are being injected above transparency, yielding equal integrated gain. [Weisbuch91]*

2. Fabrication of nanostructures

2.1. Introduction

As 2-D systems are so easily obtained in the growth direction by Molecular Beam Epitaxy (MBE) or Metal-Organic Chemical Vapour Deposition (MOCVD) techniques, an obvious way to attempt to obtain 1-D (quantum wires) or 0-D (quantum dots) systems is to pattern usual 2-D heterostructures using nanoscale lithographic techniques (section 2.2). However, great care must be taken to minimize the damage introduced by the etch procedure used in the pattern transfer onto the semiconductor. One effect of the etch damage is carrier depletion, which, before 1991 [Weisbuch91], caused quantum wires or dots below ≈100 nm geometric lateral size to be empty. Another probe of the damage is the often reported decrease in luminescence efficiency due to the numerous non-radiative states introduced during the fabrication process.

To obtain less-damaged structures, 'softer' fabrication techniques have been searched for (e.g. strain-induced potential wells, split-gate electrodes for voltage-controlled potential wells, etc. [references in Weisbuch91]). The geometric resolution of these techniques is lower than for direct etching and they allow only smooth potential variations, which are usually sufficient for low-temperature operation of electronic devices. Optical devices require substantially larger quantizing energies, since band filling or heating effects are important, owing to effective carrier temperatures quite higher than the lattice temperature, the more so at low temperatures. Therefore, abrupt variations of composition or potential energy with dimensions in the 10 nm range are necessary, and the present 'soft' techniques can not be used. Finally, a number of direct-fabrication methods are being investigated. In particular, the direct semiconductor growth methods seem to be very promising, see e.g. [Nagamune94]. In this chapter the fabrication of uniform arrays of high quality quantum dots from GaAs-$Al_{0.3}Ga_{0.7}As$ single quantum well material is described. A combination of electron-beam lithography and low-damage electron cyclotron resonance plasma etching is used.

2.2. Lithography

For the fabrication of micro- or nanostructures, first of all a mask or a writing beam that defines the pattern is needed. For example, visible to deep UV light is used for projection or contact printing. The minimum structure dimensions in this case are limited by the wavelength of the light. Direct writing of the pattern with a focussed electron beam requires no mask and can result in very high resolution structures, of the order of or smaller than 10 nm. Of course, a suitable recording medium, the resist, onto which the desired pattern is written, must be found. It must be sensitive to or altered by the writing beam in such a way that, after development, the portion exposed to the beam is removed (positive resist) or remains on the workpiece after the unexposed portion is removed (negative resist). Finally, a process is required to transfer the pattern into the semiconductor. The resist pattern is used as a mask through which material can be added (e.g. metallization or ion implantation) or removed (e.g. etching).

The combination of writing tool and resist, with the associated processes, is referred to as 'lithography', a term derived from Greek, meaning 'writing on stone', i.e. engraving, printing, etc.

2.2.1. SEM lithography

As pointed out before, electron beam lithography is a very high resolution method for direct pattern generation. The pattern is written with a highly focussed electron beam, which is controlled (deflected and turned on and off) by a computer. In our case, a modified scanning electron microscope (SEM) is used. Although the mechanical and electrical stabilities are not ideal, it is well suited for this type of work from the point of view of beam size.

2.2.2. Principle of the SEM

Electrons from a cathode gun source are accelerated by a voltage (~10 kV) between cathode and anode. In general, two or more electromagnetic lenses focus the resulting beam onto the surface of the substrate by demagnifying the electron gun source. High flexibility can be achieved, since the size and energy of the final beam can be readily varied by changing the focal length of the electron lenses and the accelerating potential of the electron gun. A deflection coil system in front of the last, probe-forming, lens scans the electron probe across the specimen. See Fig. 2.1.

The electrons hitting the sample generate secondary and other electrons. These can be collected easily by means of a positively biased collector grid thanks to their low exit energy (~few eV). Behind the collector grid, the secondary electrons are accelerated and focussed onto a scintillator and the light quanta generated are recorded by a photomultiplier. An image is formed by scanning the electron probe in a raster in synchronism with the electron beam of a separate VDU and modulating the VDU intensity by the photomultiplier signal.

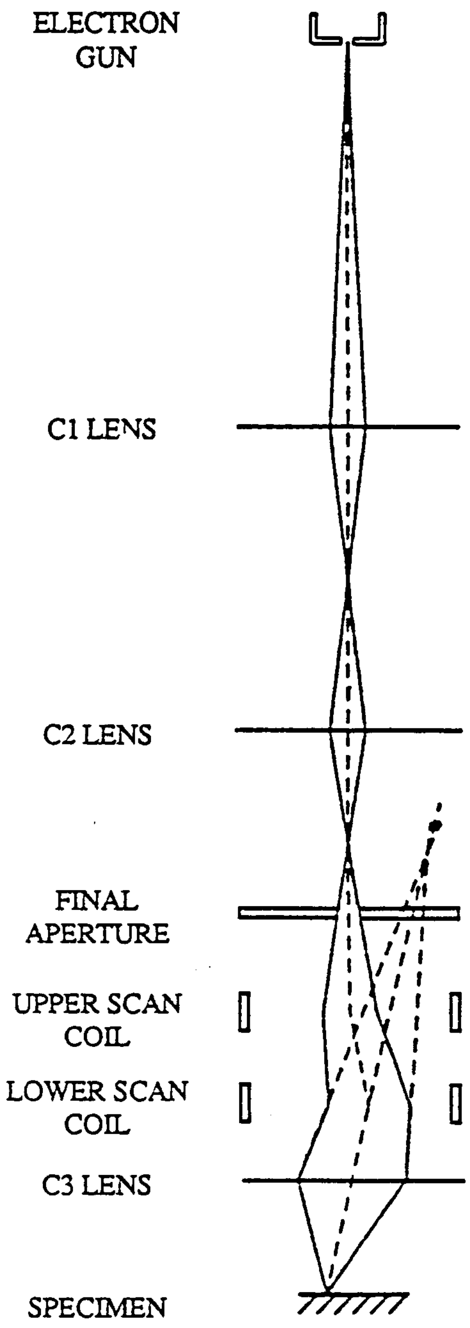


Fig. 2.1. *Schematics of the electron optics in a scanning electron microscope.*

The magnification can be increased simply by decreasing the scan-coil current and keeping the image size constant. The optical properties, e.g. the probe diameter of the Gaussian beam system in the nanometre regime are determined by a combination of lens and deflection aberrations, diffraction, and source brightness. Two different types of electron sources are often used: thermionic and field-emission guns. See Fig. 2.2.

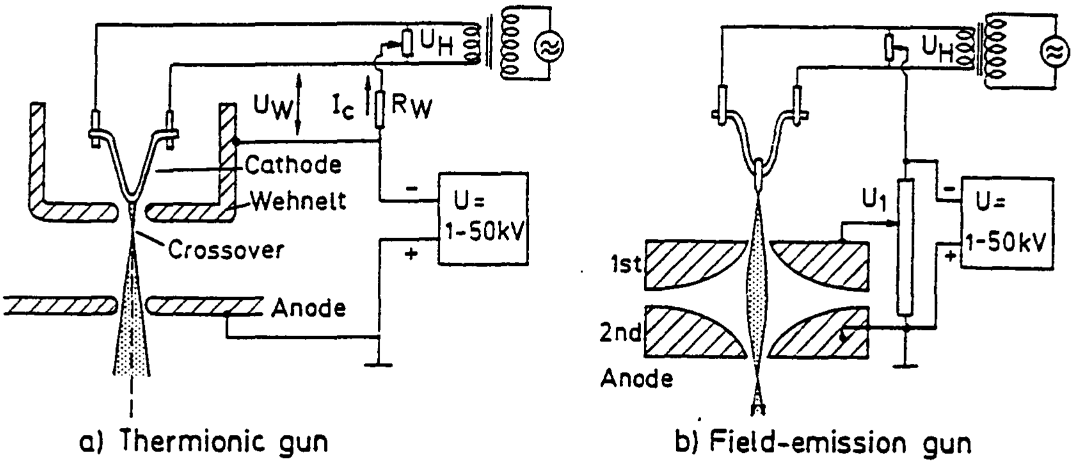


Fig. 2.2. *(a) Thermionic electron gun consisting of cathode, Wehnelt cup and anode. (b) Field-emission gun, U_I is the extracting voltage. [Reimer85]*

Thermionic electron gun

Electrons can be ejected from the cathode into the vacuum by thermionic emission. Electrons from the Fermi-level of the cathode material then overcome the work function Φ by thermionic excitation.

Here, a lanthanum hexaboride (LaB_6) rod with a polished tip is used to write the quantum dot patterns in the resist. LaB_6 (Φ=2.7 eV) is a good cathode material that can operate at 1400-2000 K (cathode tip temperature) with an emission current density j_c of the order of 20-50 Acm^{-2} (for comparison: tungsten has a higher work function, Φ=4.5 eV, resulting in working temperatures of 2500-3000 K, j_c of the order of 3 Acm^{-2} and a life-time of 2-40 hours). LaB_6 cathodes attain life-times of 200 hours or more, but need a vacuum better than 10^{-4} Pa (tungsten 1-5 $\cdot 10^{-3}$ Pa) to prevent the formation of volatile oxidation products. This can be achieved by evacuating the space in front of the cathode of the gun with an additional ion pump.

The rod is heated indirectly, by supporting it between wires of carbon fibres and applying a heating power P_c. The relation between the resulting beam current I_c and P_c is plotted in Fig. 2.3. Optimum operation, i.e. the smallest beam crossover and (therefore) probe diameter, occurs at the knee of the I_c versus P_c diagram, just before saturation sets in.

The emission current of a thermionic cathode depends very sensitively on the temperature of the cathode tip, and therefore on the supplied heating power. Long-term drift of the cathode occurs due to differences in the thermal expansion, resulting in a shortening of the distance

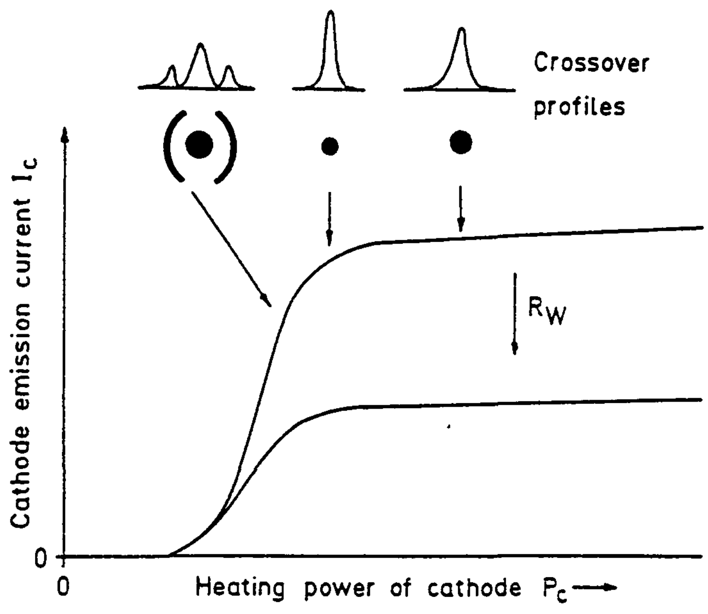


Fig. 2.3. *Cathode emission current I_c versus cathode heating power P_c for a self-biased thermionic electron gun and typical profiles of the crossover (see Fig. 2.2 (a)) at three working points. [Reimer85]*

between cathode tip and Wehnelt cup (Fig. 2.2 (a)), or due to changes in the work function. Even for a constant emission current I_c, the probe current I_p can be changed by deterioration of the gun alignment caused by small shifts of the tip position.

Field emission gun

Field emission from a tungsten tip of radius $r\approx 0.1\,\mu m$ starts when the extracting voltage U_l increases to values larger than ~2 kV ($|E|\approx U_l/r > 10^7$ Vcm^{-1}): electrons from the Fermi level E_F can penetrate the potential barrier by the quantum-mechanical tunnelling effect. Field-effect cathodes need an ultra-high vacuum better than 10^{-6} Pa otherwise the tip radius is destroyed by ion bombardment from the residual gas. A field-emission gun needs two anodes (see Fig. 2.2 (b)). The first regulates the field strength at the tip and hence the emission current. The second anode accelerates the electrons to their final kinetic energy.

Field-emission cathodes can function at room temperature but often work at $T_c\approx 1000$ K to avoid gas adsorption at the tip. This intermediate cathode temperature results in a lower energy spread of the order of 0.2-0.4 eV.

The main advantages of field-emission guns are the very substantially higher gun brightness, and the small diameter of the crossover (see Fig. 2.2 (a)) of the order of 10 nm, so that only one demagnifying lens is needed to get electron-probe diameters below 1 nm. The disadvantages of these guns are the short-term fluctuations (noise), of the order of 2-5%, and a long-term drift of the emission current, both due to desorption and adsorption of gas molecules (see Fig. 2.4). When the vacuum is sufficient, the life-time of a field-emission gun is limited only by damage to the tip caused by an electric discharge.

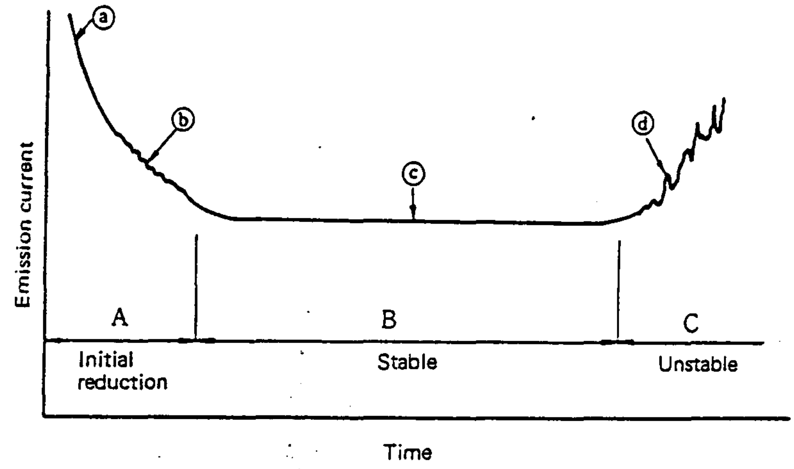


Fig. 2.4. *Emission current as a function of time, after 'flashing'.*

The noise and instability are the reason why this source is rarely used for lithography. For imaging purposes, however, these disadvantages are easily overcome and the higher brightness and smaller electron-probe diameter lead to a resolution that is much better than for a thermionic source (as small as 15 Å compared to ~80 Å).

2.2.3. Electron beam resist exposure

In the electron beam writing, highly focussed electrons hit the surface of the sample and make elastic and inelastic collisions with atoms. This is used to expose the resist, i.e. change the molecular sizes in the exposed areas, which can be translated into differences in solubility in an appropriate developer. In exposed areas, a positive resist becomes soluble because the principal effect of irradiation is breaking of the macromolecular chains whereas a negative resist becomes insoluble because of the dominance of cross-linking processes. The necessary exposures are in the range 10^{-8}- 10^{-4} Ccm^{-2}. The resist used here for the fabrication of quantum dots is a chemically amplified negative resist, with a sensitivity of <2 μCcm^{-2} (Appendix A for details).

Several factors other than the size of the electron beam determine the extent of the exposed volume (and thereby the minimum linewidth) in a layer of resist. There are several scattering processes in the material that lead to the deposition of energy at distances remote from the initial impact point of the electron. The most important interaction processes and their information volumes are shown in Fig. 2.5.

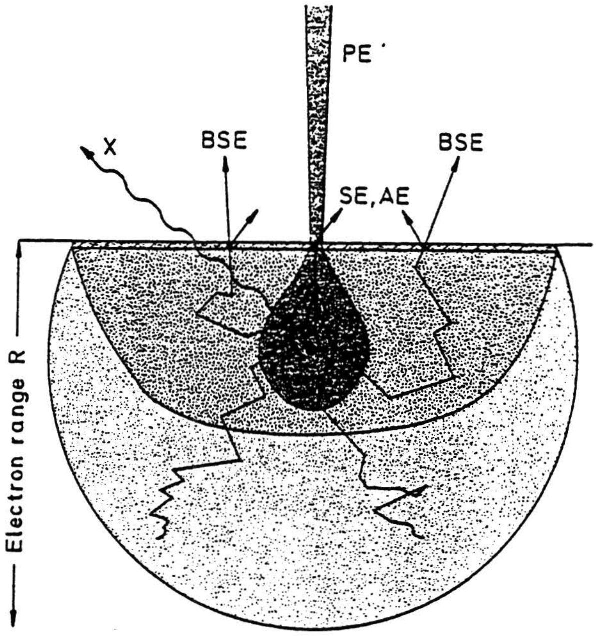


Fig. 2.5. *Origin and information depth of secondary electrons (SE), backscattered electrons (BSE), Auger electrons (AE) and x-ray quanta (X) in the diffusion cloud of electron range R for normal incidence of the primary electrons (PE). [Reimer85]*

For integrated-circuit lithography, where at present relatively low accelerating voltages and thick resists are used, electron scattering is the most important factor. For nanolithography, where higher accelerating voltages and thinner resists are used, secondary electron generation is the dominating factor, and the effects of scattering are less important. According to [Broers88] it is the delocalization of secondary-electron generation, together with the subsequent straggling of the secondary electrons in the resist, that sets the resolution limit for nanolithography. For patterns such as arrays of dots, however, the smallest feature size can be reduced below this limit (see chapter 4).

Electron scattering

Fig. 2.6 shows qualitatively the mechanisms of electron scattering. The resist is exposed by both the incident electrons and the electrons that are scattered back, e.g. from the substrate. Lateral scattering of the primary electrons as they penetrate the resist gives rise to the narrower of the distributions, the 'forward-scattering'. Backward scattering of electrons gives rise to the broader, 'backscattering' distribution.

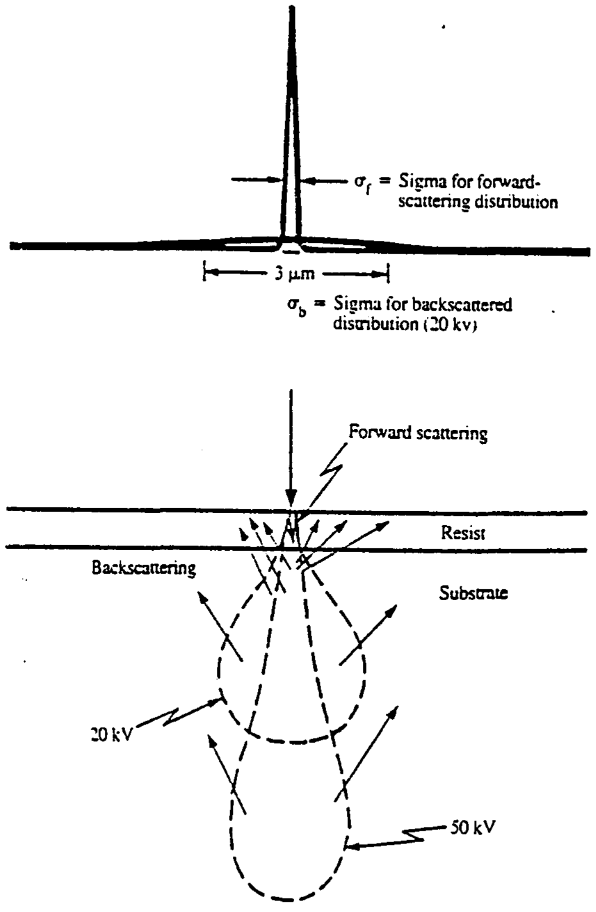


Fig. 2.6. *Electron scattering in electron resist exposure. The curves at the top of the figure show the exposure distributions due to the incident and backscattered electrons. [Broers88]*

Fig. 2.7 shows approximate distributions for the electron-beam exposure of a 1-μm thick resist layer on a silicon substrate at incident electron energies of 10, 25, and 50 kV. It indicates that the width of the forward-scattering distribution is reduced as the electron energy increases. Obviously, it is also reduced for thinner resist layers. The rates of reduction in each case are such that forward scattering becomes negligible for the combination of 50 kV and resist thicknesses below about 0.1 μm [Broers88].

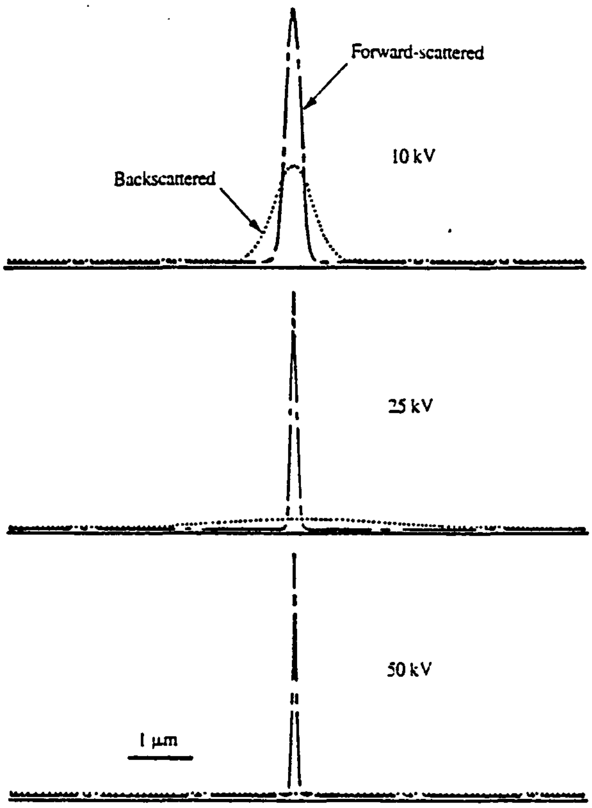


Fig. 2.7. *Exposure distributions for a 1-μm thick resist layer on a silicon substrate for 10, 25, and 50 kV. [Broers88]*

The width and intensity of the backscattered distribution also depend upon the incident electron energy. The greater the incident energy, the larger the disc from which electrons are generated.

At an electron energy of 40 keV the backscattering diameter is so large (> 3 μm, see Fig. 2.8) that the backscattered electrons only produce a background 'fog'. For small isolated features (and low exposure doses, see chapter 4), backscattering has a negligible effect on contrast or resolution. For nanolithography, the only significant effects of backscattering are that it reduces contrast for very dense patterns, and that it causes exposure variations over larger field areas (see Fig. 4.4).

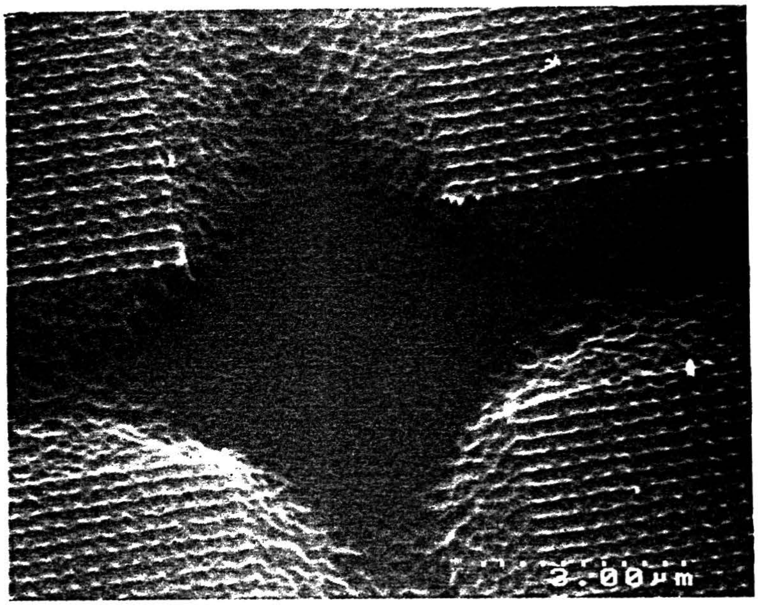


Fig. 2.8. *SEM photograph of some over-exposed 10μm-fields, each consisting of a 51x51 array of 150 nm spaced dots. The range of the backscattered electrons is about 3μm (at 40 kV).*

Secondary electrons

Electrons with energies above about 5 eV can break or make chemical bonds and thereby expose resist. This means that the free, low-energy, secondary electrons produced by inelastic interactions between high-energy electrons (either scattered or primary) and the resist atoms are important in resist exposure. These low-energy secondaries, which typically have energies up to about 20 eV, can be excited remotely from the high-energy incident electron beam and may struggle further into the resist before their energy is dissipated. As already mentioned, it is thought that these secondary electrons set the resist resolution limit. This is a fundamental problem with electron-beam exposure that can be reduced only by using resists with shorter ranges for these low energy (up to 50 eV) electrons, or which are exposed only by higher energy primary electrons. Note that a similar process also limits the resolution of the best secondary electron SEMs!

2.2.4. (Quantum dot) pattern generation

Unlike light and x-ray exposure, electron-beam lithography does not need a rigid mask, because the pattern can be digitally stored: for writing a pattern, the beam deflection coils are, instead of being raster-scanned to form an image (section 2.2.2), controlled by a computer that has been programmed to produce the desired pattern. Each time the beam has been moved (deflected), the exposure is delayed by a few milliseconds (so-called 'settling' time) to get rid of oscillations. The computer also controls a beam blanker (see Fig. 2.9) that switches the beam on and off. Instead of turning the gun on and off, the beam is deflected across a diaphragm by applying a transverse electrostatic field of e.g. a parallel plate capacitor. A magnetic field beam blanker is also possible and simpler to design, but has much slower response times [Reimer85].

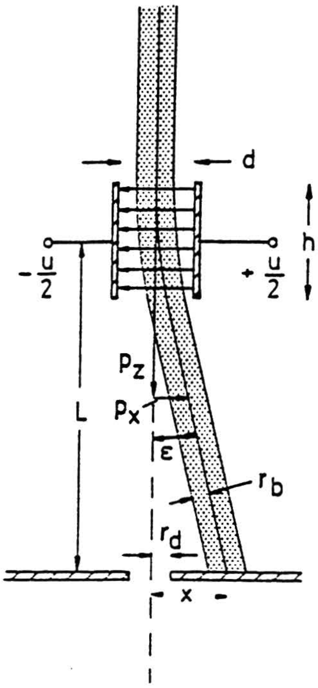


Fig. 2.9. *Schematics of a parallel plate electrostatic beam blanker.*

In this way, one field can be exposed with the pattern. The size of the field is determined by the magnification set on the SEM. For high resolution patterns relatively small writing fields are necessary: in the SEM used here the field size is 20 μm. In purpose-built lithography systems possible field sizes are much larger. Therefore, when larger total areas are required, the stage (with sampleholder) itself must be moved mechanically. For chip-production or other high-precision applications, where the fields must be positioned with the same accuracy as the pattern resolution, this requires a complex (and therefore expensive) stage control. A simple, stepper motor driven stage, however, is sufficient for experimental purposes. In the case of quantum dot patterns, a total area of approximately 300 x 300 μm is exposed. It consists of an array (15x15) of 20 μm (identical) fields, produced by 'stepping' (moving the stage) and 'repeating' the exposure. Each field consists of an array of 51x51 single pixels (total: 2601), with a spacing of ~300 nm (for a 20 μm field). A pulse generator, triggered by the computer's beam blanking signal, is used to control the exposure (or 'dwell') time for each pixel (dot). See Fig. 2.10.

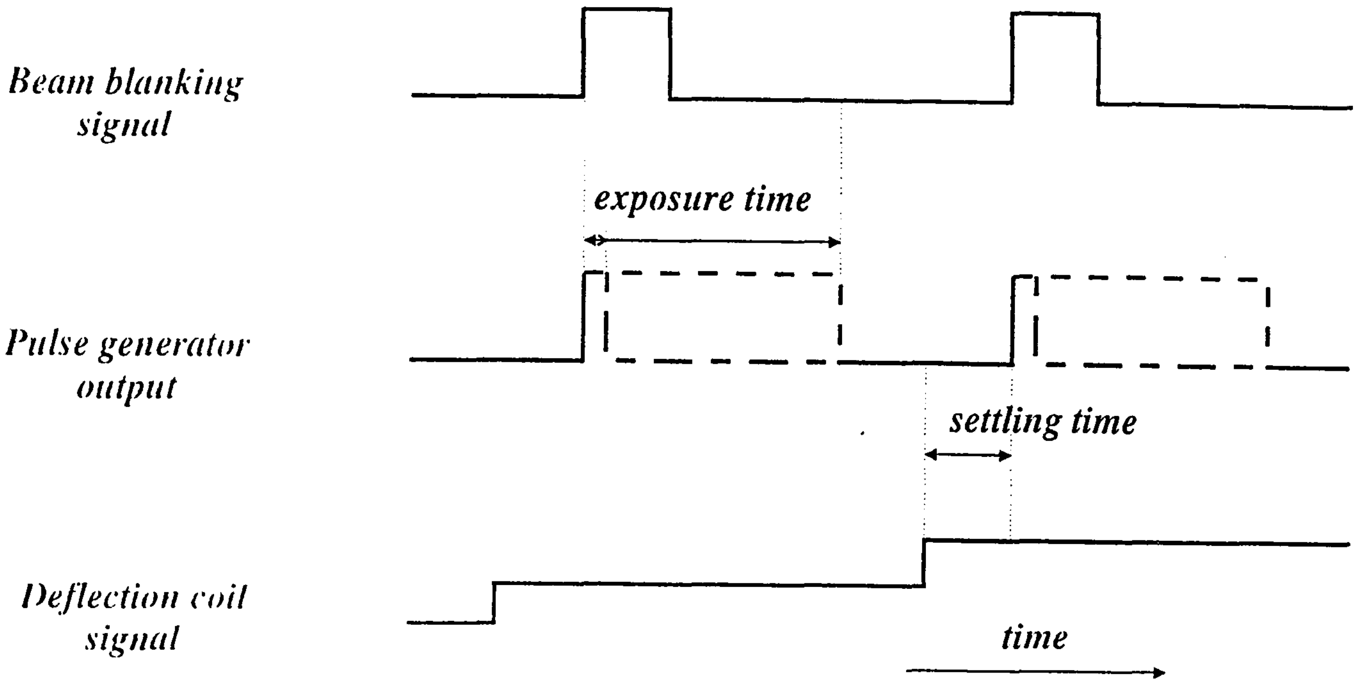


Fig. 2.10. *Illustration of the beam blanking and deflection signals.*

2.3. ECR plasma etching

To transfer the resist pattern, obtained by lithography (section 2.2), into the sample, a special etch technique, electron cyclotron resonance (ECR) plasma etching, is used. To illustrate its advantages over other etch techniques, an overview of (the principles of) plasma etching is given below.

2.3.1. Plasma etching

The plasma state encompasses a wide range of electron energies and densities and includes such phenomena as flames, low-pressure arcs, solar coronas, and thermonuclear reactions. The regime of interest to semiconductor processing is the low-pressure plasma or glow discharge. These plasmas are characterized by gas pressures on the order of 0.1 to several torr, free-electron densities of 10^9 to 10^{12} cm^{-3}, and electron energies of 1 to 10 eV. The light glow emitted from the plasma is characteristic of the electronic transitions taking place.

The plasma is able to generate chemically reactive species at relatively low temperature because of the nonequilibrium nature of the plasma state. The temperature of the chemical species (atoms, molecules, or radicals), is generally near ambient. The electrons, however, can exhibit temperatures of tens of thousands of Kelvin. Their energy is sufficient to break molecular bonds and create chemically active species in the plasma. Any of these species can be excited to higher electronic energy states by further interaction with the electrons. Hence, chemical reactions that usually occur only at high temperature can be made to occur at low or even ambient temperature in the presence of an activating plasma state. Although the ionization rate is small (typically less than 10^{-5}, but about 10^{-2} in an ECR plasma source), it is adequate to provide sufficient numbers of reactive species. It also allows most of the plasma to remain near ambient temperature.

The plasma state can create highly reactive species of oxygen, fluorine, or chlorine that will readily attack many materials: after absorption on the material surface and (kinetically assisted) chemical reactions, the volatile reaction products are pumped away. This is the process of plasma etching, which exists today in a variety of forms.

There are several reasons for the impetus it has achieved as a semiconductor manufacturing technique: apart from greater (and safer) control at reduced cost, compared with the more conventional wet chemical processes, it offers the possibility to produce fine resolution, achieved by 'anisotropic' etching. Wet chemical etching proceeds at the same rate in all directions (=isotropic), resulting in the characteristic linear-circular profile (see Fig. 2.11 (a)). Plasma etching may be completely chemical, and therefore tends to be isotropic. Or the chemical reactions on the substrate may be driven or enhanced by the kinetic energy of the incoming ions. This situation is referred to as 'kinetically assisted chemical reaction'. This type of plasma etching can be highly directional. Note that by definition, anything that is not isotropic is anisotropic. However, in the plasma etching literature, 'anisotropic' is generally used to refer to the very directional process that can result in vertical wall profiles (Fig. 2.11 (b)). But, etch conditions can also be adjusted to yield smooth sloped edge profiles (e.g. needed for metal crossovers). Gas usage in plasma etching is generally less than fluid use in wet etching, purity can be higher, and contamination less..

A disadvantage is the complex and sensitive dependence of process results on process parameters, which means that plasma processes are developed more by empirical means than by theoretical analysis. Another is the possible damage to surface layers of the semiconductor by ionic bombardment.

2.3.2. Plasma reactors

Plasma processing occurs in equipment called a plasma reactor. These reactors exist in three major configurations illustrated in Fig. 2.12: barrel (or tube) reactors, planar reactors, and

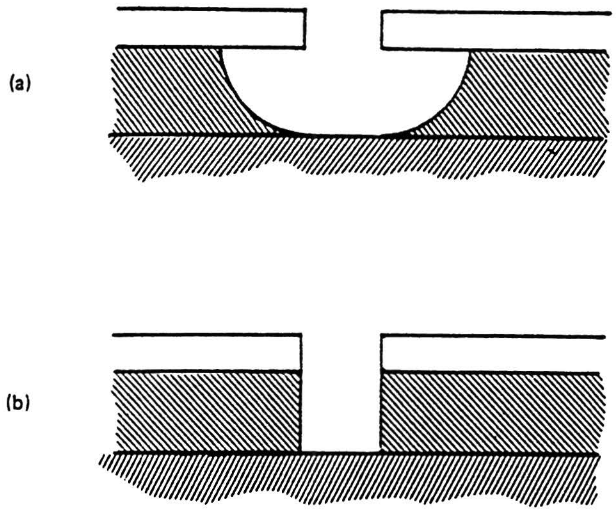


Fig. 2.11. *Plasma etching profiles: (a) isotropic and (b) directional ('anisotropic').*

downstream reactors. The first two represent the traditional approach to plasma processing. The downstream reactor is a more recent innovation and has significant advantages, as will be described below.

In barrel systems (a), the plasma is excited by a high frequency rf source, usually 13.56 MHz, using inductive coils or capacitive electrodes outside of the quartz or glass tube. There is nothing special about this number as far as the discharge is concerned. It just happens to be a frequency allotted by international communications authorities at which one can radiate a certain amount of energy without interfering with communications. The substrates are generally held in the vertical position by a wafer holder, are immersed in the plasma, and no electrical bias is applied between the wafer and the plasma. The wafer surfaces are subjected to only low-energy ion bombardment, probably less than about 30 eV. High uniformity of growth or etch rate across a wafer is difficult, if not impossible, to obtain using tube reactors. This follows from the nonuniformity in the plasma, the gas flow pattern, and the wafer temperature.

Unlike the barrel reactor, the planar reactor configuration (b) is capable of high uniformity. Wafers are generally placed flat on the lower electrode, which often is also used to heat the wafers. The rf signal which excites the plasma is applied across the two electrodes. The particular planar configuration shown in Fig. 2.12 (b) is a 'radial flow' reactor. The reactant gases are introduced either at the outer radius or on-axis, and flow radially between the electrodes.

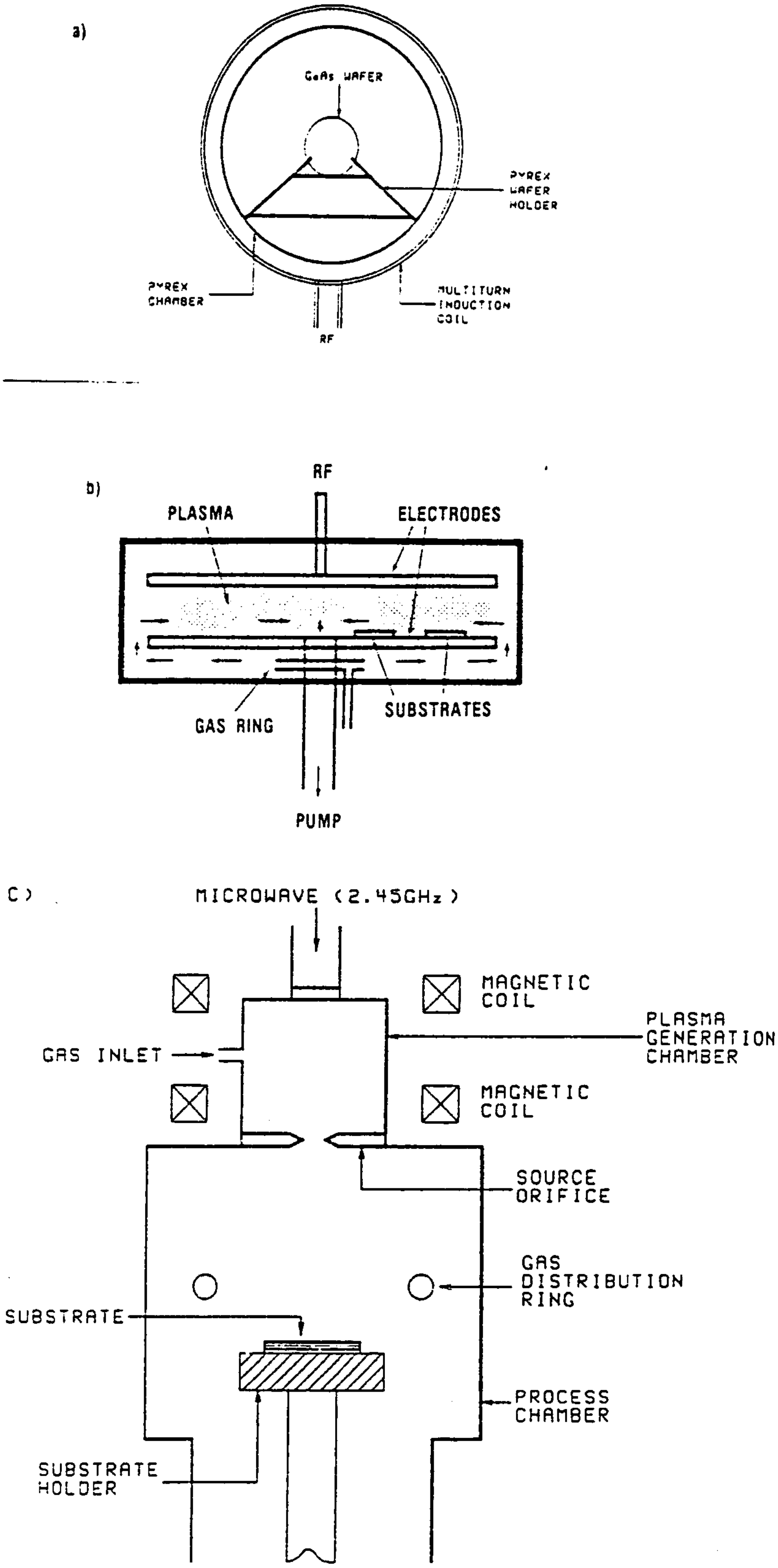


Fig. 2.12. *Three types of plasma reactors: (a) barrel; (b) planar or parallel plate; (c) downstream. [Williams90]*

In the traditional barrel and planar reactors discussed above, the wafer is situated in the plasma generation chamber and is therefore immersed within the plasma. In downstream reactors (c), the wafer is separated from the plasma chamber, so it is 'downstream' from the region of plasma generation.

There are several advantages. First, species with high kinetic energy (intrinsic to the plasma region) remain largely within the plasma chamber and do not impinge upon the wafer. This prevents (or greatly reduces) the electronic damage common to traditional planar reactors. Also, many of the active species in the discharge region decay sufficiently fast that only a few primary active species reach the wafer surface. This considerably simplifies the plasma chemistry.

Another major advantage is that secondary reactant gases can be introduced into the wafer chamber, separate from the plasma chamber, as indicated in Fig. 2.12 (c). The resulting 'downstream' chemical reactions can produce highly active free radical species not easily produced by using traditional methods.

Although downstream reactors can be constructed using traditional rf discharge plasma sources, most use microwave plasma sources operating at 2.45 GHz. Microwave discharges are electrodeless: a strong electromagnetic field is established within a resonant cavity and results in gas breakdown. Matching of the high power microwave signal to the gas load is typically accomplished by one of two methods. The first method is rather direct: both the gas and the microwave signal are simply introduced into a mechanically tuned microwave resonant cavity. Resonant-cavity-excited microwave discharges (electrodeless) couple energy directly to the gas without the high plasma sheath potentials typical of traditional rf discharge reactors using electrodes. The 13.56 MHz discharges have potentials on the order of several hundred volts, while microwave discharges have potentials on the order of ten volts, resulting in higher concentrations of atomic and radical species. This is the principal advantage of the microwave plasma source. The direct, resonant cavity sources work best at gas pressures between approximately 0.1 and 1 torr.

The second method of coupling the microwave signal to the gas in the plasma chamber uses the principle of 'electron cyclotron resonance' (ECR). The ECR concept uses a microwave energy source in a strong magnetic field to ionize gas molecules by cyclotron resonance of the outer shell electrons. The magnetic field (0.875 Tesla for 2.45 GHz) is usually applied by solenoidal coils placed around the (microwave) source cavity, as illustrated in Fig. 2.12 (c). A magnet below the process chamber is used to shape the plasma distribution in the chamber and thereby ensure that the ions arrive perpendicular to the sample.

The electron resonance and resulting secondary collisions are very efficient at producing ionization. Ionization densities are usually several orders of magnitude greater than those achieved in conventional rf discharges (ionization ratios in the range of 10^{-2} instead of 10^{-5}). This is a major advantage of the ECR downstream reactor. The greater ionization density results in very high reaction rates at lower temperature, even ambient temperature. For etch applications an rf signal (13.56 MHz) can also be applied to accelerate the ions towards the sample by coupling the dc self-bias (see [Chapman80] for an excellent discussion) to the sample holder ('chuck').

Electrical damage to the wafer can be minimized or eliminated by keeping the ion energies low. The ECR effect works best at gas pressures below 0.01 torr. The operating conditions of ECR reactors also allow (very) low pressure operation in the wafer chamber, even below 10^{-4} torr, which is below the pressure regime used in most other plasma reactors. This low pressure has several potential advantages. In etch systems, it can result in more accurate etching (replicating the mask pattern) of small (submicron) patterns and in reduced etch rate differences between large and small patterns on the same wafer ('loading effect').

From the significant advantages over traditional plasma reactors mentioned above, it is obvious that a high vacuum ECR system, such as used here, is the best choice for uniform and low-damage etching .

3. Characterization

3.1. Introduction

Fabricating nanostructures is a delicate process. To be able to optimize the different conditions, it is necessary to physically assess the processed material. The quantum well material must also be optically characterized before and after processing into quantum dots (or wires) to make comparisons of the obtained data valid.

3.2. Physical assessment

After electron beam exposure and development of the resist the sample is etched in an ECR system. The combination of the conditions used in these steps determines the size, height, and uniformity of the structures, as well as the etch-profile (also surface damage and contamination). These parameters provide important feedback information for the fabrication process. A high resolution (up to 15 Å under optimum conditions) SEM with a field-emission electron gun is used to image the structures.

The images can be recorded with a photographic unit on either polaroid or roll film. The standard photo conditions that are used for the dot-samples are an accelerating voltage of 30 kV, 45° sample tilt, and a film speed of 400 ASA. The magnification scale is indicated on each photo. See for instance Fig. 4.8 - 4.10. The test-samples (bulk GaAs) are usually sputter coated with a thin layer of gold-palladium, to prevent charging-up effects and to improve contrast.

3.3. Optical spectroscopy

As discussed above, it is common practice to examine the physical appearance of nanostructures using a scanning electron microscope (SEM). However, such a technique is not suitable for studying electrical or optical characteristics.

3.3.1. Introduction

Optical spectroscopy of semiconductors has been a lively subject for many decades. It has helped to shape the concepts of energy gaps, impurity states and resonant states among others. Of course spectroscopic techniques had been in use for some considerable time before semiconductors were developed, in the study of atomic spectra for example.

For the optical characterization of the quantum well material and the quantum dots produced from it, two spectroscopic techniques are used here: low-temperature (5 K) photoluminescence (PL) and photoluminescence excitation (PLE) spectroscopy. Both are described below.

The etched structures (quantum dots, QD) are compared to the unprocessed reference material (QW) in PL efficiency, linewidth and energy, but factors such as wafer nonuniformity can make evaluation of apparent energy shifts difficult. Therefore, the QW material is extensively mapped (3x3 array on 1x2 cm wafer) by both PL and PLE before processing, so that the characterization afterwards can be compared to the reference spectra taken within 2 mm of the area processed into dots.

3.3.2. Experimental set-up

For both PL and PLE a focussed laserbeam of certain wavelength (λ_{exc}) hits the sample. Light from the sample is collected and imaged on the entrance slit of a spectrometer, which detects the intensity as a function of wavelength (λ_{det}). By scanning either λ_{exc} or λ_{det} a spectrum is obtained, PLE and PL, respectively. In Fig. 3.1 the schematic experimental set-up is shown, which is discussed below.

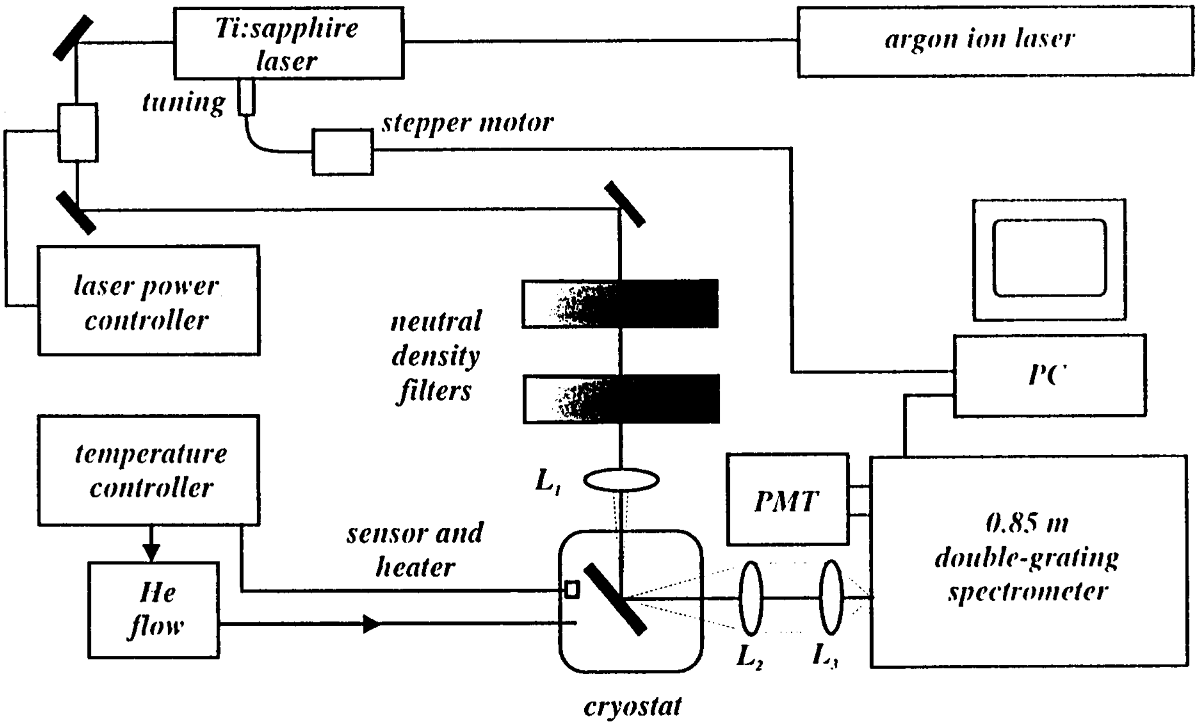


Fig. 3.1. *Schematic experimental set-up for PL and PLE spectroscopy.*

An argon ion laser with 8-9 W output power is used to pump a titanium:sapphire laser. This laser is tunable in the range from 740 to 790 nm, i.e. directly into the quantum well (see below). A laser power controller stabilizes the output power, which is typically in the range 20-150 mW, and two neutral density filters attenuate the excitation power to 150 μW.

The laser beam then is focussed onto the sample in a spot of ~50 μm FWHM by lens L_1. The sample is mounted in a helium-flow cryostat (temperature set to 5.0 K to avoid noise due to helium boiling) under nearly 45°: the specularly reflected laserlight is moved off the spectrometer entrance slit by slightly rotating the sampleholder. The luminescence light from the sample is collected by lens L_2 and imaged on the slit by L_3. A computer controls the two adjustable gratings (1800 grooves/mm) that select the detection wavelength; calibration is performed using the 546.073 nm line from a mercury lamp. The spectrometer resolution near this line is 0.005 nm, and the accuracy (over 500 nm) is 0.1 nm. Then the titanium:sapphire laser is calibrated at 740 nm using the spectrometer. All PLE spectra are corrected for the slight, about 0.5 nm, nonlinearity in the scanning of the laser (see Appendix B). Light of the selected detection wavelength is collected by a cooled GaAs photomultiplier (PMT) and the intensity measured by photon-counting electronics.

Alignment of the laserbeam on the dot area is accomplished by a combination of coarse positioning (i.e. moving the cryostat on a micrometer stage) optically under white-light illumination, followed by fine positioning until diffraction of the laser light by the dot array can be seen. Then the sampleholder is slightly and carefully rotated to move the diffraction pattern off the slit, while maintaining the alignment. Once a PL peak is observed, the alignment can be optimized: maximize the PL peak intensity by adjusting the lens positions, and possibly by slightly moving the cryostat.

3.3.3. Photoluminescence (PL)

Here, a PL spectrum is taken by scanning λ_{det} from typically 755 to 785 nm using standard conditions: λ_{exc} is fixed at 740 nm (in the quantum well, see below) and 150 μW and the spectrometer slitwidths are set to 0.4 mm, at a sample temperature of 5.0 K.

As shown in [Bestwick95] there is a decrease in the relative efficiency of in-barrier compared to in-well excitation for dots compared to the unpatterned reference. This is explicable in terms of the high density of AlGaAs surface states created in the dots, which offer a competitive recombination pathway for carriers photogenerated in the barrier. The implication of this is that PL efficiency measurements seeking to identify and elucidate intrinsic 0-D effects, i.e. those due to quantum confinement in the active region, should be performed with photoexcitation directly into active region energy states.

The PL spectrum of undoped GaAs/AlGaAs quantum wells (and dots, see Fig. 3.3) at low temperatures (5 K) consists of a single narrow line, which is brighter than in thick 3-D layers. A number of factors, occurring simultaneously or not, tend to this single recombination process and large quantum efficiency [Weisbuch91]:

- Carrier collection in QW at low temperatures is extremely efficient: carriers created in the overlayer barrier material are largely captured by the quantum wells.
- The 2-D exciton enhancement leads to efficient exciton formation due to an increased overlap of the electron and hole wavefunctions.
- Exciton luminescence is, to first order, a forbidden process, as the *k* selection rule only allows excitons with exactly the photon *k* vector to radiate. In quantum wells, excitons cannot propagate along the *z* axis as they are localized in the well. However, luminescence should be very efficient as the *k* conservation rule should be lifted thanks to the scattering by confining energy fluctuations. The same is true for quantum dots that provide a 3-D confinement for the excitons.

As opposed to PLE, the PL line shape does not represent directly the density of states (DOS) of the exciton band (see for example the discussion on the Stokes shift in section 5.3). The linewidth, however, can be used as an indication of the quality of the QW interface: spatial fluctuations of the quantized energy levels lead to broadening of the linewidth, particularly when the well width is reduced. See Fig. 3.2.

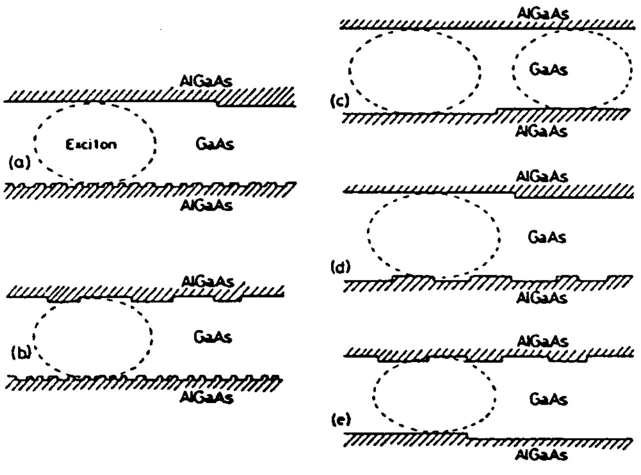


Fig. 3.2. *Atomic models of interfaces of various quantum well structures. When the lateral size of roughness is close to the exciton diameter, the PL luminescence broadens most severely. [H. Sakaki in Davies91]*

The strength of optical transitions is usually described by the 'oscillator strength', which in a bulk direct gap material is dimensionless and of the order unity. The concept of oscillator strength conveys the probability that the transition can occur, and is proportional to the number of k-states coupled to a given energy range. As discussed in [Weisbuch91], the oscillator strength per atom in 2-D does not increase but the new form of the DOS arising from the 2-D confinement results in a shared k_z for electrons and holes in the same subband. Having the same wave-vector k_z helps to concentrate the oscillator strength, when compared to 3-D. Going to 0-D, for the case where dimensions of the dot are somewhat larger than the Bohr radius of the exciton, a giant oscillator strength develops [C.M. Sotomayor-Torrès in Davies91]. However, [Weisbuch91] caution to consider the available data on weakly emitting dots and wires. They suggest that emission becomes weaker by orders of magnitude as the dimensions are reduced below 100-200 nm, based on a treatment of 'bottleneck' effects in energy and momentum relaxation [Benisty91]. Recent results in MOCVD grown dots and wires, see e.g. [Nagamune94], and also etched dots [Bestwick95], show no evidence of this effect. The results presented in chapter 5 support this: a significant increase in luminescence strength with decreasing dimensions, diameters from 60 to 25 nm, is observed.

Energy blue-shifts due to additional quantum confinement in the dots are also observed. PLE, however, is more convincing than PL, because it represents the dominant intrinsic absorption of the structures, and is not subject to factors associated with lateral carrier migration (see [Gustaffson93] and section 5.3 on Stokes shift).

3.3.4. Photoluminescence excitation (PLE)

A very convenient way to deduce absorption spectra without any sample preparation, in particular thinning and substrate removal, and good signal to noise ratio, is photoluminescence excitation spectroscopy: observing the photoluminescence at a given wavelength λ_{det}, usually in the low-energy tail of the PL peak, the excitation light wavelength λ_{exc} is scanned, with a tunable dye laser or a Ti:Sapphire laser, for instance. Peaks will appear in the spectrum as a result of increased absorption coefficient (see Fig. 1.6). In very narrow quantum wells (< 50 Å) only the n=1 heavy and light hole exciton peaks are observed, superimposed on the DOS.

The fit of the early absorption measurements led to the determination of the bandgap discontinuity ΔE between the conduction bands ΔE_c and valence bands ΔE_v. Calling $Q=\Delta E_c/\Delta E_v$, Dingle found that $Q=0.85\pm0.03$, assuming standard values for the <100> electron and hole masses, i.e. $m_e=0.067\ m_0$, $m_{HH}=0.45\ m_0$, $m_{LH}=0.08\ m_0$. In recent similar experiments on square and parabolic quantum wells, Miller was led to a reexamination of this partitioning

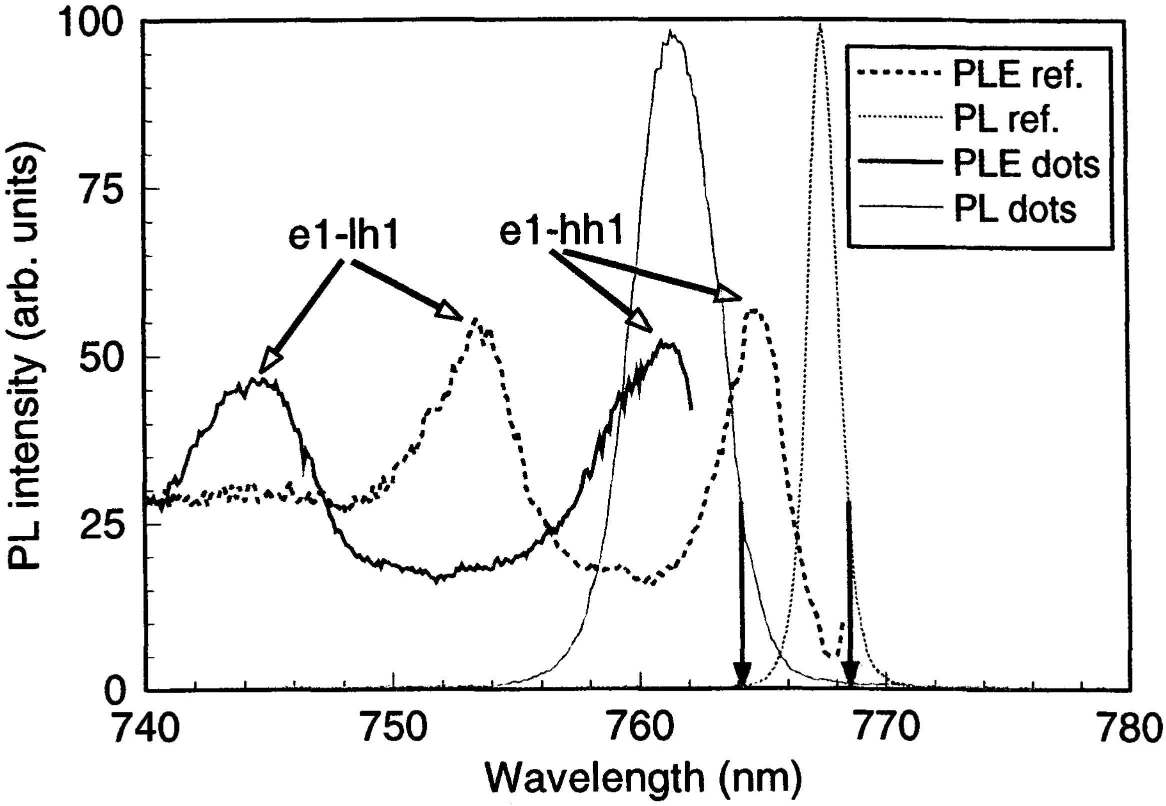


Fig. 3.3. *PL and PLE spectra for both pre-processed QW material and 25 nm diameter quantum dots. The arrows indicate the detection wavelength for the PLE spectra.*

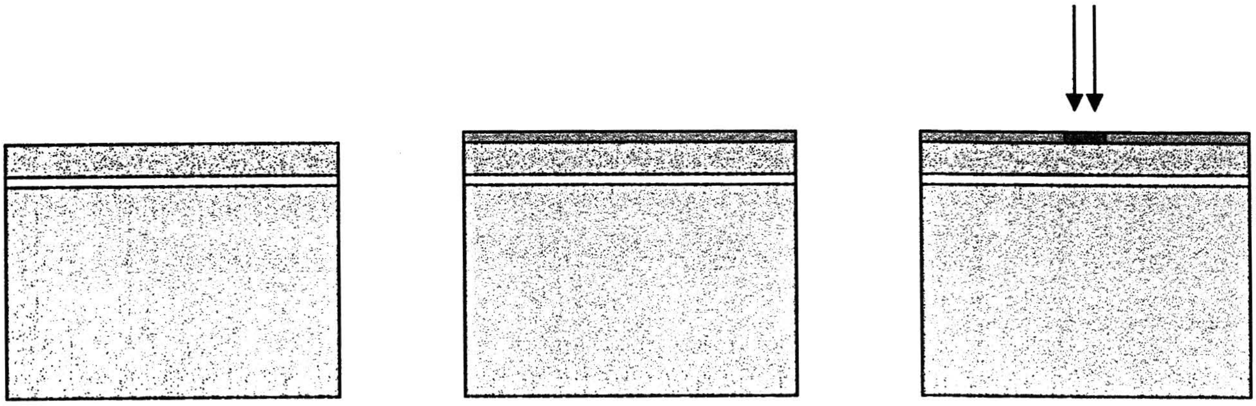
and evaluated Q=0.57, using a heavy hole mass m_{HH}=0.34 m_0 and m_{LH}=0.094 m_0. Including the exciton binding energy as a data fitting parameter, a value of 9 meV was extracted for thin wells, i.e. 50 to 100 Å, which is to be compared to the 4.2 meV value in the bulk [Weisbuch91].

The sharpness and intensity of the exciton luminescence is a measure of the sample quality: if the material is contaminated by defects, electrons and holes can be trapped in the potential wells associated with the presence of these defects. Broadening of the line width may also be a measure of the lack of uniformity of the interfaces: if the interface contains steps, the well width 'seen' by an exciton whose orbit radius exceeds the well boundary appears to be smaller for some excitons and larger for others.

The standard conditions used here for recording the PLE spectra are: Ti:Sapphire laser at 150 μW power, scanned from λ_{exc}=740 nm to about 0.5 nm below the detection wavelength λ_{det}, with spectrometer slitwidths of 0.4 mm and λ_{det} in the low-energy tail of the PL peak, at about a third of the peak intensity (see Fig. 3.3). All excitation spectra are corrected for small deviations from linearity in the scanning of the Ti:Sapphire laser (see Appendix B).

4. Reducing dot diameter

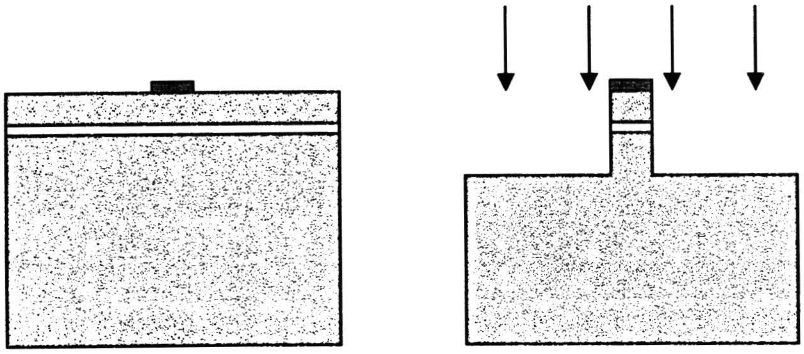
As already discussed in chapter 2, the process for the fabrication of quantum dots consists of three main steps: writing the dot array pattern in a suitable resist, developing it, and finally transferring the pattern into the semiconductor material. This is illustrated in Fig. 4.1. When trying to optimize the process, all of these steps have to be considered.



1. SQW

2. Resist application

3. Electron beam resist exposure



4. Resist development

5. ECR etching

Fig. 4.1. *Simplified fabrication process.*

4.1. Resist exposure

4.1.1. Resist

First of all, a high resolution, electron beam sensitive resist had to be found. Shipley's SAL 605, a novolak-based negative tone resist, proved to be a good choice. A thin layer of this material is applied to the GaAs/AlGaAs reference QW (Appendix A) after extensive PL and PLE mapping of the QW (section 3.3).

4.1.2. Beam conditions

The properties of the writing electron beam, i.e. beam probe diameter, intensity, energy, depend on several parameters. As shown in section 2.2, a high energy, low intensity, highly focussed electron beam is required for the smallest exposed volumes. Therefore, the highest accelerating potential and lowest probe current that are possible in the SEM, are used.

Focussing is performed at the highest magnification, within a distance of 0.5 mm of the area to be exposed (without exposing this area in doing so). This is made possible by making a thin scratch (~20 μm) on the sample from the middle to the edge. By tracing the edge of the sample (see Fig. 4.2) at a magnification of about 1000 times, the scratch can be found easily, and 'followed' down to the middle of the sample. Near the end of the scratch tiny particles (~200 nm) can be found, that serve as excellent objects for accurate focussing. The scratch also is a convenient pointer to find the dot array on the sample after processing.

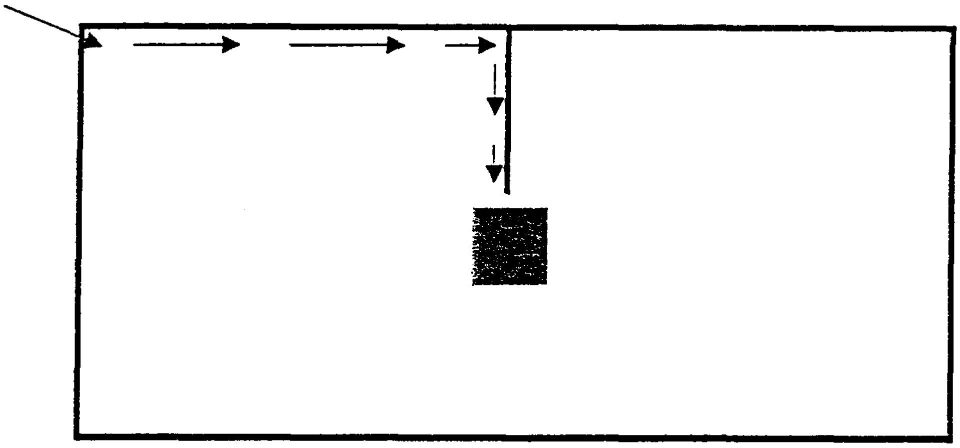


Fig. 4.2. *Sample of QW material, typically 7x4 mm in size, with scratch and dot area (not on same scale). Arrows indicate edge-tracing and following the scratch to the middle of the sample, where focussing is performed.*

The parameters that control the intensity of the electron probe are more difficult to control, as they slowly change in time (section 2.2). For example, the filament heating power slowly increases, typically in 1 hour, when the set filament current is not changed. This can drastically change the gun emission from under-saturation to over-saturation. Because the power increases quite slowly, this would not be a problem in itself. The filament holder, however, heats up much more slowly than the filament tip, which causes it to drift away from the optical axis in 5-10 minutes, and thereby significantly changes the electron probe current. Two gun alignment coils can correct for this drift, but not dynamically and not while in imaging, i.e. exposing, mode. Fortunately, the drifting stops after a certain period of warm-up time. Depending on the filament holder type, its age, etcetera, this time ranges from about 30 to 90 minutes. As mentioned above, the heating power will increase in this time, so the filament current must deliberately be set to a lower value, about 0.01 A below optimum value, to correct for this. In this way, optimum gun emission is reached after the stabilizing period, and exposure can start. Note that changing the filament current will restart the drifting. Since the focussing conditions do not change, focussing can be performed before or even during the stabilizing period.

When the sample is changed and the main chamber of the SEM pumped down again, to below $2 \cdot 10^{-6}$ torr, the above procedure has to be repeated. This means that only a few samples can be exposed in one day. It is therefore worthwhile for test purposes to expose several, up to about eight areas on the same sample, e.g. an exposure series. The total exposure time for a field varies from about 1 minute for a 3x3 test-array to about 18 minutes for a 15x15 QD-array, i.e. 300 μm on a side, and refocussing is not necessary within a distance of 0.5 to 1 mm. In this way, a lot of time and material can be saved. Making a scratch on either side of the sample, exposing around both scratches and cleaving the sample afterwards, even gives two nearly identical samples, e.g. for post-exposure testing such as PEB time and etch conditions, in almost the same time as needed for one sample.

4.1.3. Exposure dose

Having achieved a stable, high-energy electron probe of minimum diameter, the parameter determining the actual exposed volume, given the substrate and resist system, is the exposure time for each dot, i.e. the 'dwell time' of the electron probe: a larger exposure time corresponds to a larger dose, and therefore to larger dots (see below).

Because the intensity of the probe can not directly be set to a certain value, and depends sensitively on the changing beam conditions, the exposure time is only a qualitative measure of the dose arriving at the sample. The average probe current induced by the electrons hitting the sample, however, is a direct measure of this dose and is therefore used here. It can be easily monitored by a picoampère meter, which will average the signal as the dot-pattern is written (2601 dot-exposures in about 3 seconds for each field). This measured value is found to be roughly proportional to the exposure time, as expected. See Fig. 4.3. Of course, this value depends on the density of the pattern, and will therefore be different for different patterns. In this case, the exposure times can be used as a guide to find the right point exposure dose.

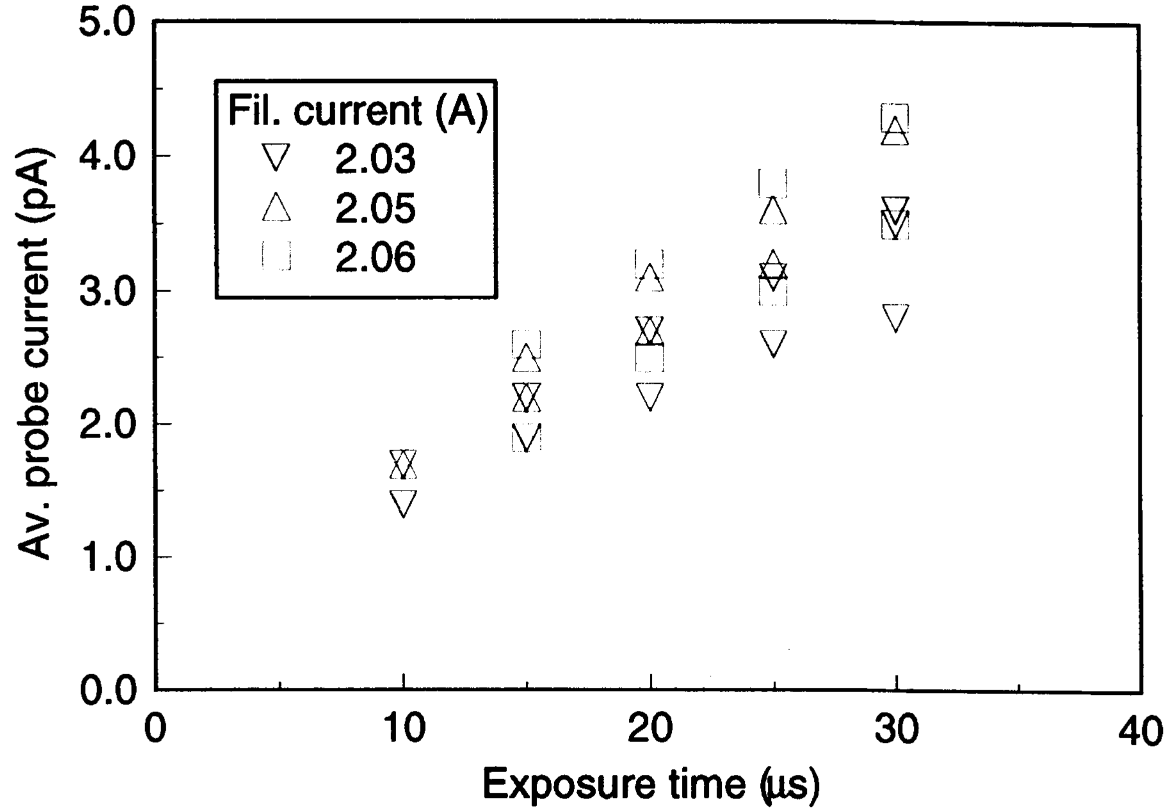


Fig. 4.3. *Average probe current versus exposure time for different samples. All points are taken at (or very close to) optimum filament current, which changes typically from 2.06 A for a new filament to 2.02 after several weeks of use.*

Fig. 4.3 illustrates that the probe intensity is indeed very sensitive to the conditions, all points being taken at or very close to optimum gun emission. The different filament currents in the graph arise from the observation that the value for optimum operation gradually decreases during the lifetime of the filament.

When a high-energy electron beam hits the resist, the resist will be exposed over an area wider than that of the incoming electron beam. This lateral exposure gives rise to what is called the proximity effect: if two patterns are exposed very close to each other, there will be some exposure of the resist between them, as illustrated in Fig. 2.8 and Fig. 4.4.

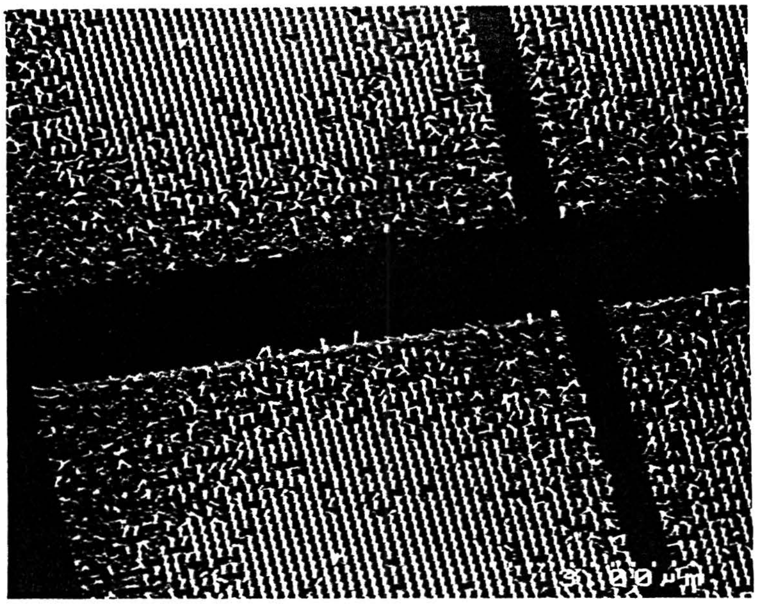


Fig. 4.4. *SEM photograph of some 10 μm fields, showing the 'proximity' effect.*

As Fig. 2.8 shows, this undesired, lateral exposure, extends as far as ~3 μm from the pattern edge, even in the thin resist layers used here. In thick resists, high exposure doses and dense patterns, the effect is more pronounced. Fig. 4.4 illustrates what can happen to dense (150 nm spacing) dot patterns: in the middle of the exposed fields dots are visible, whereas near the edge most of the dots are broken or gone, because they received a lower dose (less backscattering from nearby exposures) and, as a result, were practically etched away. The same thing happens in less dense patterns (spacing 300 nm in 20 μm fields), but can be alleviated by reducing the exposure dose, resulting in uniform arrays of dots. Further reduction of the dose will yield smaller dots, until some or all resist-dots are dissolved during resist-development, or cannot withstand the lowest-damage etch process to produce sufficiently tall pillars, i.e. through the quantum well.

To determine the range of exposure doses that will result in small (<100 nm) to ultrasmall (10-20 nm?) and uniformly sized dots, not only are several exposure series needed, but also an appropriate etch process for the GaAs/AlGaAs system, e.g. low-damage, selective, directional, etc.

4.2. Etching

4.2.1. Introduction

Given that the chlorides of arsenic, gallium and aluminium ($AsCl_3$, $GaCl_3$, $AlCl_3$) are very volatile, and chlorine-based plasmas hardly attack organic materials such as resist, Cl_2 is a widely used gas for etching GaAs/AlGaAs. Because of its high selectivity, it is very sensitive to especially organic contaminations. Therefore, a brief asher step, an oxygen plasma, is used to remove any residual resist that may remain in the unexposed portions of the resist pattern.

In numerous papers on etched quantum dots [Davis93, Sotomayor91, Sotomayor92, Wang92] the reported PL efficiency is much lower than expected: the PL intensity for dots smaller than 200 nm in diameter, corrected for the geometric fill factor, is at best the same as for the 2-D quantum well. Although surface damage is expected, this is usually ascribed to intrinsic phenomena, limiting energy and momentum relaxation, such as the 'phonon bottleneck' effect [Benisty91]. However, MOCVD-grown dots of ~35 nm diameter do exhibit a high luminescence efficiency due to a virtually damage-free fabrication method [Nagamune94]. Also, [Bestwick95] report on ~55 nm diameter ECR-etched dots with a corrected luminescence enhancement of about 10, which shows that it is possible to fabricate efficient (etched) quantum dots when care is taken to minimize surface damage. As discussed in section 2.3, ECR-etching in an ultrahigh vacuum system is a way of strongly reducing the ion energy in the plasma and thus associated surface contamination and damage. Of course, low (RF) powers, resulting in low ion energies, and low pressures must be used.

4.2.2. Etch conditions

The following method is investigated here: a combination of a short, low-power, Ar/O_2 asher step, and a Cl_2 ECR etch step, both at minimum pressures (about 5 and 1 mtorr respectively). Samples are introduced into the process chamber using a nitrogen purged vacuum load lock. Etching is started when the chamber (base) pressure is below $1\ -10^{-7}$ torr.

The etch depth, or pillar height, is expected to be proportional, with a possible off-set, to the Cl_2 etch time, and independent of the exposure dose. The SEM measurements on a series of samples, each with about four different exposures, shown in Fig. 4.5, confirm this. In the etch process, only the Cl_2 etch time was varied.

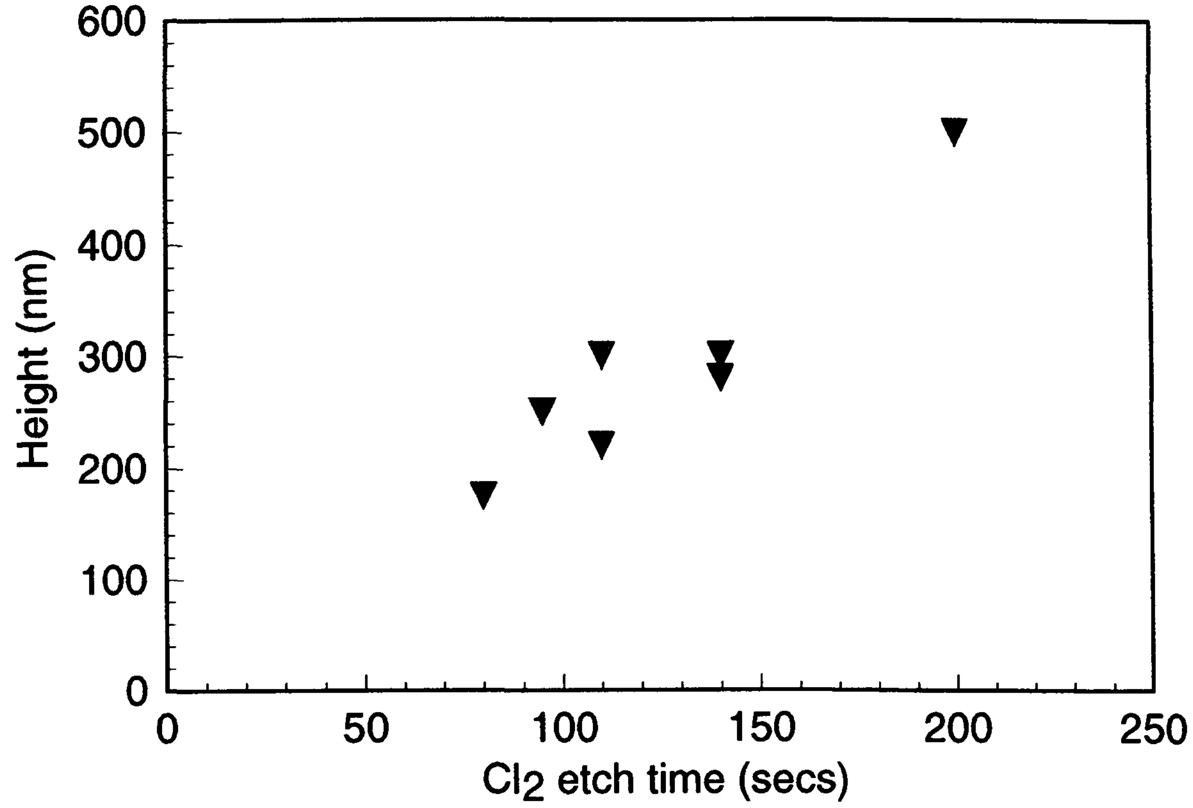


Fig. 4.5. *Measured pillar height versus Cl_2 etch time. Every point represents a different sample, each with about four different exposure doses (resulting in same height).*

At 200 seconds Cl_2 etching only a few pillars are left at the highest exposure dose, the rest are completely gone. This indicates that the resist is slowly etched away, possibly by physical sputtering, and that there probably is a slight undercut etch-profile. Obviously, shorter etch times are necessary.

The diameters of the different dots, on the same samples as above, as a function of the average probe current are plotted in Fig. 4.6, and increase with increasing dose, as expected.

For a fixed dose, the diameter increases with shorter etch times, supporting the above remark that the resist is slowly etched away. This is illustrated in Fig. 4.7. The differences in exposure between the top and bottom portions of a resist are a distinguishing characteristic of electron beam lithography: the backscattering of electrons from the substrate dominates over forward scattering in the resist, and will result in greater exposure at the bottom of the resist film than at the top. As in the proximity effect, the differences will be larger for higher exposure doses (see Fig. 4.6).

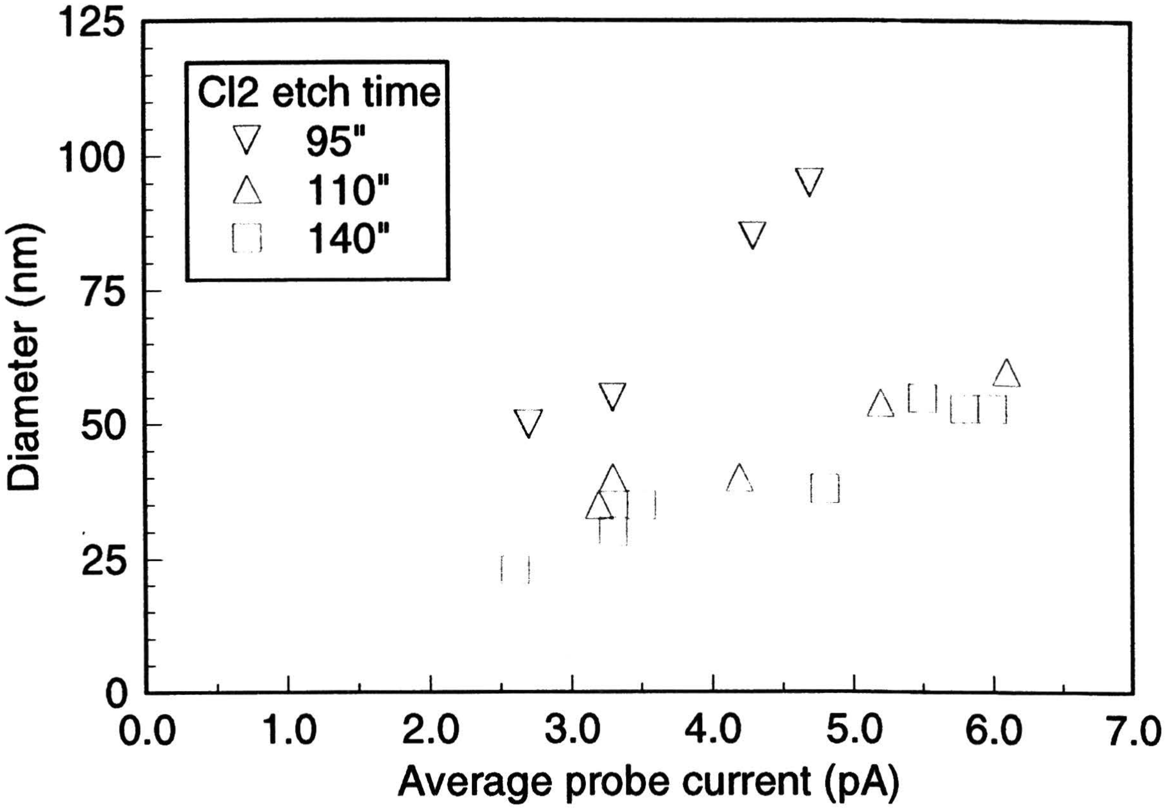


Fig. 4.6. *Dot diameter versus average probe current for different Cl_2 etch times.*

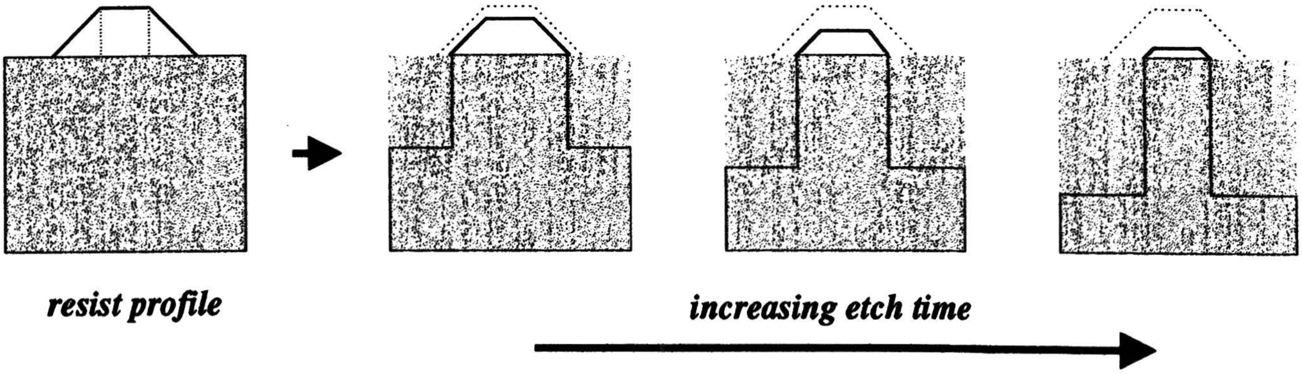


Fig. 4.7. *Schematic illustration of the influence of longer etch times on the resist profile and resulting structure diameters.*

The results of the two samples etched for 140 seconds will now be discussed as a summary and an illustration of the above. Both samples were exposed and etched on the same day, under identical conditions. The only difference is that the resist on one sample was applied two weeks before processing into dots, whereas that of the other sample was only a few hours old. Since the results are the same for both, the conclusion is that the properties of the resist do not deteriorate in a period of at least two weeks. This was confirmed in all following experiments.

At the highest exposure dose, i.e. 5.8 and 6.0 pA, there is still a sign of the proximity effect: the pillars near the edge are narrower (55 nm) than in the middle (60 nm) of the fields. Some resist is left on top of the pillars, see Fig. 4.8.

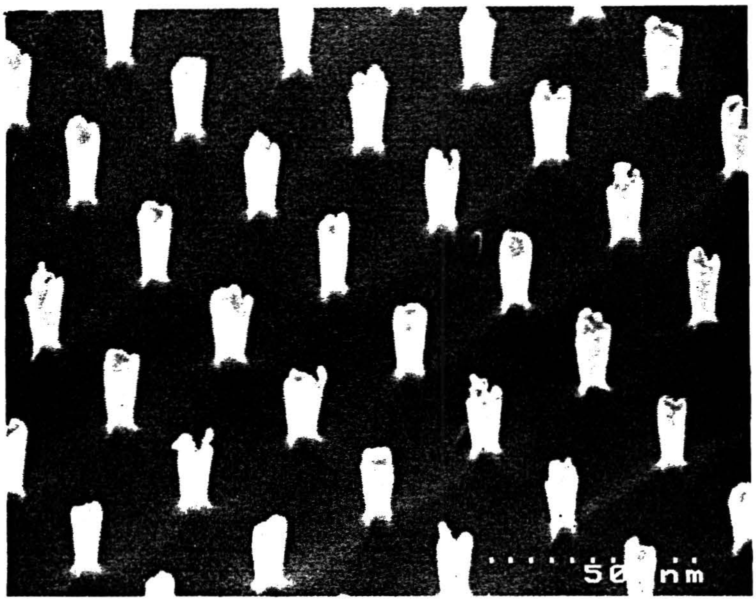


Fig. 4.8. *SEM photograph of 55 nm GaAs pillars.*

Reducing the dose to 3.3 & 3.5 pA yields uniform pillars, with a diameter of ~30 nm. Fig. 4.9 shows that the pillars have a slight undercut profile: they are larger near the top of the pillar. No resist is left in this case.

Further reduction to 2.6 pA causes about half of the pillars to break: their diameter near the surface is 20-25 nm (Fig. 4.10), corresponding to an aspect ratio of about 12!

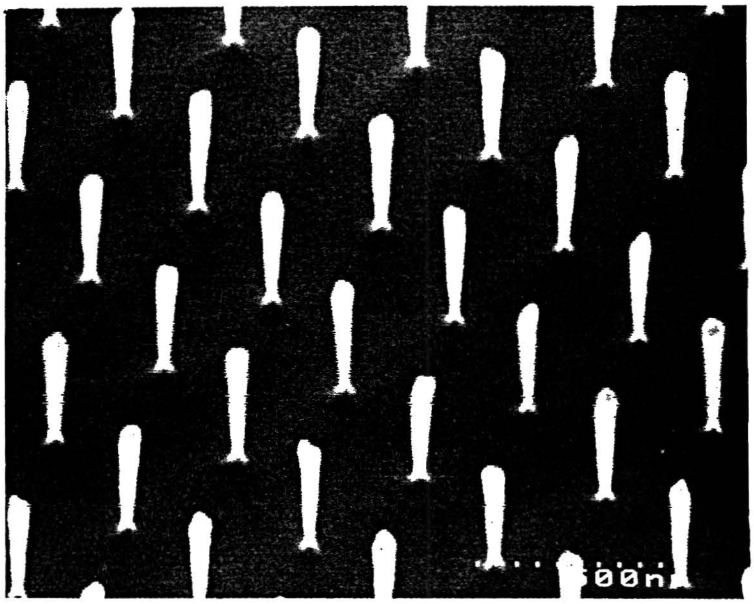


Fig. 4.9. *SEM photograph of 30 nm GaAs pillars.*

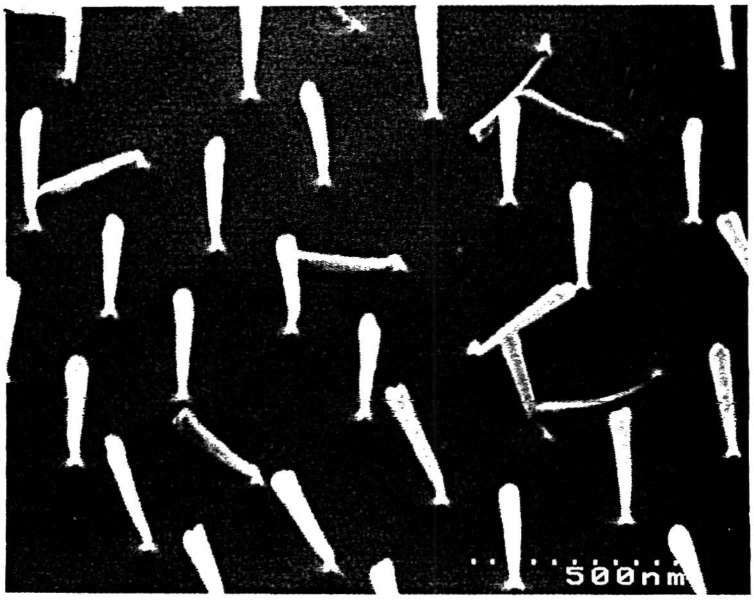


Fig. 4.10. *SEM photograph of 20-25 nm GaAs pillars, 50% broken.*

From this example it is clear that for narrower pillars both height and amount of undercut must be reduced. Simply reducing the Cl_2 etch time is not the solution, as the dot diameter will increase. The same effect as the one illustrated in Fig. 4.7 can be achieved by increasing the oxygen cleaning step, without affecting the pillar height. So, decreasing the Cl_2 etch time to get shorter pillars, but still etching through the quantum well, and increasing the O_2 etch time will result in smaller dots: see Fig. 4.11 and Table 4.1 for details.

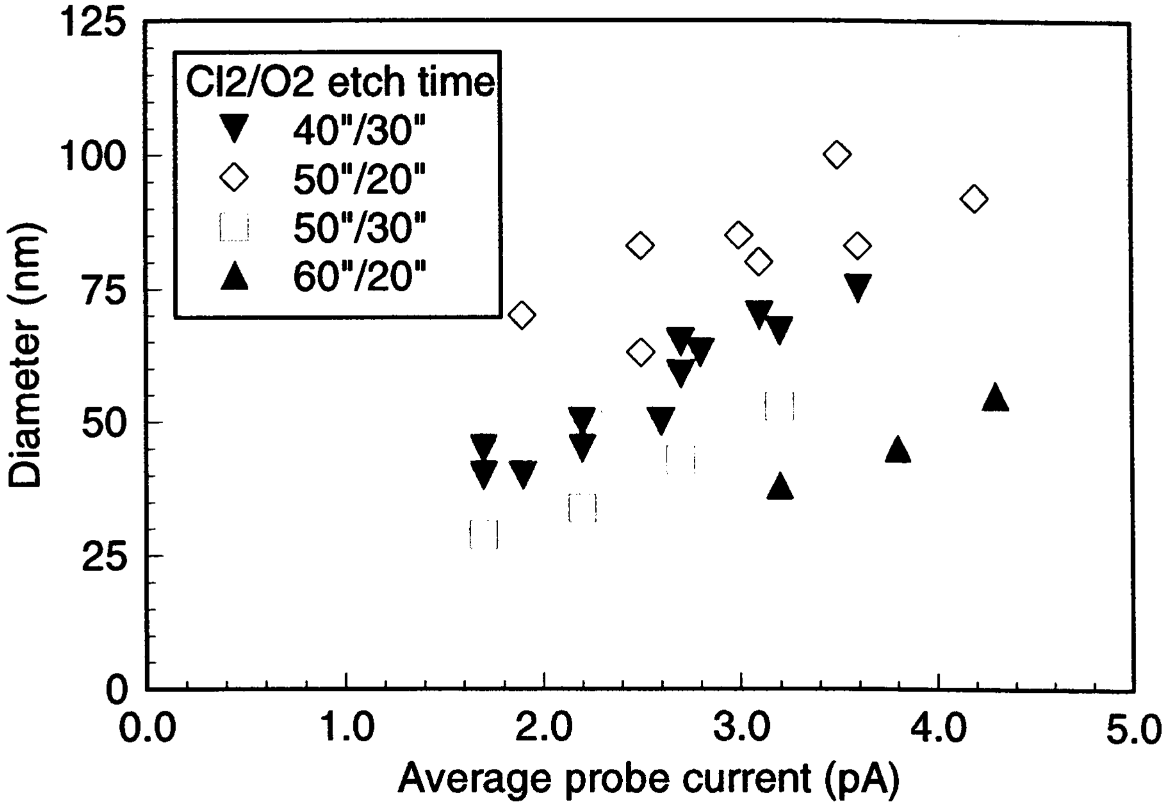


Fig. 4.11. *Dot diameter versus average probe current for different Cl_2/O_2 combinations.*

Table 4.1. *Pillar details for different Cl_2/O_2 combinations.*

Cl_2/O_2	profile tapering?	height (nm)	min. diam.	comments
60/20	yes	250	25 nm	50 nm at top
50/20	some	200	65 nm	
50/30	slight	170	30 nm	20 % broken
40/30	almost none	140	40 nm	good

For a given etch process the dot diameter can also be reduced by using shorter post-exposure bake (PEB) times for the resist. The crosslinking chemistry requires both acid (catalyst locally generated by exposure) and heat to produce insolubility. A shorter PEB will therefore cause areas with less acid to remain soluble. This results in smaller dots of resist for intermediate PEB times. During PEB the acid will also diffuse somewhat, but according to [Perkins93] this will only be a minor effect. It was found that after etching, for average probe currents below about 4 pA, no resist was left for PEBs of 2 minutes or less. See Fig. 4.12. Here, the Cl_2/O_2 = 40"/30" process was used. Note that the samples used above had PEBs of 5 minutes.

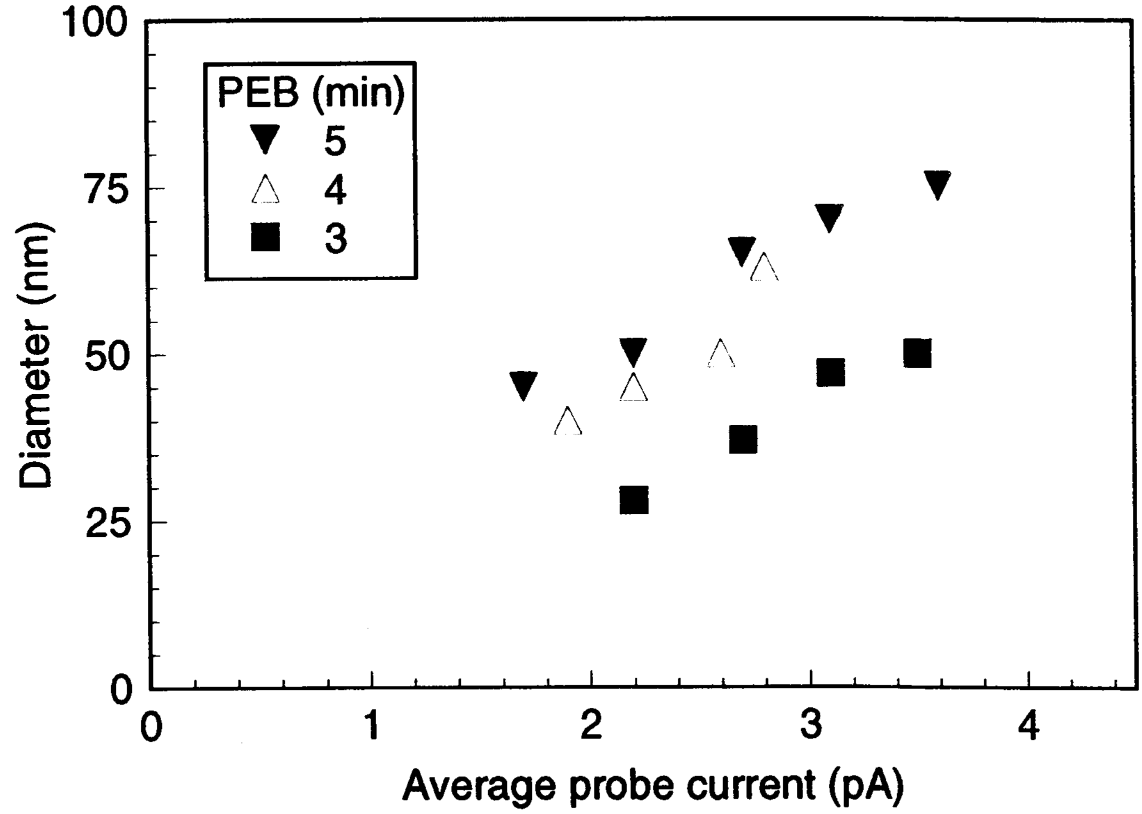


Fig. 4.12. *Dot diameter versus average probe current for different PEB times.*

4.3. Discussion

The resist application gives no problems, and the resist film remains stable for at least two weeks when the samples are stored in a cool place. Although the electron beam exposure is not straightforward, it yields reproducible results when care is taken to optimize the beam conditions. An exposure dose range is determined, giving uniform dot arrays, with controllable diameters.

The etch process, however, proves to be very sensitive to parameters that can not always be controlled, such as residual gas in the chamber, and contaminations. For example, it was found that running an ECR process using other gases than Cl_2 and O_2 caused dramatic changes in the etch characteristics of identically processed samples that were etched before and after this process. Apparently, pumping the process chamber down to a vacuum better than $1 \; -10^{-7}$ torr, and even running the Cl_2/O_2 process a few times before introducing the sample, is not sufficient to 'clean' the system to fixed initial conditions. It may therefore be necessary to recalibrate the process from time to time in order to get satisfactory results, especially when other processes are run: only the Cl_2 time, to determine the desired etch depth, and possibly the O_2 time need to be adjusted. The other process parameters, such as gas flow, RF power, pressure, magnet supply, etcetera need not be changed. Etching test-samples with exposure series under varying conditions, makes it possible to (re)optimize the process in about five attempts. Examining the structures in the SEM before etching the next sample, is a great help in determining whether to increase or decrease the time and by what amount.

5. Luminescence from GaAs/(Al,Ga)As quantum dots

5.1. Introduction

A number of factors have complicated investigations of the intrinsic electronic and optical properties of 1-dimensional and zero-dimensional semiconductor nanostructures. These structures, involving carrier confinement on a length scale of about 10 nm, are difficult to produce, and imperfections in the fabrication process can mask intrinsic effects. Lateral size fluctuations, for example, are a potential problem in any process producing the arrays or grids of quantum dots or wires required for practical applications. A high density of surface states, impurities, and damage, with a consequent reduction in radiative efficiency, can be introduced by the commonly applied fabrication techniques, especially those involving plasma-based dry etching [Clausen89, Maile89]. Assuming these effects can be minimized, care must still be taken in the interpretation of optical data used to characterize the structures. One example is that a blue-shift in the PL in patterned quantum well structures, compared to the unprocessed quantum well material, can result not from a shift in energy levels caused by additional confinement, but by inhibition of lateral transport to potential minima associated with well width fluctuations [Gustaffson93]. Another is the conflicting evidence, both for decreased [Wang92, Sotomayor92] and for increased [Marzin94, Davis93, Nagamune94, Galeuchet91], PL efficiency with decreasing dot diameter in studies seeking to investigate the proposed 'phonon bottleneck' inhibited relaxation mechanism [Benisty91].

5.2. Samples

Highly uniform and efficient arrays of free-standing quantum dots with diameters in the range 25-60 nm have been made from GaAs/AlGaAs 5 nm single quantum well (SQW) material. Electron beam lithography and low-damage electron cyclotron resonance (ECR) etching were used to fabricate the dots, which are believed to be the smallest ever produced in this way. The SQW material has been characterized using low temperature (5 K) PL and PLE before and after processing into quantum dots.

Before presenting the optical data obtained from the quantum dot samples, the different samples and the SQW material from which they were fabricated, will be discussed.

5.2.1. SQW material

The single quantum well structures used here, S97 and S53, were grown by gas-source molecular beam epitaxy on a 50 mm diameter semi-insulating <100> GaAs substrate. The 5 (5.5) nm SQW (S97, S53) was sandwiched between a 50 (100) nm $Al_{0.3}Ga_{0.7}As$ top barrier layer and a 20 (40) nm $Al_{0.3}Ga_{0.7}As$ lower barrier, all grown on a 0.5 μm GaAs buffer layer and with a 5 nm GaAs top protective cap. None of the layers were intentionally doped. After growth, and before any processing, nine PL and PLE mapping points were taken across each of the 1x2 cm samples. The measured energies of the PL maxima and the clear e1-hh1 (HH) and e1-lh1 (LH) exciton transitions in PLE (see below) were found to deviate by within 0.3 meV for all, indicating high uniformity of both samples. The PL linewidth was 3.0 and 2.3 meV FWHM and the measured Stokes shift (section 5.3) was 7.4 and 2.2 meV for S97 and S53 respectively, as indicated in Fig. 5.1. Also note the clearly resolved continuum edges for both free exciton peaks in S53. Using an envelope function model of proven applicability ([Bastard82] and [Dawson85]) to fit the measured PLE energies yields a well width of 4.47 nm for S97 and 5.74 nm for S53. The input parameters are: $m_e^* = 0.0665\ m_0$; $m_{HH}^* = 0.3400\ m_0$; $m_{LH}^* = 0.0940\ m_0$ (see section 3.3.4); $T = 5.0$ K; x (Al-fraction) = 0.28; exciton binding energy = 9 meV and a conduction-valence band offset ratio of 67 : 33.

5.2.2. Quantum dot samples

From the spectra shown in Fig. 5.1 it can be seen that S53 was a higher quality sample than the still very good S97: very small linewidth and Stokes shift, and clearly resolved continuum edges for both free exciton peaks all attest to the high quality. Its top barrier, however, is much thicker (100 nm instead of 50 nm), making it extra difficult to produce small dots, since larger etch depths are required.

Aiming for the very smallest dots, one dot sample was fabricated from S53 with 33.5 nm dot diameter; the rest of the available material was used up on tests. Further work was done on bulk GaAs until dot diameters from 20 to >60 nm could be reproduced in S97, the second-best SQW material in store at the time: samples with uniform arrays of dots with diameters of 60, 45, and 25 nm were fabricated. Appendix C contains SEM photographs and spectra of all four quantum dot samples. SEM examination showed that no unetched SQW material remained on the sample, and that the uniformity in diameter of the dots was typically ±2.5 nm. The etched surfaces of the dots were not intentionally passivated, and remained exposed to air.

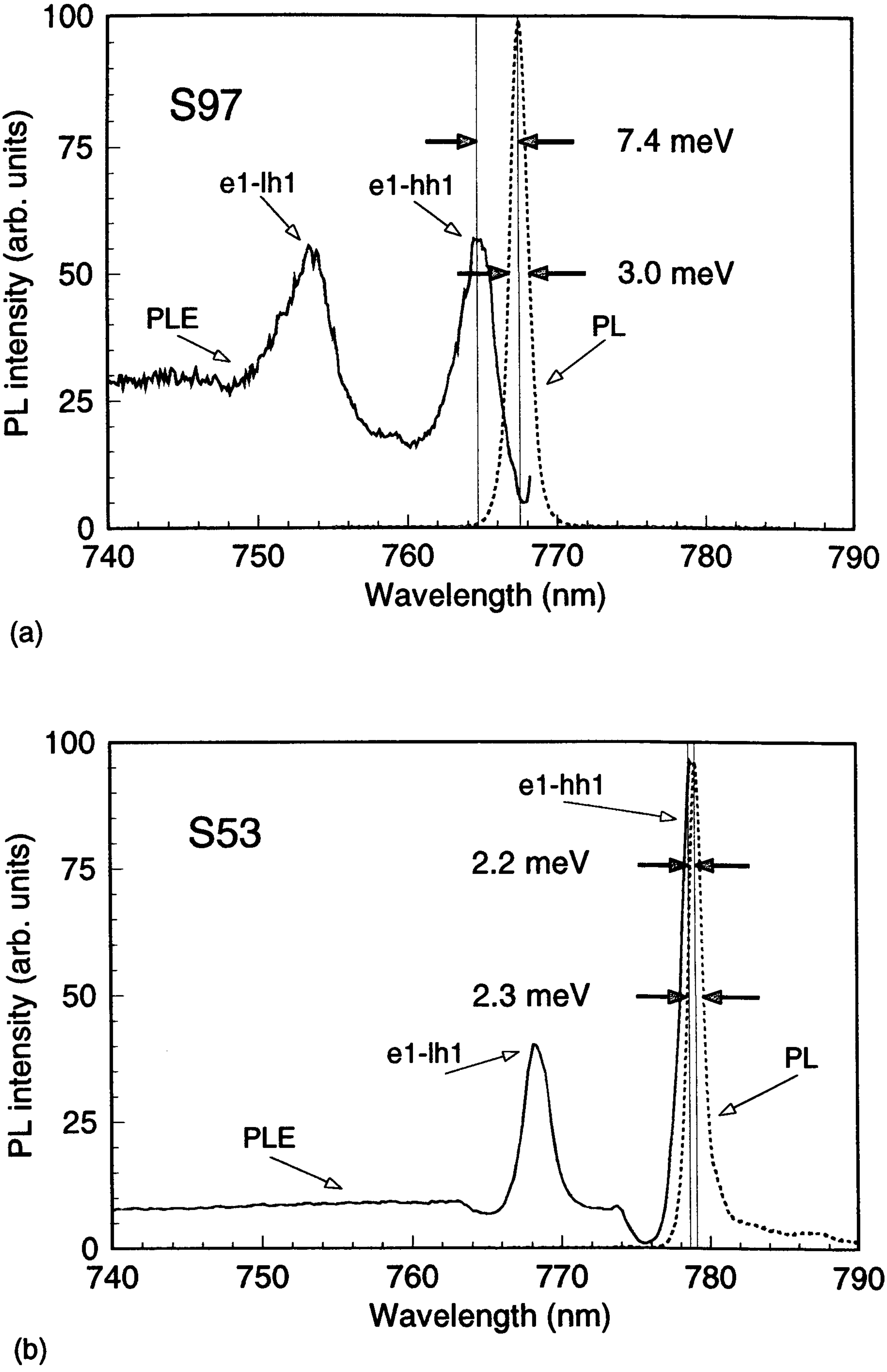


Fig. 5.1. *PL and PLE spectra for (a) S97 and (b) S53, indicating PL linewidth and Stokes shift.*

An array of quantum dots of 18-20 nm in diameter was also produced, but about 80 % of the dots were gone or broken, making it very difficult to find any luminescence: no diffraction pattern was visible, so only coarse positioning could be performed, and the expected luminescence signal would be weak, even if the alignment of the collecting optics was optimal. A few attempts were made, but, unfortunately, without success.

5.3. Optical data

The PL and PLE spectra for both the 25 nm dots and the reference SQW material, taken within 2 mm of the area subsequently processed into quantum dots, are shown in Fig. 5.2. They are normalized to their own PL maxima. In fact, the integrated dot PL intensity is less than 2 times smaller in the 25 nm dots, and about 4 times smaller in the 60 nm dots than that of the reference material (see below). The spectra of all the quantum dot samples and their reference spectra can be found in Appendix C, together with the SEM photographs. No luminescence was found when the excitation spot was moved off the dot array.

Standard excitation and detection conditions were used, with the excitation power fixed at 150 μW and the PL excited at 740 nm, i.e. directly into the quantum well. 'In-well' excitation is used so that the effect of non-radiative recombination at the free surfaces of the AlGaAs barrier layers is minimized. Therefore, PL peak intensities can be compared directly. The focussed laser spot size (~50 μm FWHM) was intentionally smaller than the total dot array area (300x300 μm), so that the correction factor for the integrated intensity is simply the geometric fill factor of the dots. It still covers an estimated 20,000 dots (dot spacing 300 nm).

Blue-shifts

Clear free exciton features have been observed in the excitation spectra from the dots. Fig. 5.2 shows clearly that both the e1-hh1 (heavy hole, HH) and e1-lh1 (light hole, LH) free exciton peaks, together with the PL emission, are blue-shifted with respect to the reference material. This is due to the additional confinement of the carriers, as described in section 1.3. The increase in HH-LH splitting, as expected from the much smaller light hole effective mass and therefore larger energy shift, is also observed. Both energy shifts and splitting increase with decreasing dot diameter: Fig. 5.3 shows the different blue-shifts (PL, HH, and LH) as a function of dot diameter. The measured positions of the heavy and light hole excitons shift by less than 1 meV in 60 nm dots, but are shifted by 8 and 20 meV respectively in 25 nm diameter dots. Because PLE directly probes the density of allowed states, the HH and LH

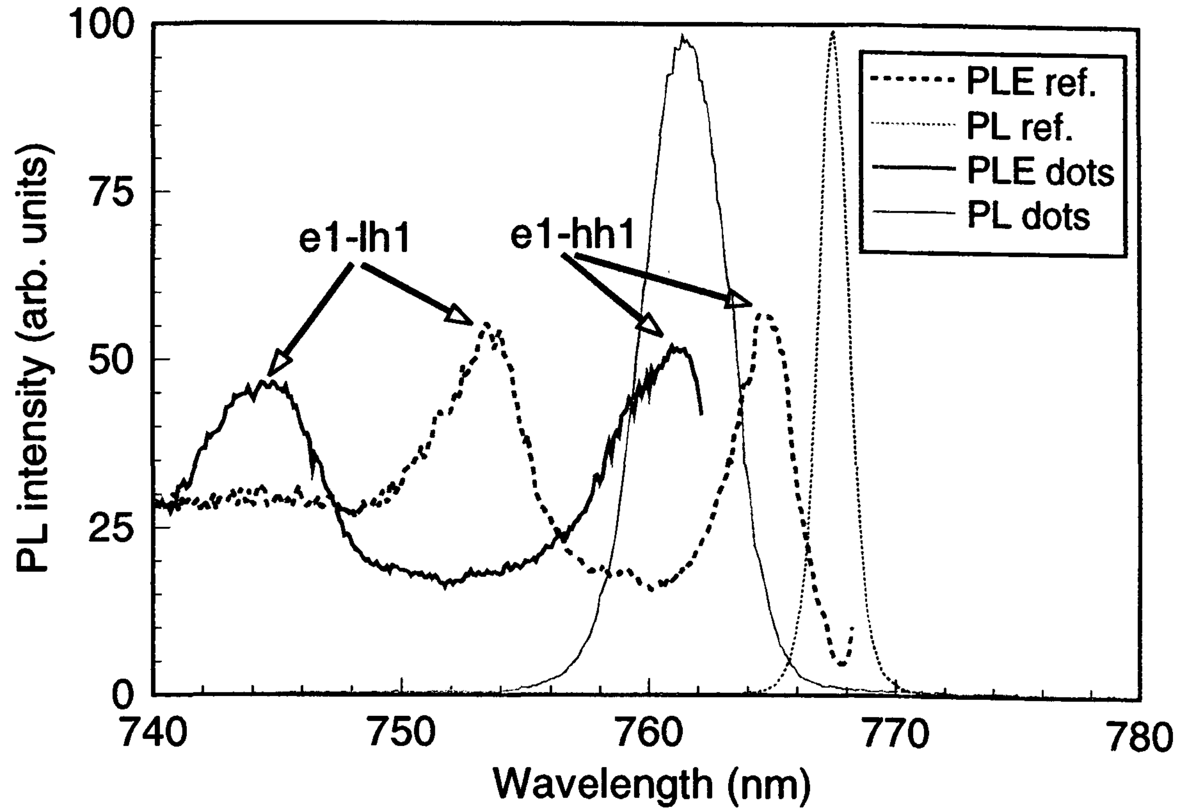


Fig. 5.2. *PL and PLE spectra for both the 25 nm dots and the reference SQW material.*

shifts should be equal to the sum of the additional electron and hole (heavy and light, respectively) lateral confinement energies (equation (1.12)), and are plotted in Fig. 5.3.

The calculated values for both HH and LH shifts are much larger: about 4 meV for 60 nm to about 17 meV for 25 nm dots. According to [Bryant91] (see also section 1.3) the Coulomb interaction between the electrons and holes significantly enhances the exciton binding energies in this size range. In order to estimate the order of magnitude of these excitonic effects, the exciton binding energies by [Xia89] in spherical GaAs dots in air are plotted in Fig. 5.4, together with the calculated electron and hole eigenstates minus the observed energy shifts. Note that the bulk values for the effective masses are used in the calculations.

Because the experiments were done on quantum discs rather than spheres, there are a few differences to be considered:

- In our case, polarization effects due to the dielectric mismatch at the air/GaAs interface will be much smaller, because only the sidewalls of the disc are in contact with air.

- The thickness of the disc is small compared to the diameter D, so that the effective diameter will be smaller, and changes like $D^{2/3}$, resulting in a shift and a change in slope.

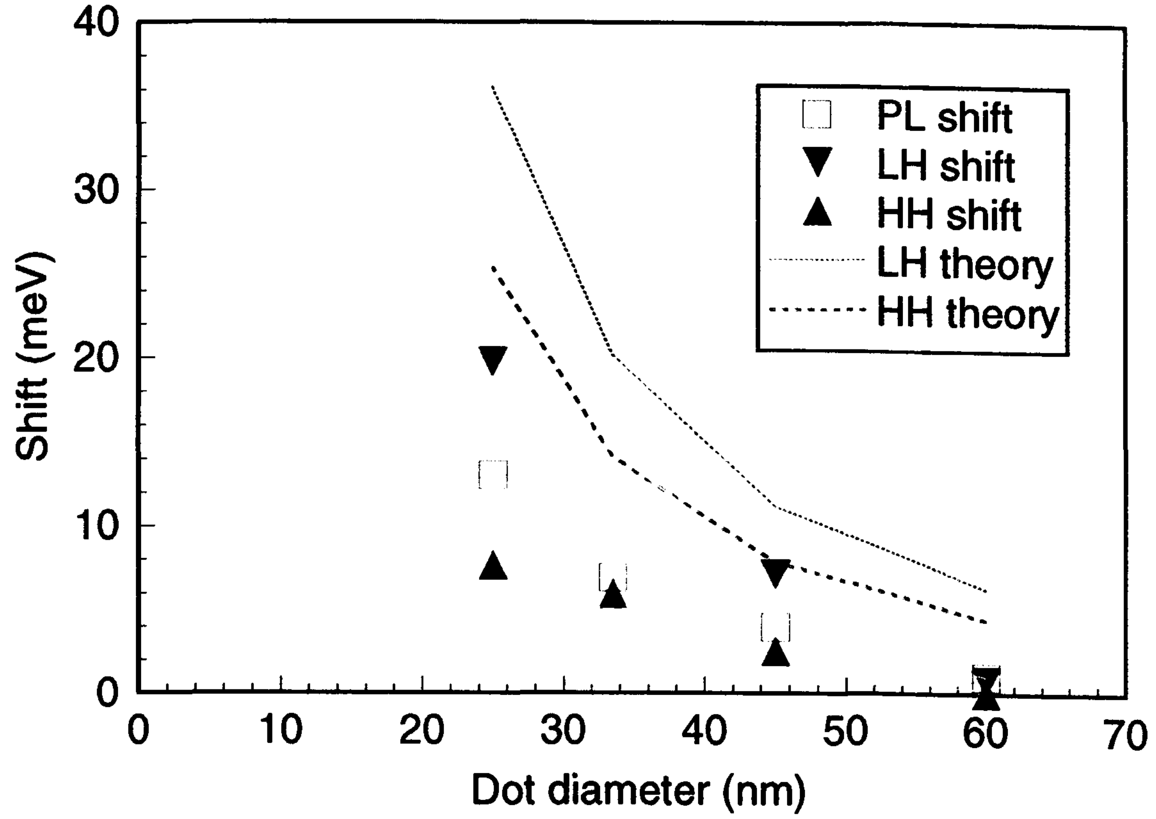


Fig. 5.3. *PL, HH, and LH blue-shifts with respect to SQW reference. The broken lines show the values calculated by considering single particle states only.*

- Wavefunction leakage into the barrier will result in a larger effective diameter for the discs.

The excellent agreement for both HH and LH, when neglecting polarization, is probably due to cancellation of errors in the effective diameters. It is, however, a good first order approximation, describing both trend and order of magnitude. Therefore, we believe that excitonic effects play an important role in the explanation of the significant discrepancy between the electron and hole eigenstates and the observed energy shifts, but further theoretical work would be needed to confirm this.

Fig. 5.2 shows that there is a difference between the PL and the HH peak position of the reference SQW, which is called the Stokes shift. This shift is observed, because PL also has a contribution from lateral carrier migration: upon laser excitation carriers are generated in the quantum well, that will rapidly relax to the lowest energy state, which is the e1-hh1 (HH)

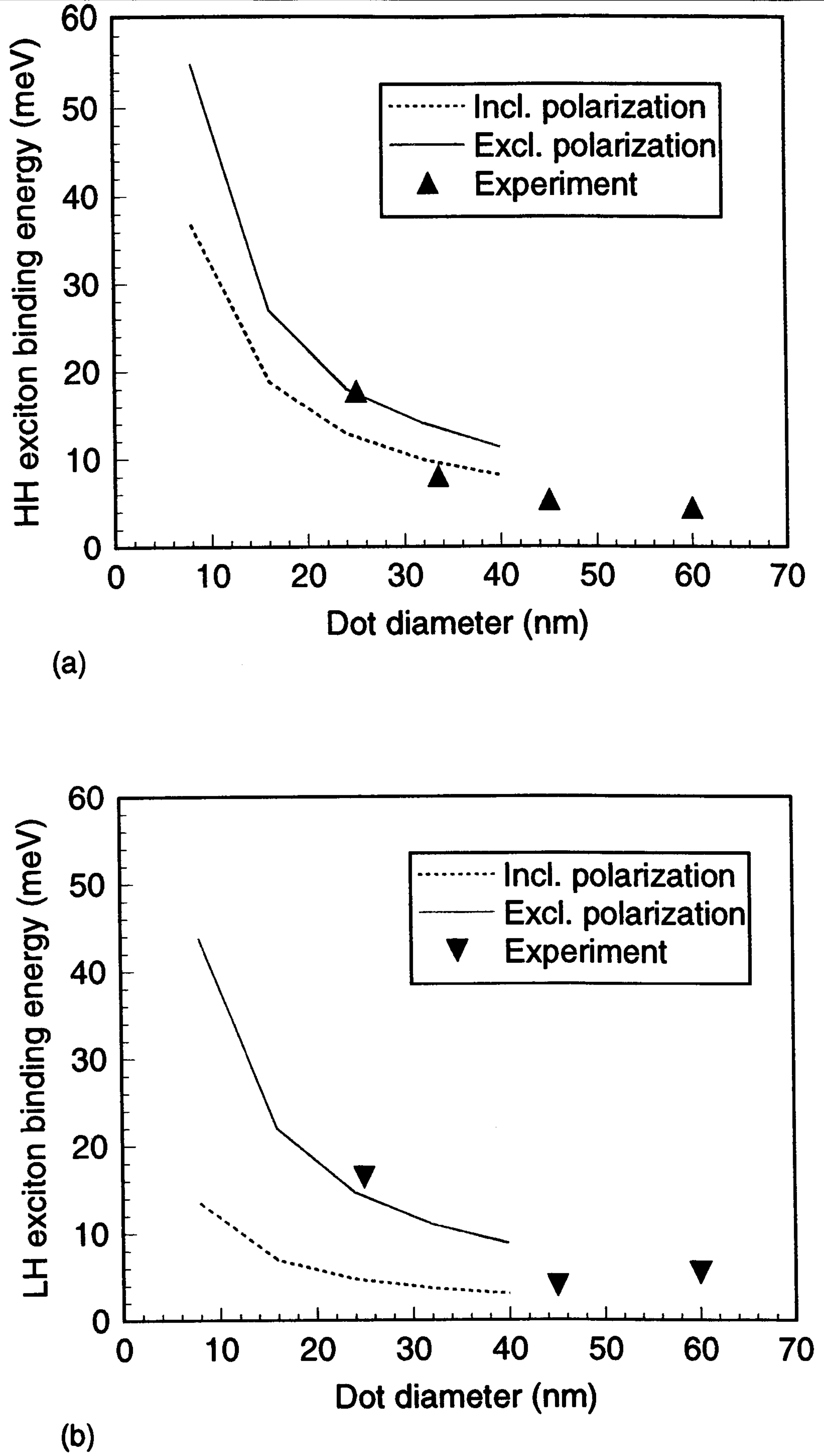


Fig. 5.4. *Exciton binding energy versus dot diameter according to [Xia89] for (a) HH and (b) LH. The triangles indicate the experimental values.*

exciton, before they undergo radiative recombination. In PLE, the spectrum is determined by changes in the absorption coefficient as a function of the excitation wavelength, resulting in changes in the observed luminescence intensity. It therefore directly probes the density of states. In PL, however, the formed HH excitons can laterally migrate to larger well widths (even in very high quality SQW there are fluctuations in well width of the order of a monolayer) before they recombine, thereby giving a luminescence peak at a lower energy. The energy difference between PL and HH, the Stokes shift, therefore is a measure of the well-width uniformity of the SQW.

This process of carrier migration can also explain why the PL peak shifts more than the HH peak in the dot samples, and increasingly so for smaller dot diameters (Fig. 5.3): the more the excitons are laterally confined in the dots, the more the transfer of excitons to larger widths is reduced. The higher energy contribution, corresponding to smaller well widths, will therefore increase, while the lower energy contribution decreases. For spectrally unresolved monolayer splittings this change in ratio will be observed as an additional PL blue-shift and a broadening of the peak. In other words, the Stokes shift will decrease by patterning the SQW into dots (or wires), and will become smaller for smaller dot diameters. This is illustrated in Table 5.2, listing the Stokes shifts for both reference and dot samples.

Table. 5.2. *Observed Stokes shifts in reference and dot samples.*

Dot diameter (nm)	Ref. shift (meV)	Dot shift (meV)	SQW material
60	7.4	6.3	S97
45	7.4	5.9	S97
33	2.2	1.2	S53
25	7.4	1.9	S97

PL intensity

As mentioned before, it is possible to compare the PL intensities of the reference material and the dot samples directly (same excitation and detection conditions), and the geometric fill factor of the dots can be used to calculate the integrated dot PL intensity.

The measured PL peak intensities of the dots, with respect to their reference material, drop by a factor of 6 for the 60 nm dots, to 4.8 for the 25 nm dots. The integrated intensity, however, only drops by a factor of 4 and 1.8 for 60 and 25 nm dots respectively. When the geometric fill factor of the dots is taken into account (2.1 % for 60 nm, to 0.36 % for 25 nm), the intensity is enhanced by a factor of 12 for the 60 nm dots, which increases to 150 for the 25 nm dots, as is shown in Fig. 5.5.

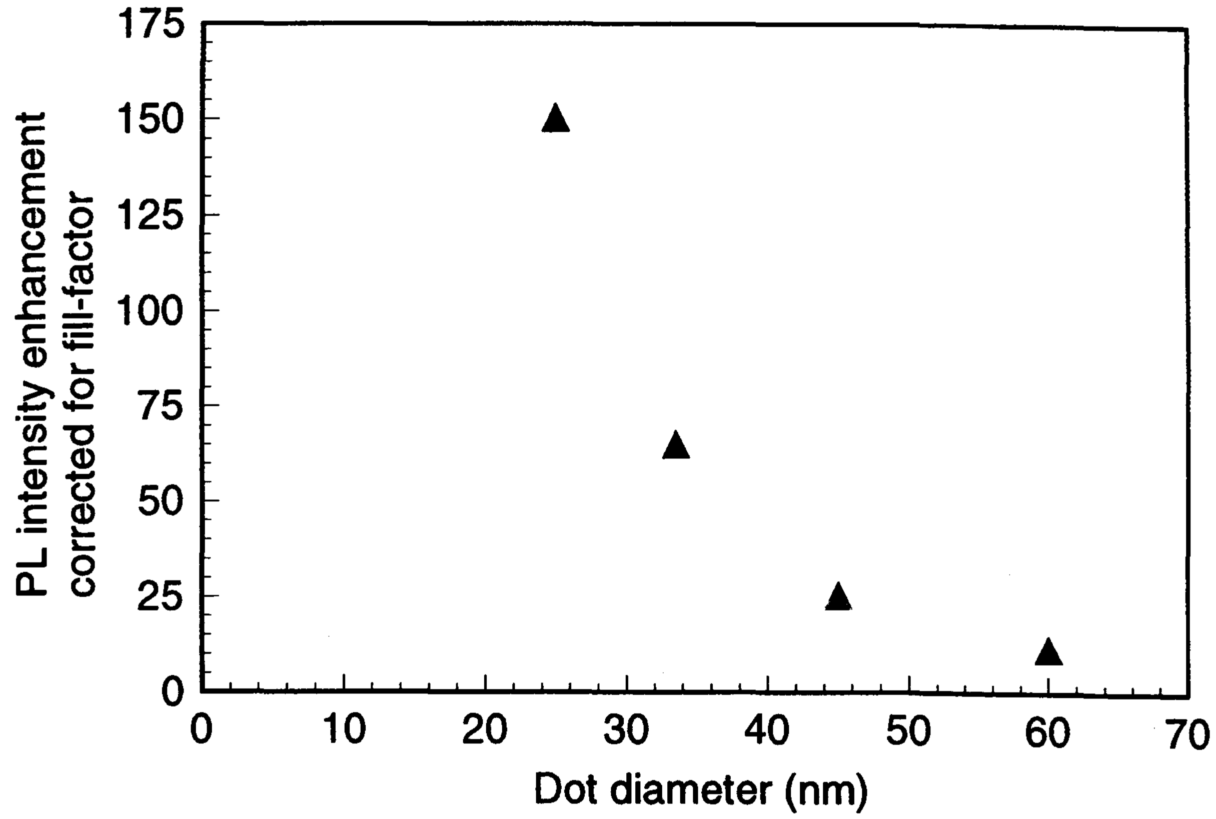


Fig. 5.5. *Intensity enhancement factor with respect to reference SQW versus dot diameter.*

The efficient luminescence of these quantum dots suggests that the surface damage layer, due to the ECR etch process is thin compared to the dot diameter. This is plausible since, in the absence of channelling, normal incidence 145 eV ions from the plasma would be expected to penetrate only a few monolayers into the surface [Burenkov86].

The increase in intensity with smaller dot diameter is expected from the increased spatial overlap between the electron and hole wavefunctions, which concentrates the oscillator strength as the 0-D situation is approached. Additional contributions invoke electrodynamic effects [Marzin94] (better coupling with incoming laser light) or the suppression of in-plane diffusion in the patterned samples. Diffusion lengths at 5 K for free excitons in GaAs quantum wells are on the order of microns [Takahashi94]. This could allow them to find nonradiative defects in the quantum wells in a manner not possible for quantum dots; the same mechanism can allow significant carrier migration out of the optical excitation/collection spot in the quantum wells. The effect in both cases is to lead to an apparent decrease in the luminescence efficiency of the unpatterned reference.

The PL data show no evidence of the orders of magnitude reduction in quantum efficiency for dots below 100-200 nm predicted by the 'phonon bottleneck' mechanism [Benisty91]. It is expected that the increase in intensity will break down as the dot size approaches the exciton diameter (100-200 Å): in that case correlation of electrons and holes is no longer possible, because carrier motion is 'frozen out' by the confinement, so excitons can no longer exist.

Linewidth

In the ideal situation, i.e. no fluctuations in size (well width, dot diameter), no defects, damage or contamination, the linewidth is determined only by the carrier lifetimes (uncertainty principle). In real material, however, defects, size fluctuations, etcetera are present, that will broaden the peak.

The PL linewidths of the SQW reference samples are 3.0 and 2.3 meV FWHM for S97 and S53, respectively. In the dot samples, the linewidths increase somewhat with decreasing diameter (see Fig. 5.6) but remain in the range 9 to 4 meV: [Nagamune94, Arakawa93] (MOCVD) find 20-25 meV and [Davis93, Marzin94] (etched) about 40 meV FWHM for comparable diameters!

The significant broadening of these linewidths is usually attributed to extrinsic effects, such as size distributions in the structures: dots of different diameter result in different energy shifts of the centre of gravity of the '2-D' peak (PL, but also HH and LH), resulting in a broadening. Another contribution comes from the fact that the dot pillars are not perfect cylinders, causing extra scattering and therefore additional broadening. The lateral carrier confinement, and the resulting reduction in Stokes shift discussed above, will also broaden the peak.

The size fluctuation of the dot samples is estimated from SEM measurements (absolute accuracy within ±2 nm), and is found to be less than ±2 nm for the 60 nm dots, gradually increasing to ±3 nm for the 25 nm dots. These estimates are used to calculate the contribution to the broadening from fits through the different measured blue-shifts (Fig. 5.3). In other words, the measured size fluctuation is 'translated' into a fluctuation in energy blue-shift, i.e. an increase in linewidth. The results are summarized in Table 5.3. The error in the calculated values ('fit') is about 0.5 meV for the largest dots, and about ±2 meV for the 25 nm dots.

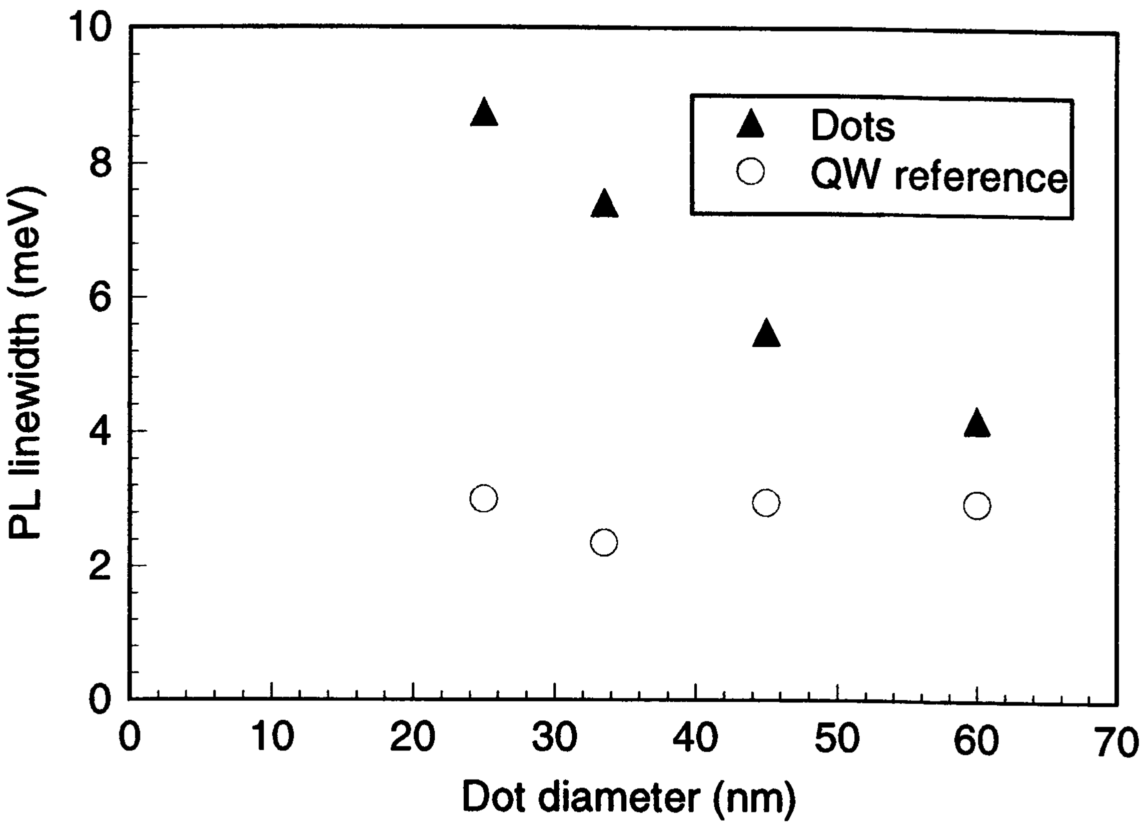


Fig. 5.6. *PL linewidth versus dot diameter.*

Table. 5.3. *Measured FWHM linewidths (meV) for PL, HH, and LH in reference SQW ('ref.') and dot samples ('dot'), and the observed and calculated increase in linewidth ('obs.' and 'fit').*

size in nm	PL width (meV)		increase in linewidth		HH width (meV)		increase in linewidth		LH width (meV)		increase in linewidth	
	ref.	dot ±1	obs. ±1	fit	ref.	dot ±2	obs. ±2	fit	ref.	dot ±2	obs. ±2	fit
60	3.0	4	1	1	6.3	6	0	0.3	8.5	8	0	1
45	3.0	5	2	1.5	5.7	9	3	1.5	8.3	12	4	2.5
33	2.3	7	5	2.5	3.0	7	4	1.5	4.5	-	-	4
25	3.0	9	6	~6	5.4	9	4	~3	7.8	11	3	~6

Although this method is simple and somewhat crude, it does show that the broadening is due mainly to (small) size fluctuations introduced in the fabrication process. The relatively large broadening in the 33 nm dots is probably due to additional scattering effects: comparing the SEM photographs in Appendix C shows that these dots have a more irregular profile.

6. Conclusions

The results presented in this report show that it is possible to fabricate very small, uniform and efficient quantum dots, using electron beam lithography and low-damage electron cyclotron resonance (ECR) plasma etching. Uniform arrays of free-standing quantum dots with dimensions in the range 25-60 nm have been made from GaAs/AlGaAs 5 nm single quantum well material. The 25 nm dots are believed to be the smallest ever produced in this way.

For the optical characterization of the quantum well material and the quantum dots produced from it, two spectroscopic techniques are used here: low-temperature (5 K) photoluminescence (PL) and photoluminescence excitation (PLE) spectroscopy. The quantum dots are compared to the unprocessed reference material (QW) in PL efficiency, linewidth and energy, but factors such as wafer nonuniformity can make evaluation of apparent energy shifts difficult. Therefore, the QW material is extensively mapped (3x3 array on 1x2 cm wafer) by both PL and PLE before processing, so that the characterization afterwards can be compared to the reference spectra taken within 2 mm of the area processed into dots.

Standard excitation and detection conditions were used, with the excitation power fixed at 150 μW and the PL excited at 740 nm, i.e. directly into the quantum well. 'In-well' excitation is used so that the effect of non-radiative recombination at the free surfaces of the AlGaAs barrier layers is minimized. Therefore, PL peak intensities can be compared directly. Although the etched surfaces of the dots are not intentionally passivated and remain exposed to air, strong photoluminescence is measured from the features. When corrected for the area of the quantum well remaining in the dots, the intensity increases rapidly as the dot diameter is reduced. After correcting for this 'fill-factor', the integrated PL intensity of the 25 nm dots is more than 150 times greater than the reference unprocessed SQW, which is a remarkable increase. This suggests that the surface damage layer, due to the ECR etch process, is thin compared to the dot diameter, and shows the advantages of ECR etching over traditional etch techniques.

The PL linewidth from the dots, and the heavy (HH) and light (LH) hole exciton linewidths, as observed in PLE, increase somewhat with decreasing dot diameter. However, they remain in the range 9 to 4 meV (FWHM) for PL and HH, and 12 to 8 meV for LH, compared with the 3 meV PL, 6 meV HH and 8 meV LH linewidths form the unprocessed single quantum well material, indicating high uniformity. The fluctuations in dot diameter estimated from SEM measurements are ± 2 nm, and this size distribution is thought to be the main contribution to the broadening of the linewidth in both PL and PLE.

PLE shows clear free heavy and light hole excitonic features in the quantum dots. They are increasingly blue-shifted with decreasing dot diameter, due to the additional lateral confinement of the carriers. The measured positions of the heavy and light hole excitons shift by less than 1 meV in 60 nm dots, but are shifted by 8 and 20 meV respectively in 25 nm diameter dots. PLE is preferable to PL when determining blue-shifts because, representing predominantly the absorption spectrum, it is less sensitive to local fluctuations in the quantum well width. This study is the clearest observation yet of light and heavy hole excitons in a range of quantum dot sizes.

The measured blue-shifts are 4 to about 16 meV smaller for LH, and 4 to 18 mev for HH than the values calculated by considering electron and hole eigenstates only. We believe that this difference is due to an enhanced exciton binding energy: comparison with theoretical studies [Xia89] yields a very good, although first order, agreement in both trend and order of magnitude for both heavy and light hole excitons.

The quantum dots discussed in this report are free-standing and remain exposed to air. In order to make a device with an active region containing these quantum dots, the etched pillars must be 'buried', e.g. by MBE overgrowth. This might necessitate applying a passivating layer on the dots while in the vacuum etch chamber, to protect the surface from oxidation. Studies of such passivation and regrowth over quantum dots are probably the next step in investigating the properties and potential applications of quantum dots formed by lithography and etching.

References

[Banyai87] L. Banyai, Phys. Rev. B **36**, 6099 (1987).

[Bastard82] G. Bastard, Phys. Rev. B **12**, 7584 (1982).

[Benisty91] H. Benisty, C.M. Sotomayor-Torrès, and C. Weisbuch, Phys. Rev. B **44**, 10945 (1991).

[Bestwick95] T.D. Bestwick, M.D. Dawson, A.H. Kean, and G. Duggan, submitted to Appl. Phys. Lett.

[Broers88] A.N. Broers, IBM J. Res. Develop. Vol. **32** no. 4 (july 1988).

[Burenkov86] A.F. Burenkov, F.F. Komarov, M.A. Kumakhov, and M.M. Temkin, *Tables of Ion Implantation Spatial Distributions* (Gordon and Breach, New York, 1986).

[Chapman80] B.N. Chapman, *Glow Discharge Processes, sputtering and plasma etching* (John Wiley & sons, 1980).

[Clausen89] E.M. Clausen, Jr., H.G. Craighead, J.M. Worlock, J.P. Harbison, L.M. Schiavone, L. Florez, and B. van der Gaag, Appl. Phys. Lett. **55**, 1427 (1989).

[Davies91] J.H. Davies, A.R. Long, *Physics of Nanostructures* (Glasgow University and SUSSP, St Andrews, 1991).

[Davis93] L. Davis, K.K. Ko, W.-Q. Li, H.C. Sun, Y. Lam, T. Brock, S.W. Pang, P.K. Bhattacharya, and M.J. Rooks, Appl. Phys. Lett. **62**, 2766 (1993).

[Dawson85] P. Dawson, G. Duggan, H.I. Ralph, K. Woodbridge, and G.W. 't Hooft, Superlatt. Microstruct. **1**, 231 (1985).

[Galeuchet91] Y.D. Galeuchet, H. Rothuizen, and P. Roentgen, Appl. Phys. Lett. **58**, 2423 (1991).

[Gustaffson93] A. Gustaffson, X. Liu, I. Maximov, L. Samuelson, and W. Seifert, Appl. Phys. Lett. **62**, 1709 (1993).

[Haug90] H. Haug, and S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (World Scientific, London, 1990).

[Herman92] J.S. Herman, and F.L. Terry, Jr., Appl. Phys. Lett. **60** (6), 716 (1992).

[Jaros90] M. Jaros, *Physics and Applications of Semiconductor Microstructures* (Clarendon Press, Oxford, 1989, reprinted 1990).

[Kittel76] C. Kittel, *Introduction to solid state physics* (John Wiley & sons, New York, 1976).

[Maile89] B.E. Maile, A. Forschel, R. Germann, D. Grützmacher, H.P. Meier, and J.-P. Reithmaier, J. Vac. Sci. Technol. B **7**, 2030 (1989).

[Marzin94] J.-Y. Marzin, A. Israel, and L. Birotheau, Solid State Electron., **37**, 1091 (1994).

[Nagamune94] Y. Nagamune, M. Nishioka, S. Tsukamoto, and Y. Arakawa, Appl. Phys. Lett. **64**, 2495 (1994).

[Perkins93] F.K. Perkins, E.A. Dobisz, and C.R.K. Marrian, J. Vac. Sci. Technol. B **11**(6), (1993).

[Reimer85] L. Reimer, *Scanning Electron Microscopy. Physics of Image Formation and Microanalysis* (Springer-Verlag, Berlin, 1985).

[Schiff68] L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, Tokyo, 1968).

[Sotomayor91] C.M. Sotomayor-Torrès, P.D. Wang, W.E. Leitch, H. Benisty, and C. Weisbuch in *Optics of Excitons in Confined Systems* (Institute of Physics Publishing, London, 1991).

[Sotomayor92] C.M. Sotomayor-Torrès, W. Leitch, D. Lootens, P. Wang, G. Williams, S. Thomas, H. Wallace, P. van Daele, A. Cullis, C. Stanley, P. Demeester, and S. Beaumont, in *Nanostructures and Mesoscopic Systems* (Academic, San Diego, CA, 1992), p.455.

[Takahashi94] Y. Takahashi, S.S. Kano, K. Muraki, S. Fukatsu, Y. Shiraki, and R. Ito, Appl. Phys. Lett. **64**, 1845 (1994).

[Wang92] P. Wang, C.M. Sotomayor-Torrès, H. Benisty, C. Weisbuch, and S. Beaumont, Appl. Phys. Lett. **61**, 946 (1992).

[Weisbuch91] C. Weisbuch, B. Vinter, *Quantum semiconductor structures. Fundamentals and applications* (Academic Press inc., 1991).

[Williams90] R. Williams, *Modern GaAs Processing Methods* (Artech House, London, 1990).

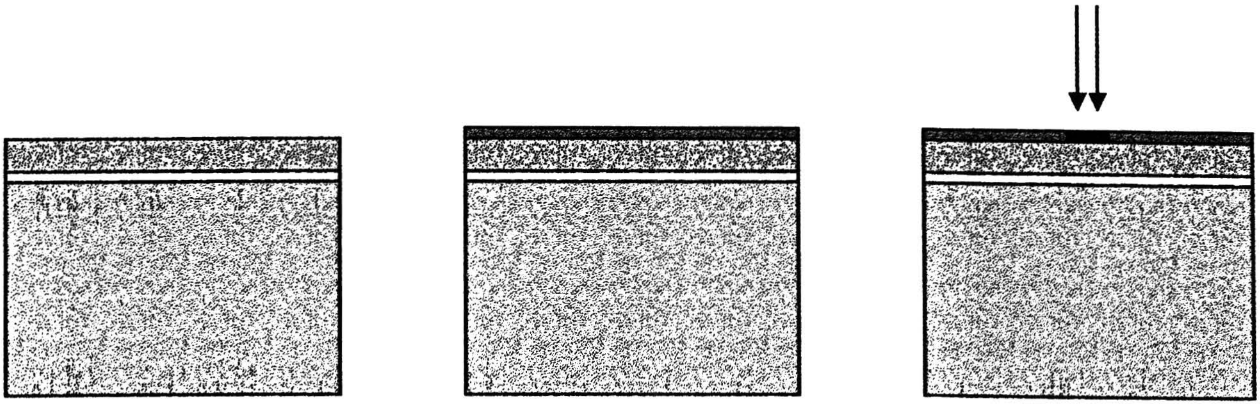
[Xia89] J.-B. Xia, Phys. Rev. B **40**, 8500 (1989).

Acknowledgements

First of all, I wish to thank professor J. Wolter (Eindhoven University of Technology, Netherlands) for offering me this five month project at Sharp Laboratories of Europe (SLE) in Oxford, United Kingdom, and Dr. C.C. Bradley (M.D.) for agreeing to my visit. Thanks also go to everyone at SLE, especially the opto group, for making my stay a pleasant and memorable one, and for sharing (new) culinary experiences.

In particular, I want to thank my supervisors T. Bestwick, M. Dawson, and G. Duggan for their advice, help, and many fruitful discussions, and also J. Rorison for her help with the theory on optical transitions in quantum dots.

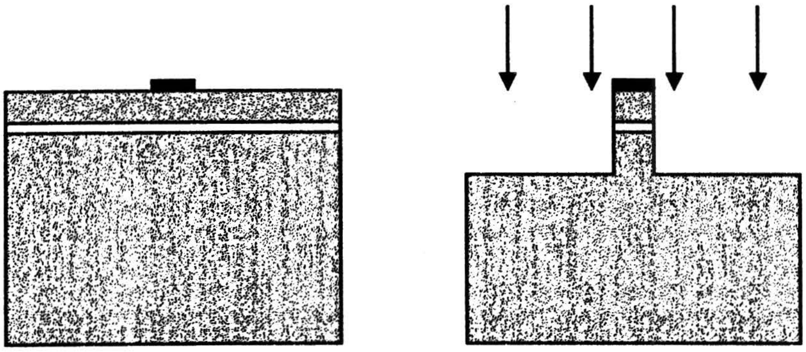
Appendix A. Quantum dot fabrication



1. SQW

2. Resist application

3. Electron beam resist exposure



4. Resist development

5. ECR etching

Fig. A.1-5. *Sideviews of samples, resist application, exposure, etching.*

The resist used here for the quantum dot pattern generation is a high-resolution, negative-tone novolak-based resist, Shipley's Microposit SAL 605. It uses an energy sensitive, acid catalyzed crosslinking system:

- Exposure:
  generation of acid upon exposure to an electron beam.
- PEB (post exposure bake):
  thermally-activated crosslinking in those areas where the catalyst has been generated.
- Development:
  dissolution of the un-crosslinked film.

Appendix B. Laser correction

The spectrograph software package that is used, assumes that the excitation source is scanned linearly. In fact, the titanium:sapphire laser exhibits a slight, almost parabolic deviation for which the PLE spectra must be corrected.

To determine the correction, the laser was scanned through the entire wavelength range used in the optical measurements and the peak position measured by the calibrated spectrometer (the resolution of the spectrometer is 0.005 nm (at 546.1 nm Hg-line), and the long-term accuracy (over 500 nm) is ±0.1 nm). The correction in wavelength, i.e. the difference between the real (measured) and the set wavelength, is plotted in Fig. B.1,

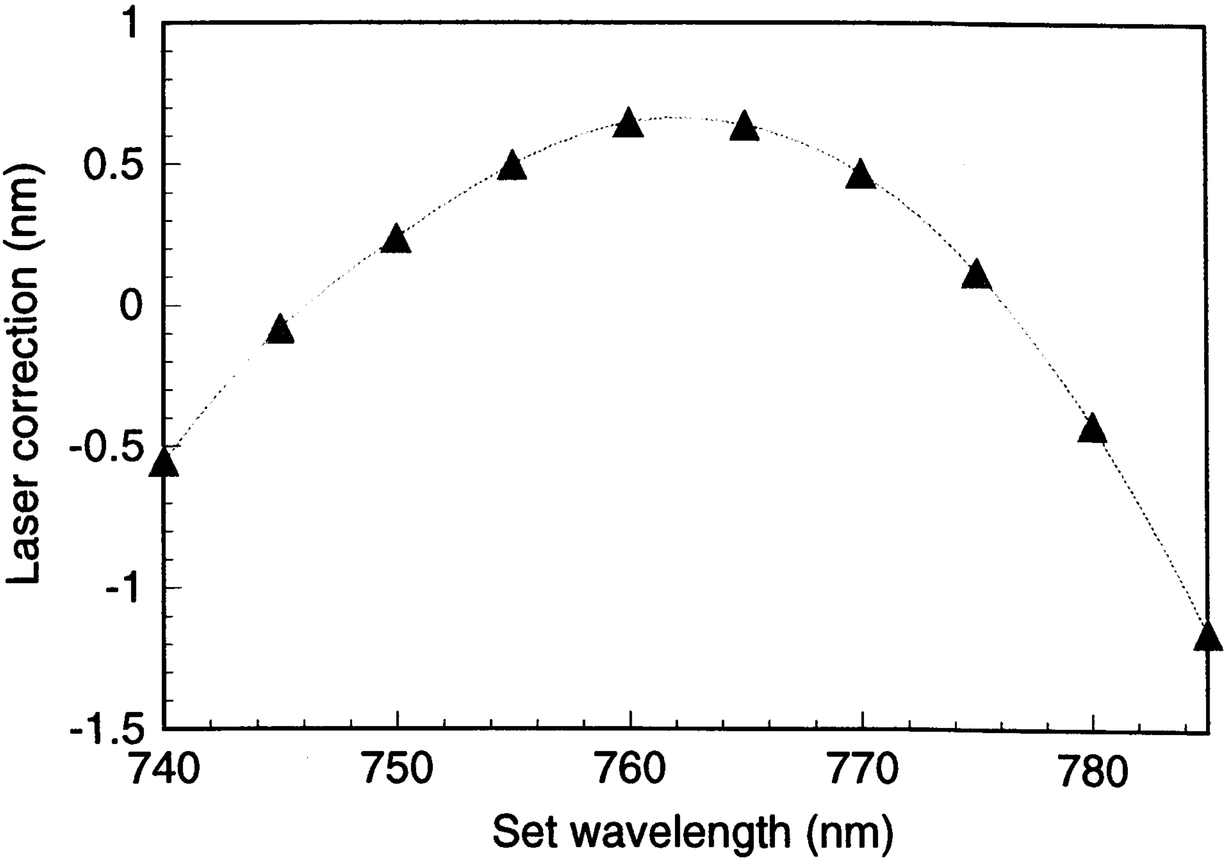


Fig. B.1. *Laser correction (nm) versus set laser wavelength.*

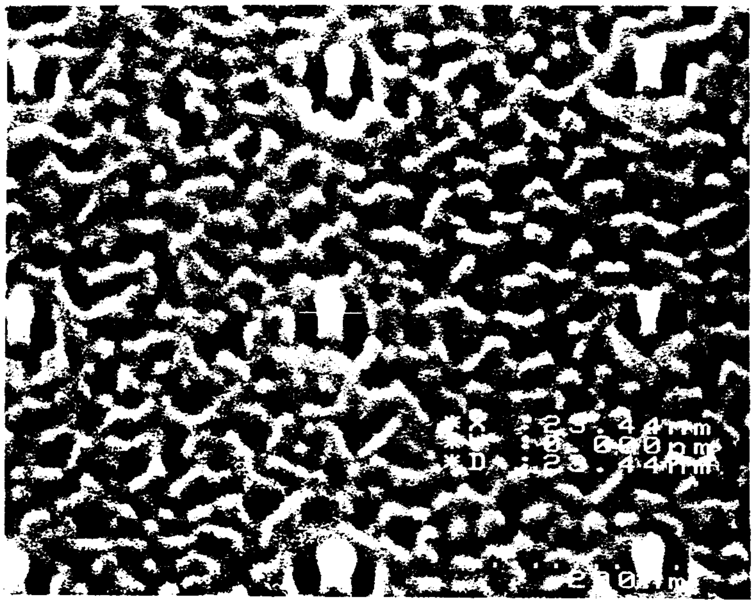
as well as the curve

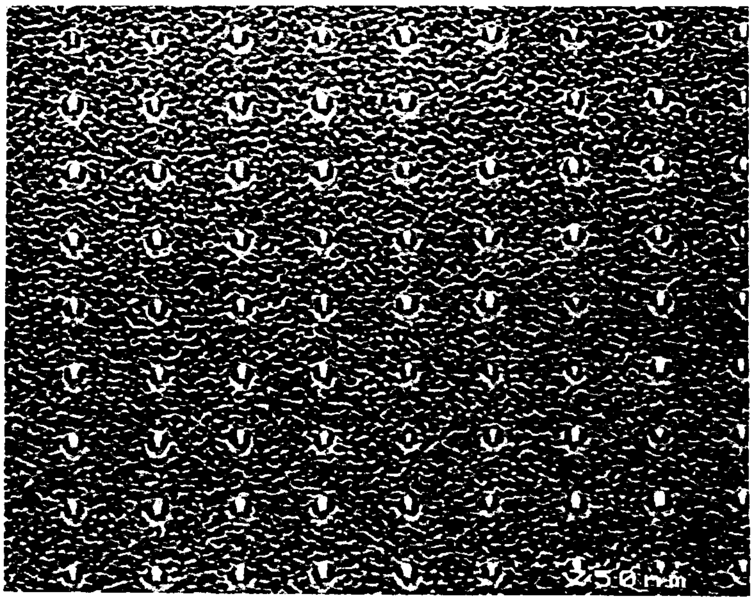
$$\lambda_{corr.} = a_4\lambda^4 + a_3\lambda^3 + \ldots + a_1\lambda + a_0$$

a polynomial of order four, fitted through the measured peak positions. This polynomial is used to correct all PLE spectra.

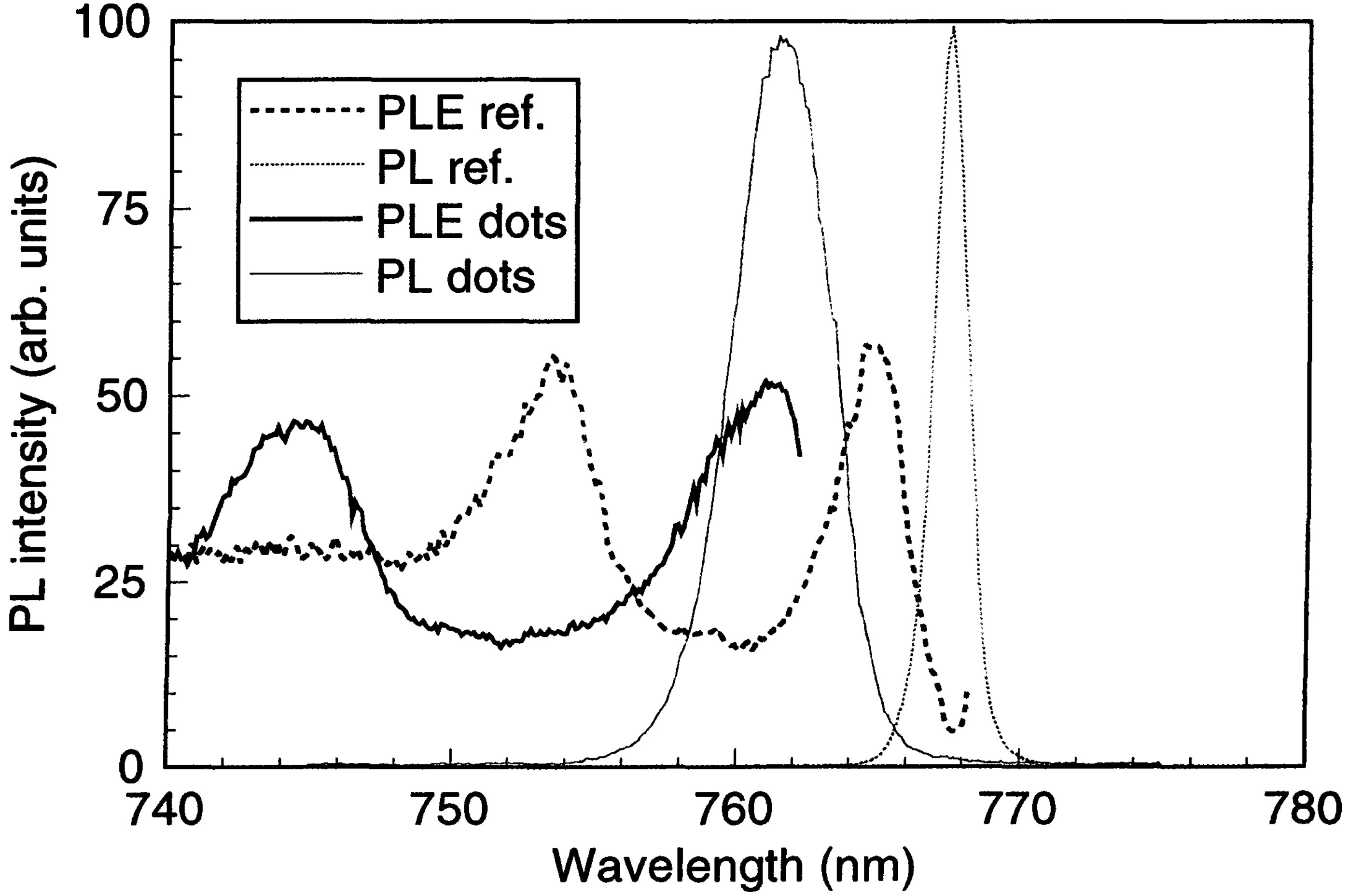
Appendix C. Quantum dot SEM photographs and spectra

25 nm

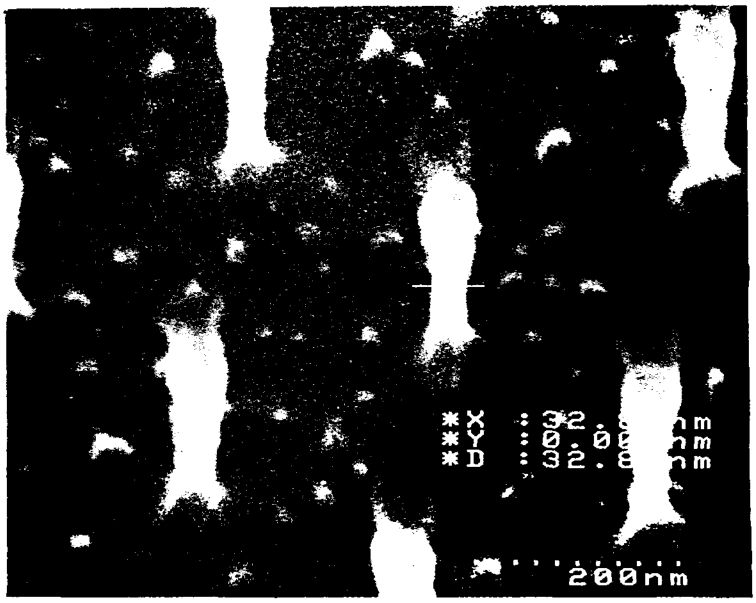


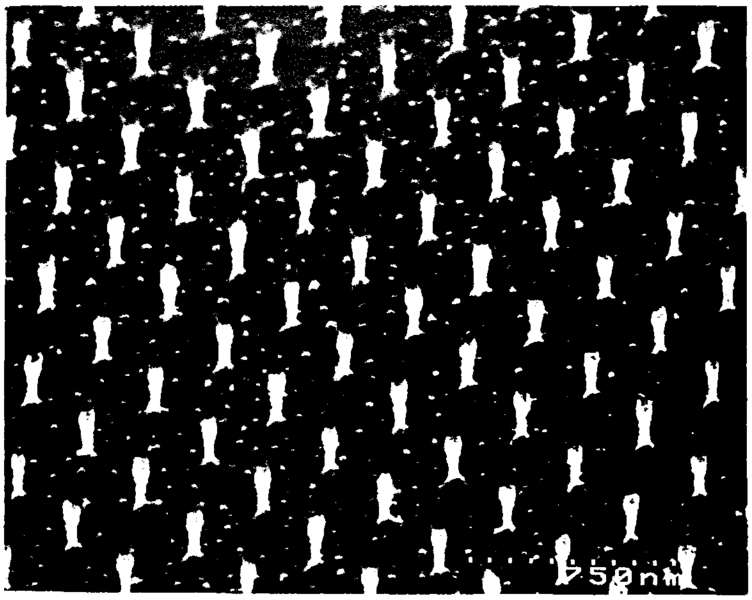


Dot diameter = 25 nm

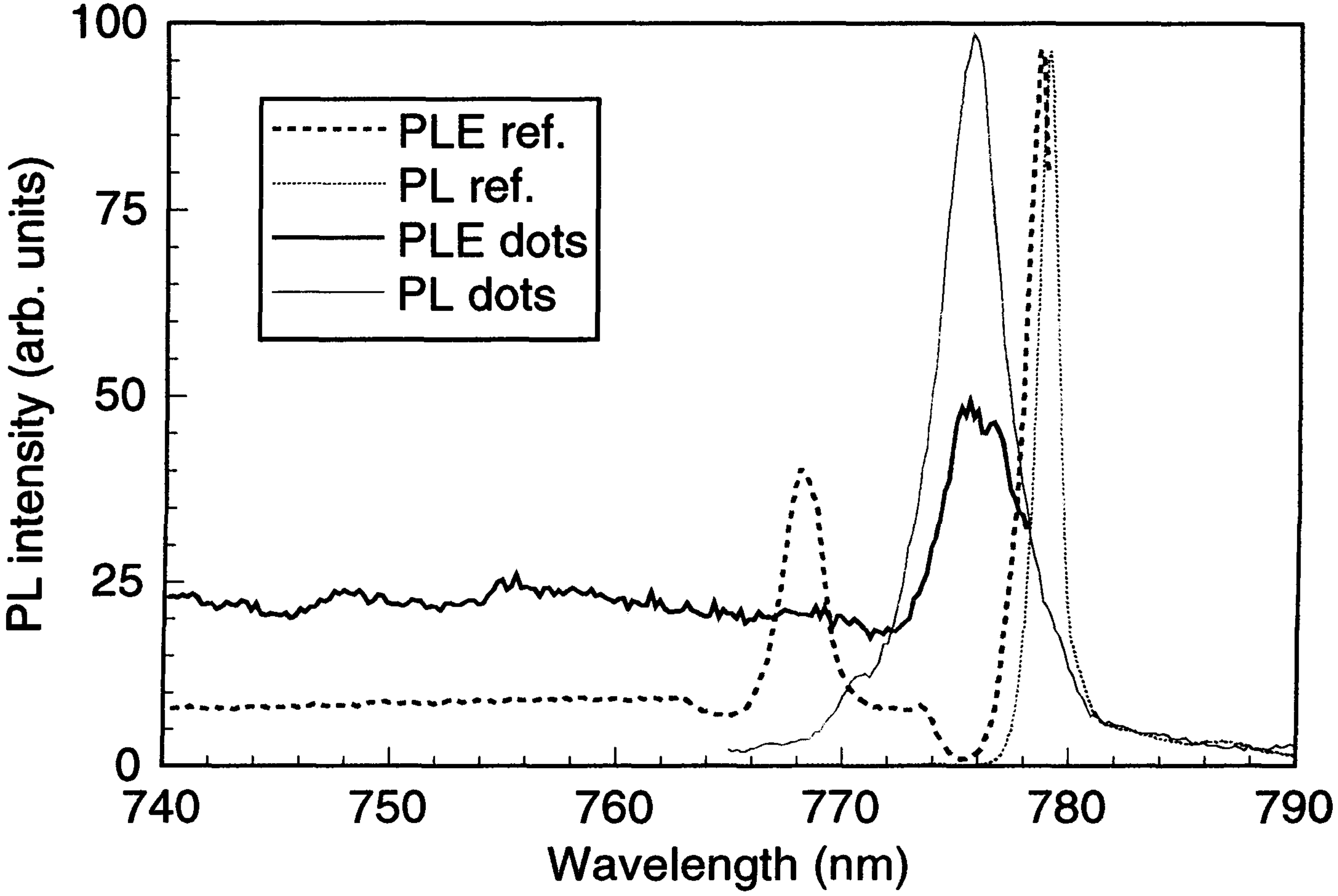


33 nm

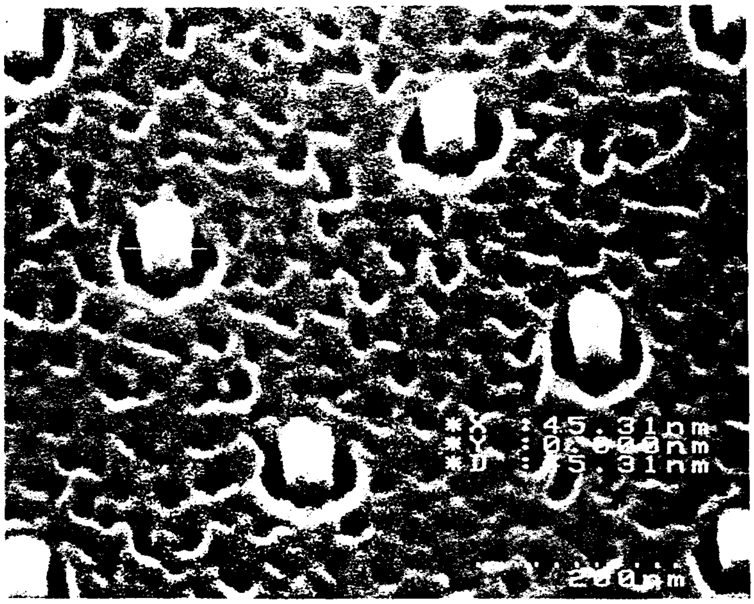


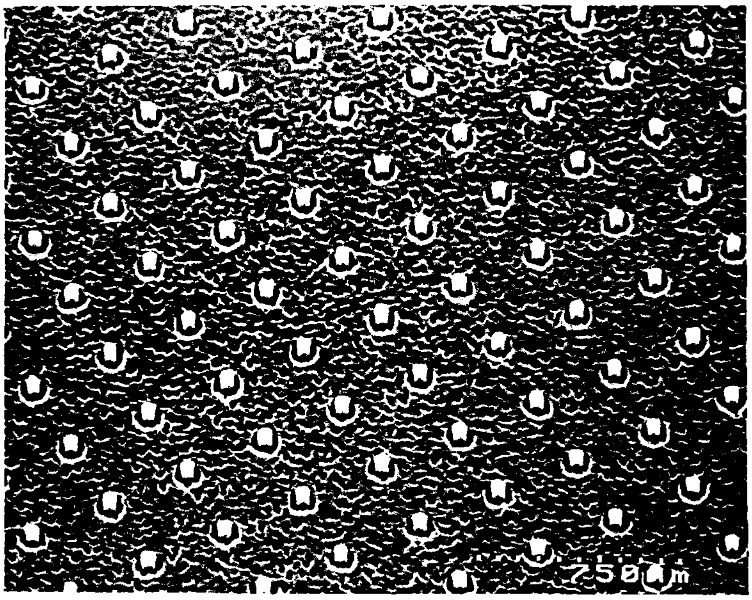


Dot diameter = 33 nm

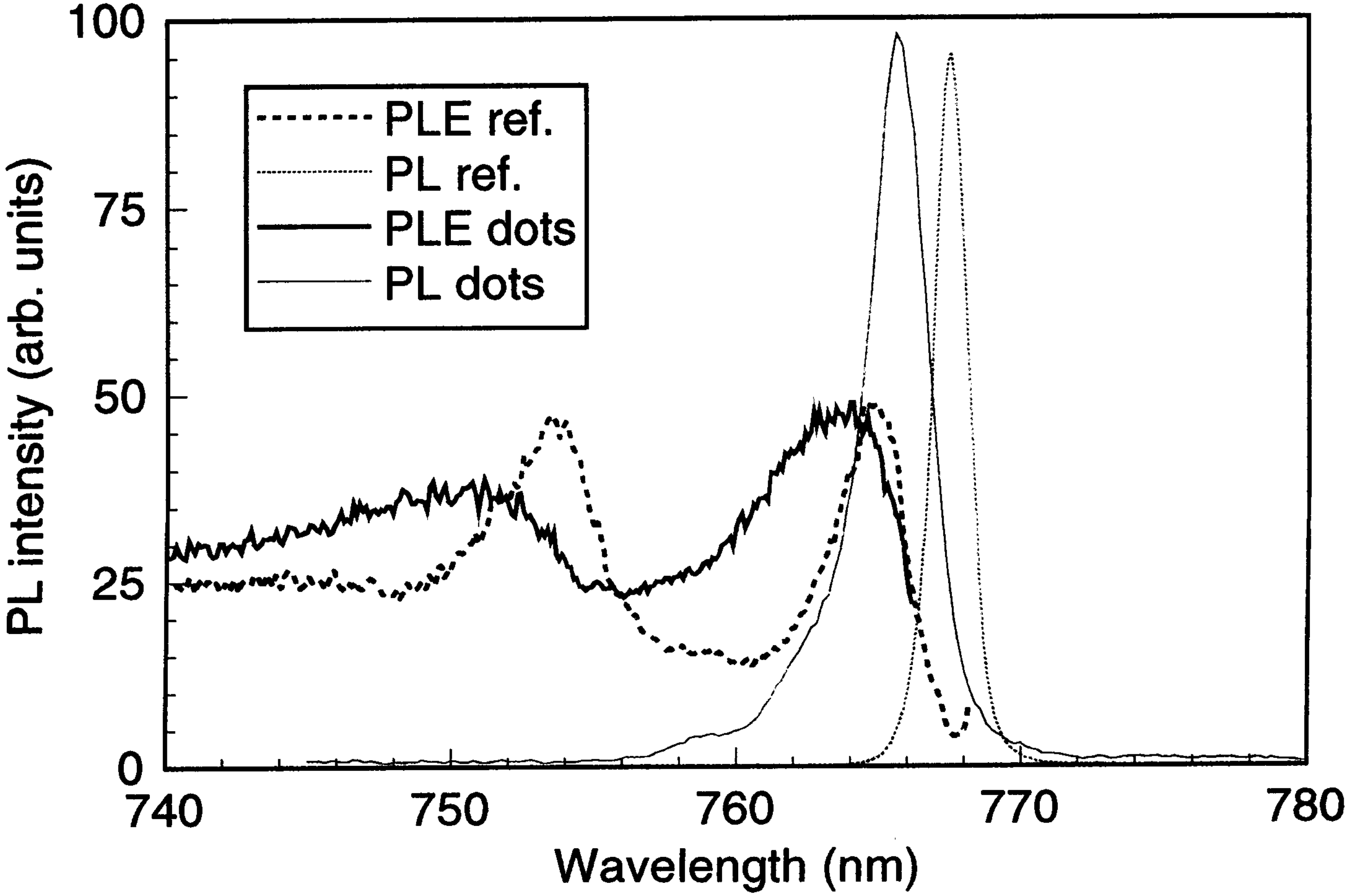


45 nm

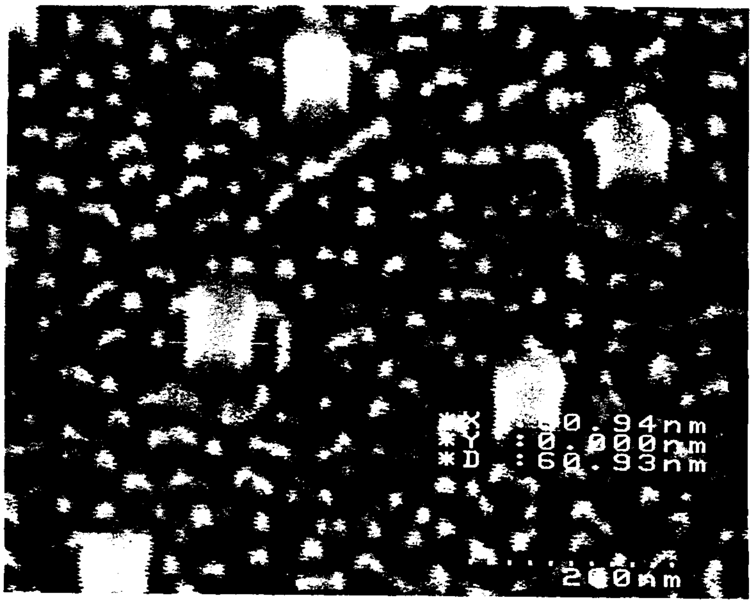


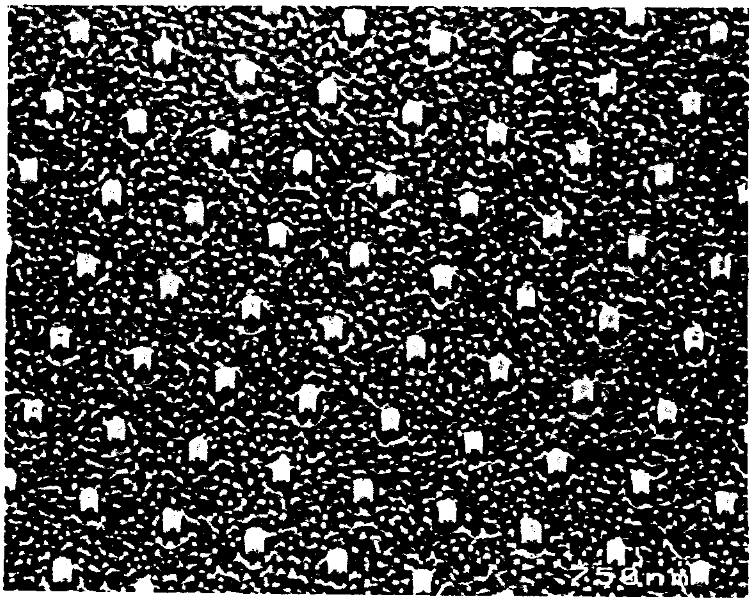


[redacted]

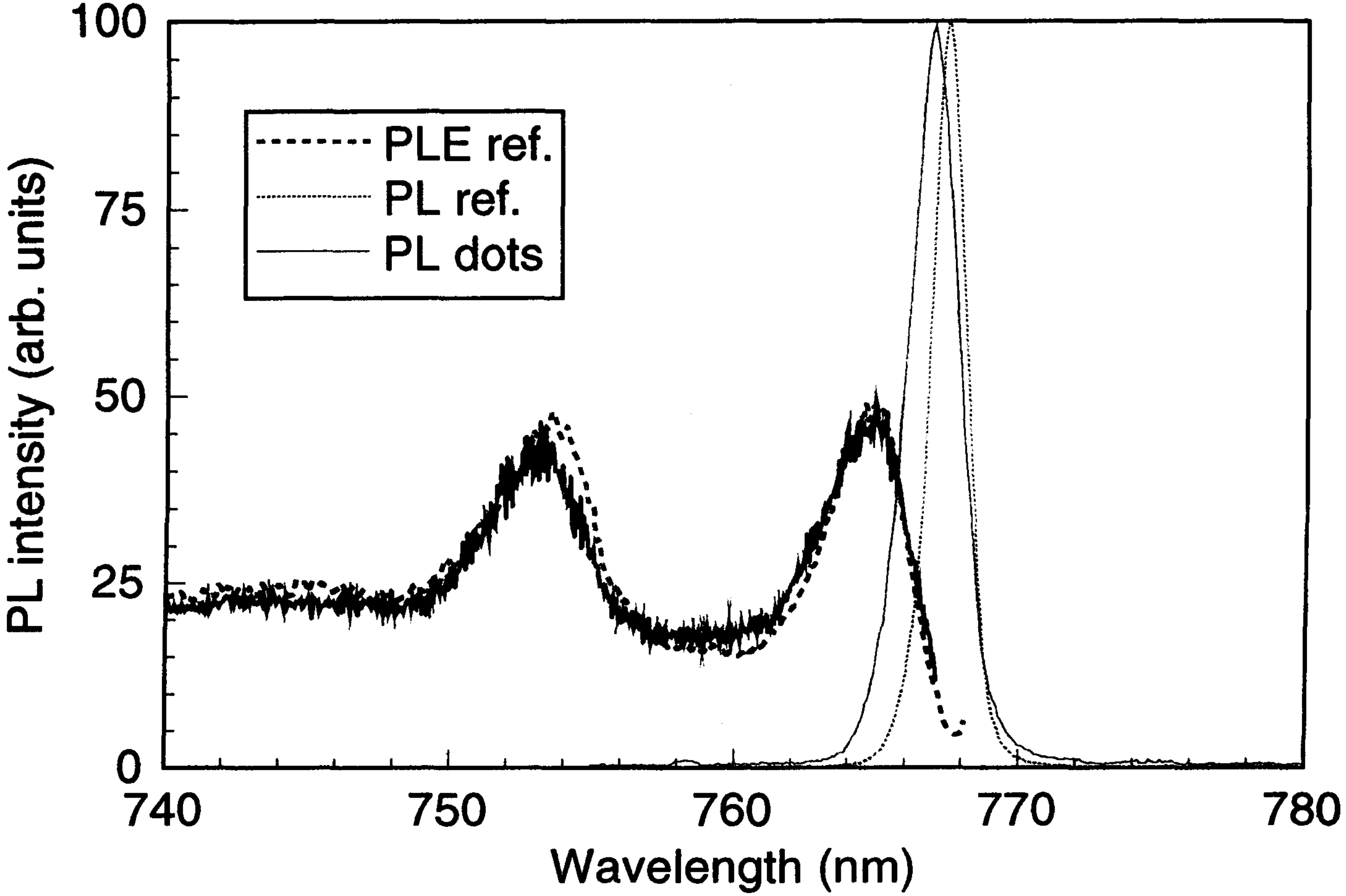

PLE ref.
PL ref.
PLE dots
PL dots
100
75
50
25
0
PL intensity (arb. units)
740
750
760
770
780
Wavelength (nm)

60 nm





[redacted]


PLE ref.
PL ref.
PL dots
PL intensity (arb. units)
0
25
50
75
100
740
750
760
770
780
Wavelength (nm)