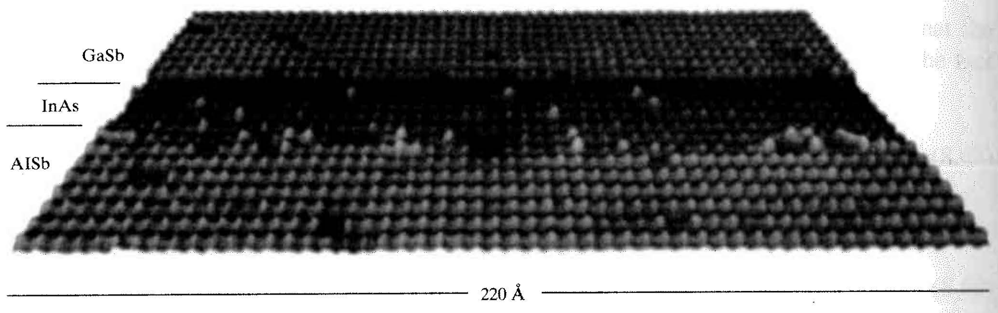
Chapter 9

MODELS OF SEMICONDUCTOR QUANTUM WELLS, QUANTUM WIRES, AND QUANTUM DOTS



GaSb-InAs-AlSb heterostructure. (From Harper, J., et al., "Microstructure of GaSb-on-InAs Heterojunction Examined with Cross-Sectional Scanning Tunneling Microscopy," *Appl. Phys. Lett.* 73 (1998): 2805.)

In Section 4.7, the idea of a quantum well, a quantum wire, and a quantum dot was introduced using a simple particle-in-a-box model. The main differentiating characteristic among the three structures was the size of the structure in each coordinate, with respect to a particle's de Broglie wavelength at the Fermi energy (i.e., with respect to λ_F). Assuming the electron E–k relation

$$E = \frac{\hbar k^2}{2m_e^*}, \tag{9.1}$$

which, for instance, typically holds near the bottom of the conduction band in a semiconductor, we have

$$\lambda_F = \frac{h}{\sqrt{2m_e^* E_F}}. \tag{9.2}$$

If the electron's environment is large compared to λ_F, then the electron will behave approximately as if it is free. If the electron's wavelength is on the order of, or is large compared with, its environment, then it will behave in a confined fashion.[†]

To appreciate the size scales involved, consider that in three dimensions the Fermi wavelength is, using (8.36),

$$\lambda_F = \frac{2\pi}{\left(3N_e\pi^2\right)^{\frac{1}{3}}}. \tag{9.3}$$

For copper, $N_e \simeq 8.45 \times 10^{28}$ /m^3, such that $\lambda_F = 0.46$ nm ($\lambda_F = 0.52$ nm for gold), whereas for GaAs[‡] assuming a doping level such that $N_e \simeq 10^{22}$ m^{-3}, $\lambda_F = 94$ nm. More generally, in typical metals,

$$\lambda_F \sim 0.5 - 1 \text{ nm}, \tag{9.4}$$

and in typical semiconductors,

$$\lambda_F \sim 10 - 100 \text{ nm}, \tag{9.5}$$

although this value depends on the doping level. Thus, confinement effects tend to become important in semiconductors at much larger dimensions than for conductors. This can also be considered from the effective mass point of view, due to the small effective mass in semiconductors.

To summarize, the three structures are shown in Fig. 9.1, where

- if $\lambda_F \ll L_x, L_y, L_z$, then we have an effectively three-dimensional system—the system, in all directions, is large compared to the size scale of an electron;
- if $L_x \leq \lambda_F \ll L_y, L_z$, then we have an effectively two-dimensional system—a two-dimensional electron gas or quantum well;
- if $L_x, L_y \leq \lambda_F \ll L_z$, then we have an effectively one-dimensional system—a quantum wire;
- if $L_x, L_y, L_z \leq \lambda_F$, then we have an effectively zero-dimensional system—a quantum dot.

In the previous discussion, we simply considered the structures to be empty space bounded by hard walls. However, many important structures are made from semiconducting materials, and, in any event, it is often crucial to take into account the material properties of the object. In this chapter, we reconsider quantum wells, wires, and dots, taking into account material properties of the structure.

[†] Coherence length, which will be described in Section 10.2, is another factor in determining system dimensionality. Furthermore, for quantum dots, often the exciton radius is taken as the measure of quantum confinement, as described in Section 9.3.1.

[‡] For semiconductors, the use of (8.36) in (9.2) corresponds to approximating the material as a bulk electron gas having electron density equivalent to the doping density. See also Problem 9.2 for another method, as well as the discussion in Section 9.3.1.

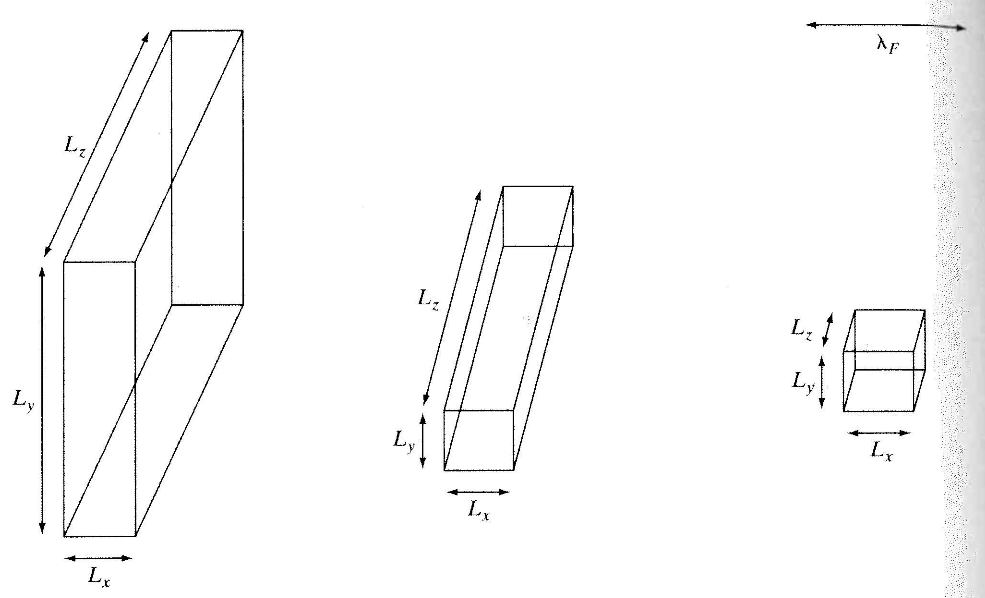


Figure 9.1 Effectively two-, one-, and zero-dimensional regions of space, where it is assumed that a spatial dimension can be neglected when $L \leq \lambda_F$. The size of λ_F is depicted in the upper-right corner.

9.1 SEMICONDUCTOR HETEROSTRUCTURES AND QUANTUM WELLS

Crystal growth techniques can produce atomically abrupt interfaces between two materials (see, for example, the image at the start of this chapter). especially if the lattice types and lattice constants of the two materials are similar. For example, we could sandwich a small bandgap material, such as GaAs (perhaps tens of angstroms thick) between thick layers of a large bandgap material, such as AlGaAs, as shown in Fig. 9.2. Note that in this model, the AlGaAs is a semi-infinite half space, and the GaAs is an infinite planar slab having thickness L_x.

This sandwich is called a *semiconductor heterostructure*. Quantum wells are formed in both the conduction and valence bands, as shown in the real-space energy band diagram in Fig. 9.3.

To a good degree of accuracy, semiconductor heterostructures composed of crystalline materials can be analyzed by replacing the actual crystal structure with a potential energy profile, and using the value of effective mass appropriate to the material. This approximation

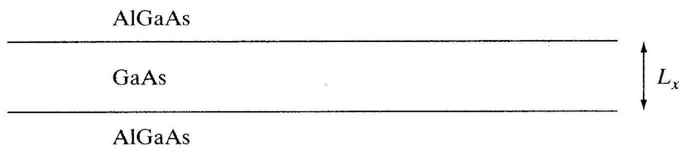


Figure 9.2 Semiconductor heterostructure composed of AlGaAs (large bandgap material) and GaAs (small bandgap material).

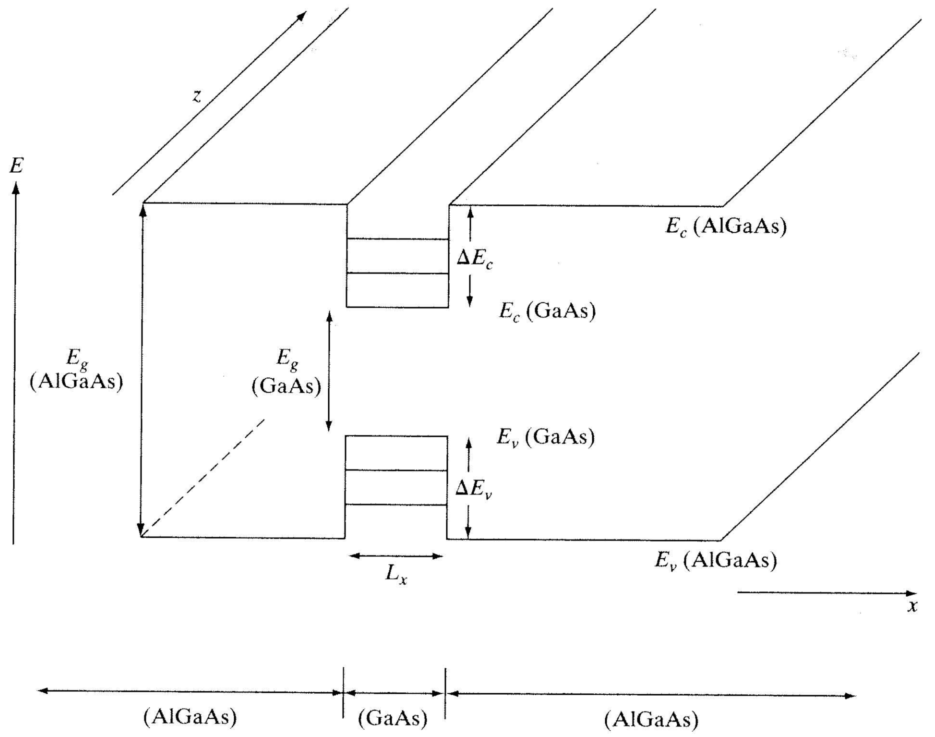


Figure 9.3 Semiconductor heterostructure made from AlGaAs and GaAs. Quantum wells form in both the conduction and valence bands, resulting in subband energy levels.

presupposes that the Bloch functions (5.7) in each region are the same, or similar. This is often a reasonable approximation for typical III-V heterostructures, i.e., those constructed from elements in groups III and V of the periodic table, and II-VI heterostructures are also of interest.

We use the effective mass Schrödinger's equation (5.41),

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 + V(x)\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{9.6}$$

where m^* is the effective mass of the particle at a given point in the heterostructure, so that $m^* = m^*(x)$. Since the heterostructure is comprised of piece-wise constant regions, we can solve (9.6) in the usual way. That is, we solve (9.6) in each region, and match the solutions across each interface using the continuity conditions of Schrödinger's equation (3.143).

The effective mass Schrödinger's equation actually does not strictly hold for alloys such as AlGaAs, which do not have pure periodic (crystalline) structure. However, it often suffices to treat the alloy as having a periodic nature. For instance, AlGaAs is modeled as having properties between AlAs and GaAs. This is called the *virtual-crystal approximation*. The bandgap of the alloy is approximately the arithmetic average of the bandgaps of the two materials. For AlAs, $E_g \simeq 2.2$ eV, and for GaAs, $E_g \simeq 1.5$ eV, and so AlGaAs has a bandgap on the order of $E_g \simeq 1.85$ eV. More specifically, AlGaAs is an alloy, written as $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$, where α is the mole fraction of Al. The band gap can be empirically approximated as

$$E_g \simeq 1.426 + 1.247\alpha \tag{9.7}$$

for $\alpha < 0.45$. (See Table VI in Appendix B for properties of AlGaAs). For this range of α, $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ is a direct bandgap material, whereas for $0.45 < \alpha < 1.0$, it is an indirect bandgap material ($\alpha = 0.3$ is a common value, resulting in $E_g \simeq 1.80$). Furthermore, AlAs and GaAs have similar lattice constants (Table V in Appendix B), and thus can be combined without significant lattice strain.

Using the separation of variables technique presented in Section 3.2, we solve (9.6) by writing

$$\psi(x, y, z) = \psi_x(x)\,\psi_y(y)\,\psi_z(z), \tag{9.8}$$

which results in

$$\left(\frac{1}{\psi_x}\frac{\partial^2}{\partial x^2}\psi_x + \frac{1}{\psi_y}\frac{\partial^2}{\partial y^2}\psi_y + \frac{1}{\psi_z}\frac{\partial^2}{\partial z^2}\psi_z + \frac{2m^*}{\hbar^2}(E - V(z))\right) = 0. \tag{9.9}$$

Using the separation argument, we obtain

$$\frac{1}{\psi_y}\frac{\partial^2}{\partial y^2}\psi_y = -k_y^2 \to \psi_y(y) = Ce^{ik_y y} + De^{-ik_y y}, \tag{9.10}$$

$$\frac{1}{\psi_z}\frac{\partial^2}{\partial z^2}\psi_z = -k_z^2 \to \psi_z(z) = Ae^{ik_z z} + Be^{-ik_z z}, \tag{9.11}$$

leading to

$$\left(-k_y^2 - k_z^2 + \frac{1}{\psi_x}\frac{\partial^2}{\partial x^2}\psi_x + \frac{2m^*}{\hbar^2}(E - V(x))\right) = 0. \tag{9.12}$$

Therefore, in the unconstrained directions (y and z), the wavefunction is represented by plane waves. Since k_y and k_z can be allowed to take on positive or negative values, for our purposes it is sufficient to take

$$\psi_y(y)\,\psi_z(z) = Ae^{ik_y y}e^{ik_z z}. \tag{9.13}$$

It remains to solve†

$$\left(-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2}{2m^*}\left(k_y^2 + k_z^2\right) + V(x)\right)\psi_x(x) = E\psi_x(x). \tag{9.14}$$

---

† Another approach is to assume plane waves in the unconfined coordinates, in accordance with our experiences with free (unconfined) electrons, such that,

$$\psi(x, y, z) = e^{ik_y y}e^{ik_z z}\psi_x(x),$$

which, when substituted into (9.6), leads to the same result.

In what follows, we will want to make use of the various one-dimensional problems considered in Sections 4.1, 4.3, and 4.5. However, there we solved the one-dimensional Schrödinger's equation (3.140),

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x). \tag{9.15}$$

This equation and (9.14) will have the same form, and, thus, the solutions of (9.15) can be used as the solutions of (9.14), if we define an effective potential

$$V_e(x, k) = \frac{\hbar^2}{2m^*(x)}\left(k_y^2 + k_z^2\right) + V(z), \tag{9.16}$$

such that (9.14) becomes

$$\left(-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + V_e(x, k)\right)\psi_x(x) = E\psi_x(x). \tag{9.17}$$

The effective potential depends on the confining potential $V(x)$, and also on the longitudinal wavenumber $k_l = \sqrt{k_y^2 + k_z^2}$. Furthermore, the effective mass will depend on position x as well.

This is as far as we can go without considering a specific model of the heterostructure, which we will do in the next section. In summary, we have

$$\psi(x, y, z) = Ae^{ik_y y}e^{ik_z z}\psi_x(x), \tag{9.18}$$

where k_y and k_z are continuous variables (as in the free-electron case), and ψ_x (and k_x) will be determined by solving (9.17) subject to boundary conditions in the x-coordinate.

9.1.1 Confinement Models and Two-Dimensional Electron Gas

To continue with our analysis of the heterostructure we need to determine the wavefunction ψ_x, and wavenumber k_x. The effective potential V_e depends on the longitudinal wavenumber k_l, and serves to modify the potential seen by the electron from $V(x)$ to $V_e(x, k)$.

The simplest approximation is to assume that the well (here referring to the well in either the conduction or valence band) is infinitely deep, that is,

$$V = 0, \quad 0 \le x \le L_x, \tag{9.19}$$
$$V = \infty, \quad x < 0,\ x > L_x,$$

which is the hard-wall model considered in Section 4.3.2. Schrödinger's equation inside the well is

$$\left(-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2 k_l^2}{2m^*}\right)\psi_x(x) = E\psi_x(x), \tag{9.20}$$

with k_l being the longitudinal wavenumber and m^* the effective mass in the well region (i.e., the GaAs region in Figs. 9.2 and 9.3). The solution of (9.20) is

$$\psi_x(x) = Fe^{ik_x x} + Ge^{-ik_x x}, \tag{9.21}$$

where

$$k_x^2 = \frac{2m^*}{\hbar^2}E - k_l^2. \tag{9.22}$$

Applying the boundary conditions

$$\psi_x(0) = \psi_x(L_x) = 0 \tag{9.23}$$

to (9.21), we obtain, as found in Section 4.3.1,

$$k_x = k_{x,n} = \frac{n\pi}{L_x}, \tag{9.24}$$

$n = 1, 2, 3, \ldots$, so that

$$\psi(x) = \left(\frac{2}{L_x}\right)^{1/2} \sin\frac{n\pi}{L_x}x. \tag{9.25}$$

From (9.18),

$$\psi(x, y, z) = A\left(\frac{2}{L_x}\right)^{1/2} \sin\left(\frac{n\pi}{L_x}x\right) e^{ik_y y}e^{ik_z z}. \tag{9.26}$$

Energy is obtained from (9.22) as

$$E = \frac{\hbar^2}{2m^*}\left(\left(\frac{n\pi}{L_x}\right)^2 + k_y^2 + k_z^2\right) \tag{9.27}$$

$$= E_n + \frac{\hbar^2}{2m^*}\left(k_y^2 + k_z^2\right), \tag{9.28}$$

where

$$E_n = \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{L_x}\right)^2. \tag{9.29}$$

The given energy levels are the subbands previously obtained in Section 4.7 (i.e., E_n is the bottom of the nth subband), and depicted as the quantized energy levels in Fig. 9.3. The resulting collections of electrons confined to the well form a two-dimensional electron gas; the structure is effectively two-dimensional in the sense that electrons can only move freely in two dimensions (the x- and y-directions). Two-dimensional electron gases formed at semiconductor heterojunctions typically have thickness values on the scale of a few nanometers, resulting in relatively large, well-separated subband energies. The energy–wavevector dispersion diagram is shown in Fig. 9.4 (a).

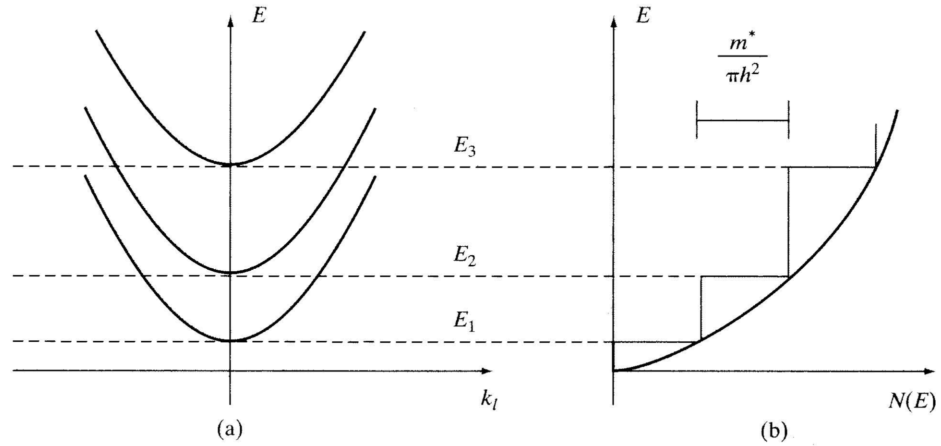


Figure 9.4 (a) Energy versus continuous wavevector $k_l = \sqrt{k_y^2 + k_z^2}$, showing three subbands. (b) Energy versus two-dimensional density of states. For comparison, the continuous parabolic curve is the density of states for three-dimensional free electrons.

Earlier, we found that the density of states in two dimensions is constant with respect to energy,

$$N(E) = \frac{m^*}{\pi\hbar^2}. \tag{9.30}$$

In the present case, however, associated with each subband is a density of states. The total density of states for any energy is the sum of all subband density of states at or below that energy, leading to the density of states depicted in Fig. 9.4(b), and given by

$$N(E) = \frac{m^*}{\pi\hbar^2}\sum_j H(E - E_j), \tag{9.31}$$

where H is the Heaviside function,

$$H(E) = \begin{cases} 1, & E > 0, \\ 0, & E < 0. \end{cases} \tag{9.32}$$

Absorption in an $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As/GaAs}$ quantum well structure is shown in Fig. 9.5, where the effects of the step-like two-dimensional density of states is evident.

If the Fermi energy is a bit greater than the jth subband energy, $E_F > E_j$, then approximately j two-dimensional subbands will be filled. More specifically, from (8.49), the density of electrons in the jth two-dimensional subband will be given by

$$n_j = \int_{E_j}^{\infty} N(E)\, f(E, E_F, T)\, dE \tag{9.33}$$

$$= \frac{m^*}{\pi\hbar^2}\int_{E_j}^{\infty} f(E, E_F, T)\, dE = \frac{m^* k_B T}{\pi\hbar^2}\ln\left(1 + e^{\left(\frac{E_F - E_j}{k_B T}\right)}\right), \tag{9.34}$$

and the total density of electrons per unit area is $n = \sum_j n_j$.

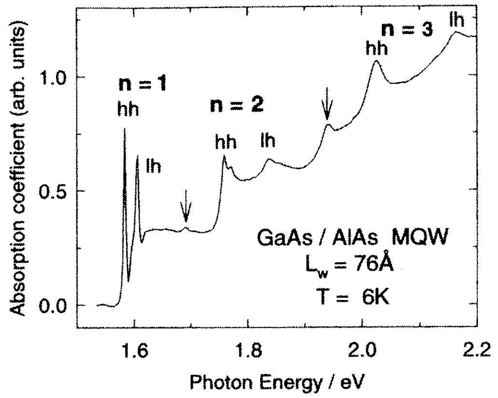


Figure 9.5 Absorption in an AlGaAs/GaAs multiple quantum well structure (40 wells, each of width 7.6 nm), showing clearly the effects of the step-like two-dimensional density of states. The relatively sharp peaks at the onset of absorption are due to excitons (hh and lh indicate heavy holes and light holes, respectively). (Based on a figure from N. Mayhew, D. Phil. thesis, Oxford University, 1993. (Published in A.M. Fox, "Optoelectronics in Quantum Well Structures," *Contemporary Physics* 37, 111–125 (1996) (http://www.tandf.co.uk/journals)). Reprinted with permission from A.M. Fox.)

It should be noted that for the subband model to be appropriate, the thermal energy $k_B T$ should be much less than the difference between energy subbands; otherwise, the discrete nature of the subbands will be obscured. Therefore, we must have

$$\frac{\hbar^2}{2m^*}\left(\frac{\pi}{L_x}\right)^2 \gg k_B T. \tag{9.35}$$

Considering GaAs with $m^* = 0.067 m_e$ for electrons at room temperature, we find that (9.35) yields $L_x \ll 15$ nm. At $T = 4$ K, $L_x \ll 128$ nm.

9.1.2 Energy Band Transitions in Quantum Wells

As with a bulk semiconductor, energy band transitions can occur between states in a quantum heterostructure. These transitions tend to govern the optical properties of the device, and can be engineered for desired properties. Since quantum wells are formed in both the valence and conduction bands, various transitions can be envisioned as described next.

Interband Transitions. Interband transitions are from the n^{th} state in the valence band well to the m^{th} state in the conduction band well (or vice versa). From Fig. 9.3, it is evident that these can occur for incident photon energies[†]

$$\hbar\omega = E_g + E_n + E_m \tag{9.36}$$

[†] Here we assume bandedge-to-bandedge direct transitions (i.e., at $k = 0$). See Section 5.4.5 for more details on other types of transitions.

$$= E_g + \frac{\hbar^2\pi^2}{2m_h^* L_x^2} n^2 + \frac{\hbar^2\pi^2}{2m_e^* L_x^2} m^2, \tag{9.37}$$

where E_g is the gap energy in the well, and E_n and E_m are the energies in the valance and conduction well subbands, respectively, measured from the band edges.

The shift (from the bulk value E_g) in the required photon energy for absorption is characteristic of quantum-confined structures, and the possible tuning of the shift by adjustment of the well thickness and well material is one advantage of confined structures over bulk materials. For transitions among the lowest states ($n = m = 1$) we have

$$\hbar\omega = E_g + \frac{\hbar^2}{2}\left(\frac{\pi}{L_x}\right)^2\left(\frac{1}{m_h^*} + \frac{1}{m_h^*}\right) \tag{9.38}$$

$$= E_g + \frac{\hbar^2\pi^2}{2m_r^* L_x^2}, \tag{9.39}$$

where m_r^* is the reduced mass, $m_r^{-1} = m_e^{*-1} + m_h^{*-1}$.

It should be noted that not all transitions are possible, and only certain pairs of n and m lead to permissible transitions. This is governed by Fermi's golden rule,[†] and leads to the idea of *selection rules* for deciding which transitions are permissible. The allowed transitions are dependent on the type of incident energy; for example, it turns out that for incident light polarized in the plane of the well (the z–y plane in Fig. 9.3), only transitions $n = m$ are allowed, although this rule is only strictly applicable to the infinite-wall model. These transitions are depicted in Fig. 9.6 for a $\text{Al}_\alpha\text{Ga}_{1-\alpha}\text{As/GaAs}$ system.

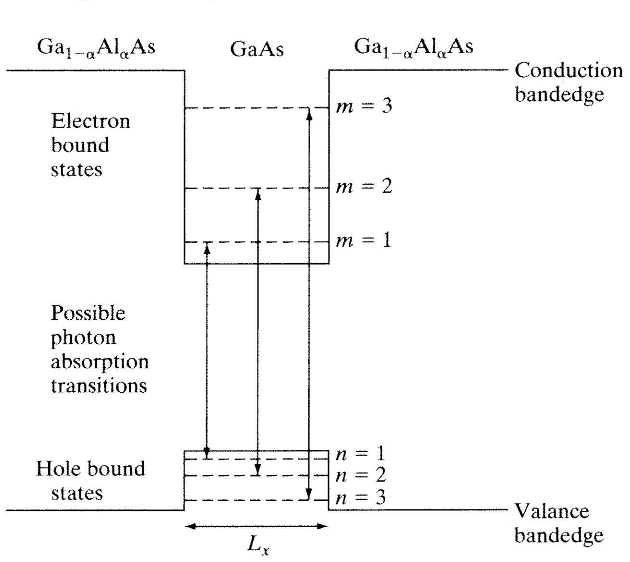


Figure 9.6 Allowed transitions in an infinite-well model of an $\text{Al}_\alpha\text{Ga}_{1-\alpha}\text{As/GaAs}$ quantum well structure for incident light polarized in the plane of the well.

[†] Fermi's golden rule is discussed in many solid state and quantum mechanics texts; a good introduction can be found in [17].

Excitonic Effects. Excitonic effects are more prominent in quantum well structures then in bulk materials. As described in Section 5.4.5, for bulk semiconductors, excitons (bound electron–hole pairs) are typically only important at very low temperatures. Room temperature thermal energy can easily overcome the exciton binding energy (5.55), which is on the order of a few meV. However, the situation changes significantly in quantum-confined structures. In quantum wells, the binding energy is enhanced since electrons and holes are forced to be closer together. This is obvious for a suitably small well (obviously the electron and hole cannot be separated by 14 nm, the bulk GaAs exciton radius as given in Section 5.4.5, in a 5 nm well!), although the effect of the finite size of the structure is felt by the exciton even if the well size is bigger than the exciton radius. In fact, as described later for quantum dots, the relative size between the exciton radius and the well width can be used to gauge the importance of confinement effects. The closer spacing between the electron and hole in a confined exciton results in quantum well excitons being generally stable at room temperatures. Excitonic effects are seen in Fig. 9.5 as the relatively sharp peaks at the onset of absorption.

Intersubband Transitions. Aside from transitions from the valence band well to the conduction band well, transitions between subbands in each of these wells can occur. These are known as *intersubband transitions*, and obviously involve much lower energies than interband transitions since the transition energies are simply

$$\Delta E = E_\alpha - E_\beta = \frac{\hbar^2}{2m^*}\left(\frac{\pi}{L_x}\right)^2(\alpha^2 - \beta^2) \tag{9.40}$$

for the subbands indexed by α and β.

Intersubband transitions in the conduction band well are depicted in Fig. 9.7.

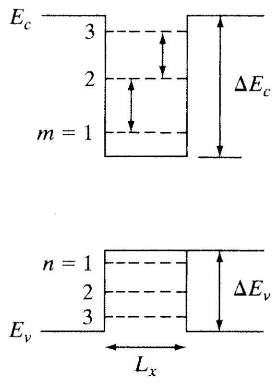


Figure 9.7 Depiction of intersubband transitions in the conduction band of a quantum well.

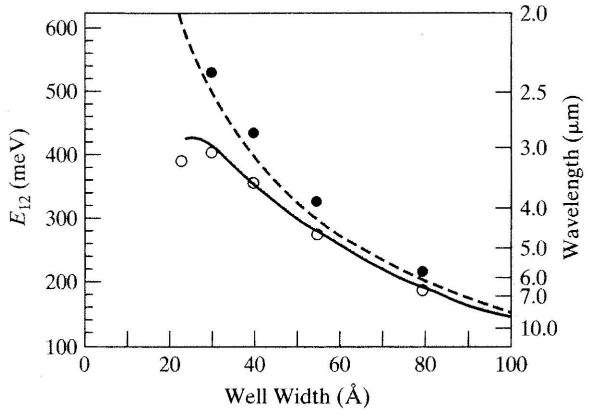


Figure 9.8 Measured quantum well width dependence on the lowest intersubband transition ($\alpha = 2$, $\beta = 1$) peak energies for $In_{0.5}Ga_{0.5}As/Al_{0.45}Ga_{0.55}As$ (hollow circles) and $In_{0.5}Ga_{0.5}As/AlAs$ (solid circles) quantum wells. Solid and dashed lines are calculated results for $Al_{0.45}Ga_{0.55}As$ (solid line) and AlAs (dashed line) quantum wells using an effective mass model. (Based on a figure from Chui, H. C., et al., "Short Wavelength Intersubband Transitions in In GaAs/AlGaAs Quantum Wells Grown on GaAs," *Appl. Phy. Lett.* 64 (1994): 736. © 1994, American Institute of Physics.)

Modeling the well as being infinitely high and considering light polarized in the direction of confinement[†] (the x-direction in Fig. 9.3), we find that selection rules dictate that only transitions corresponding to $(\alpha - \beta)$ being an odd number are allowed. Therefore, the lowest transition occurs for

$$\Delta E = E_2 - E_1 = \frac{3\hbar^2}{2m^*}\left(\frac{\pi}{L_x}\right)^2, \tag{9.41}$$

which for a 10 nm GaAs conduction band well is $\Delta E = 0.168$ eV. This corresponds to $f = 40.62$ THz, or $\lambda = 7.39$ μm, and, thus, intersubband transitions can be used for infrared detectors and emitters.

As an example, the width dependence of the peak energies of a quantum well for the lowest intersubband transition ($\alpha = 2$, $\beta = 1$) is shown in Fig. 9.8. The structures are $In_{0.5}Ga_{0.5}As/Al_{0.45}Ga_{0.55}As$ (hollow circles) and $In_{0.5}Ga_{0.5}As/AlAs$ (solid circles) quantum wells. Circles are measured results, and solid and dashed lines are calculated results for $Al_{0.45}Ga_{0.55}As$ (solid line) and AlAs (dashed line) quantum wells using an effective mass model.

Comments on the Model. The simple infinite well model is a fairly gross approximation to a realistic heterostructure, since the difference in energy bands in the different

---

[†] In quantum well structures such as $Al_\alpha Ga_{1-\alpha}As/GaAs$, intersubband transitions will essentially only be induced by light polarized in the direction of confinement, which necessitates features such as gratings to couple light into the well.

materials making up the heterostructure can be as small as a few tenths of an eV. For example, the bandgap for GaAs is 1.5 eV, and for AlGaAs E_g is on the order of 1.8 eV. Modeling this difference as being infinitely large is then not a very good approximation (although the infinite-well model gives an order-of-magnitude approximation; see Problem 4.17 and Problem 9.8). Furthermore, it can be appreciated that the result (9.28),

$$E = E_n + \frac{\hbar^2}{2m^*}k_l^2, \tag{9.42}$$

and, in particular, the existence of energy subbands, is quite general, and is independent of the specific confining potential in the x coordinate. For example, we could have solved (9.17) subject to a finite potential well as considered in Section 4.5, a parabolic well as considered in Section 4.5.2, or a triangular well as considered in Section 4.5.3. From an energy standpoint, the main difference is that in those cases, E_n would not be given by (9.29), but by the solution of (4.83) or (4.85) for the finite well case, by (4.98) for the parabolic well case, or (4.100) for the triangular well case. Other potential profiles can be considered, obviously. Thus, qualitatively, we expect our solution to be reasonable, but for good quantitative results, we would need to account for the specific confining potential.

However, if we assume a more realistic finite well model, a complication arises, since an electron's effective mass will be different in each different region (the well itself, and the exterior regions). We can take this into account by using the appropriate effective mass in the boundary conditions for the derivative of the wavefunction, as given in (3.143). This aspect was ignored in our previous analysis of finite wells in Chapter 4, although it is considered in Problems 9.7 and 9.8 in this chapter. For example, as discussed in Problem 9.7, if we assume a finite square well profile, accounting for effective mass results in equations for the energy values that are a bit more complicated than (4.83) and (4.85), but which can still be easily solved numerically in the same manner as the constant mass equations.

Yet another complication for the finite potential well model is that we need to solve (9.17) involving the effective potential,

$$V_e(x,k) = \frac{\hbar^2 k_l^2}{2m^*} + V(x). \tag{9.43}$$

For example, consider the two adjacent regions at either side of a finite square well. On one side, we have the large bandgap material, which we'll call the barrier, and on the other side, we have the smaller bandgap material of the well. We may be interested in either the conduction band, the valence band, or both bands. In the barrier region, we take $V(x) = E^b$, where E^b is the bandedge in either the conduction or valence band (i.e., E^b is either E_c^b or E_v^b), and in the well, the corresponding bandedge is $V(x) = E^w$. As shown in Fig. 9.3, the potential difference between the two materials is simply the difference in bandedges. (The absolute reference of each band is not important.) Therefore, in the barrier, Schrödinger's equation (9.17) is

$$\left(-\frac{\hbar^2}{2m_b^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2 k_l^2}{2m_b^*} + E^b\right)\psi_x(x) = E\psi_x(x), \tag{9.44}$$

and in the well,

$$\left(-\frac{\hbar^2}{2m_w^*}\frac{\partial^2}{\partial x^2} + \frac{\hbar^2 k_l^2}{2m_w^*} + E^w\right)\psi_x(x) = E\psi_x(x), \tag{9.45}$$

where $m_{b,w}^*$ is the appropriate effective mass in each region. Because $m_b^* \neq m_w^*$, in general, the actual potential difference seen by a particle is not simply $E^b - E^w$ (although not described at that time, the potential V_0 used in Section 4.5 represents this difference in bandedges), but is

$$V_0 = E^b - E^w + \frac{\hbar^2 k_l^2}{2}\left(\frac{1}{m_b^*} - \frac{1}{m_w^*}\right).$$

Therefore, the solution actually depends on the value of the longitudinal wavenumber k_l. This serves to perturb the results from the k-independent potential V_0 considered in Section 4.5. However, for any given value of k_l, the solution can still be obtained.

Degenerately Doped Heterojunctions. Earlier, we considered forming a two-dimensional electron gas by sandwiching a very thin layer of a smaller bandgap material (such as GaAs) between thick layers of a larger bandgap material (such as AlGaAs, more precisely denoted by $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$). There is another method of forming a 2DEG that is of considerable interest, illustrated in Fig. 9.9.

In this method, the GaAs/$\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ heterojunction is formed by n-type heavily doped (degenerate) $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$, grown adjacent to *undoped* GaAs, typically using molecular-beam epitaxy. Silicon is often used as the dopant, forming donor atoms in the $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ material. A small amount of energy can liberate electrons from their donors, and these will be attracted to the GaAs material by the smaller bandgap. However, these electrons will still be attracted to their ionized donors in the $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ material, and thus will form a thin layer of electrons at the interface.

Furthermore, when there is an undoped layer between a heavily doped large bandgap material and a small bandgap material, a 2DEG forms in the undoped layer. This is known as *modulation doping*. Because the donors are separated from the electron gas layer, scattering from the donor atoms is avoided, resulting in very high mobility values.[†] Contributing to the high mobility is the good lattice match between AlGaAs and GaAs (Table V in

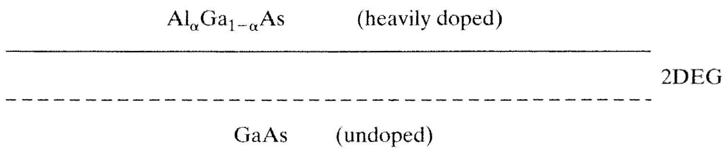


Figure 9.9 Degenerately doped heterojunction of $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ and GaAs.

[†] Mobility is discussed on page 320 , although, essentially, large (small) mobility means that electrons travel a relativity long (short) distance before colliding with the lattice or impurities.

Appendix B), forming an interface without many defects or much strain. For example, intrinsic GaAs at room temperature has an electron mobility of approximately $\mu_e = 8{,}600$ cm^2/V-s, whereas for intrinsic Si, $\mu_e = 1{,}350$ cm^2/V-s, and for copper, $\mu_e = 43.5$ cm^2/V-s. With a doping concentration of $N_D = 10^{18}$ cm^{-3}, still at room temperature, $\mu_e = 3{,}000$ cm^2/V-s for GaAs and $\mu_e = 300$ cm^2/V-s for Si. The decrease in mobility for the doped materials is due to scattering from the dopant atoms. In contrast, in modulation-doped AlGaAs/GaAs heterostructures, one can achieve $\mu_e = 7{,}000$ cm^2/V-s at room temperatures, and as high as $\mu_e = 10^6$–10^7 cm^2/V-s at low temperatures. Thus, modulation doping can offer a significant advantage over traditional doping methods. In fact, modulation doping is the basis for the *high electron mobility transistor* (HEMT), which is a common commercial structure. Other quantum well/heterostructure applications include resonant tunneling diodes, quantum well lasers, quantum well photodetectors, and optical modulators.

9.2 QUANTUM WIRES AND NANOWIRES

The preceding analysis assumed that electrons were confined along one coordinate, x, where the confinement was provided by the difference in bandgaps between different materials. With this assumption, we developed the concept of a two-dimensional electron gas and of energy subbands. The next logical step is to consider what happens if we confine electrons in a second direction, say along the y coordinate, as depicted in Fig. 9.10. The resulting (effectively one-dimensional) structure is called a quantum wire, as discussed previously in Section 4.7.2.

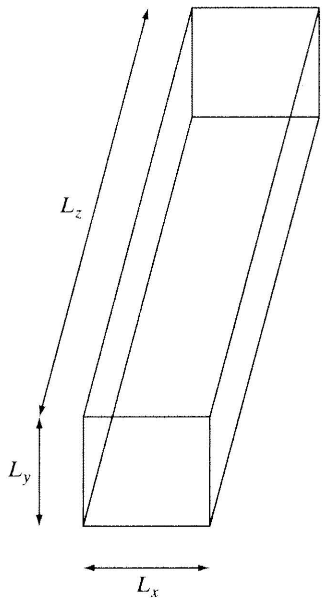


Figure 9.10 Quantum wire; $L_{x,y} \leq \lambda_F$, $L_z \gg \lambda_F$.

To analyze the quantum wire, we start with Schrödinger's equation,

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 + V(x,y)\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{9.46}$$

and express the wavefunction as

$$\psi(\mathbf{r}) = e^{ik_z z}\psi(x,y). \tag{9.47}$$

That is, the electron will be free to move along the z coordinate, i.e., along the wire, but will be confined in other directions. Substituting (9.47) into (9.46), we obtain

$$\left(-\frac{\hbar^2}{2m^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \frac{\hbar^2 k_z^2}{2m^*} + V(x,y)\right)\psi(x,y) = E\psi(x,y). \tag{9.48}$$

Given a certain potential $V(x,y)$, we can solve (9.48) analytically if V has a particularly simple form, or numerically, or using some approximate method. Although the details may be complicated for general confining barriers, given our experience in the previous section, we expect to obtain an energy relation analogous to (9.42), having the form

$$E = E_{n_x,n_y} + \frac{\hbar^2}{2m^*}k_z^2, \tag{9.49}$$

where k_z is now the longitudinal wavenumber. This will be a continuous parameter since electrons are free along the z coordinate. The discrete indices n_x and n_y correspond to subband indices. The subband energy E_{n_x,n_y} will be given by

$$E_{n_x,n_y} = \frac{\hbar^2}{2m^*}\left(k_{x,n_x}^2 + k_{y,n_y}^2\right), \tag{9.50}$$

where k_{x,n_x} and k_{y,n_y} will be discrete, and will depend on the specific form of the potential $V(x,y)$. For example, for an infinite confining potential (i.e., a "hard" barrier) at $x = 0, L_x$ and $y = 0, L_y$, then

$$k_{x,n_x} = \frac{n_x\pi}{L_x}, \qquad k_{y,n_y} = \frac{n_y\pi}{L_y}, \tag{9.51}$$

with $n_x, n_y = 1, 2, 3, \ldots$, such that

$$E_{n_x,n_y} = \frac{\hbar^2\pi^2}{2m^*}\left(\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2\right). \tag{9.52}$$

We previously determined the density of states in one dimension to be (8.15),

$$N(E) = \frac{1}{\pi\hbar}\left(\frac{2m^*}{E - V_0}\right)^{1/2} \tag{9.53}$$

for $E > V_0$. As with the two-dimensional electron gas (quantum well case), for the one-dimensional quantum wire, each subband will have a density of states given by the one-dimensional result (9.53),

$$N_{n_x,n_y}(E) = \frac{1}{\pi\hbar}\left(\frac{2m^*}{E - E_{n_x,n_y}}\right)^{1/2}. \tag{9.54}$$

If the energy level E is low, such that only the first subband is filled, the preceding density of states holds, and the system is one dimensional. As energy increases and more subbands are filled, the system is quasi-one dimensional. In this case the density of states is found by summing over all subbands, resulting in

$$N(E) = \frac{1}{\pi\hbar}\sum_{n_x,n_y}\left(\frac{2m^*}{E - E_{n_x,n_y}}\right)^{1/2} H\left(E - E_{n_x,n_y}\right),$$

where H is the Heaviside function (9.32),

This quasi-one dimensional density of states is depicted in Fig. 9.11. The discontinuities in the density of states are known as *van Hove singularities*.

As a concrete example, the density of states for two different single-wall carbon nanotubes is given in Fig. 9.12. As discussed in Section 5.5 and Problem 5.24, the (9, 0) zigzag tube is metallic, and has a finite density of states at the Fermi level $E_F = 0$. The (10, 0) tube is semiconducting, and the density of states is zero at the Fermi level.

Transitions and Excitonic Effects. Optical transitions in quantum wires arise from allowed transitions (governed by appropriate selection rules) between energy bands of the wire, and can be associated with the locations of van Hove singularities. As with quantum well structures, excitonic effects are expected to be very important in quantum wires even at room temperature, due to larger binding energies associated with confinement. For example, although binding energies for bulk semiconductors are typically a very small fraction of the bandgap energy (and thus constitute a relatively small perturbation in optical properties (Problem 5.20)), exciton binding energies in semiconducting quantum wires can be a large fraction of the bandgap. In SWNTs, the binding energy can be close to half of the bandgap,[†] significantly altering optical properties. Although there is still some debate about the relative importance between excitonic and van Hove singularity effects,

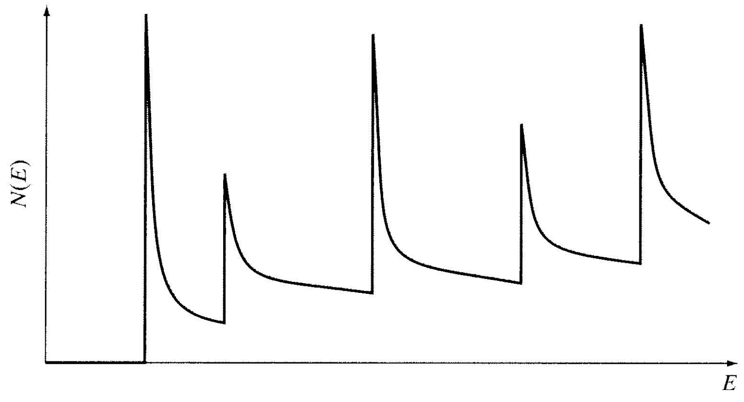


Figure 9.11 One-dimensional density of states.

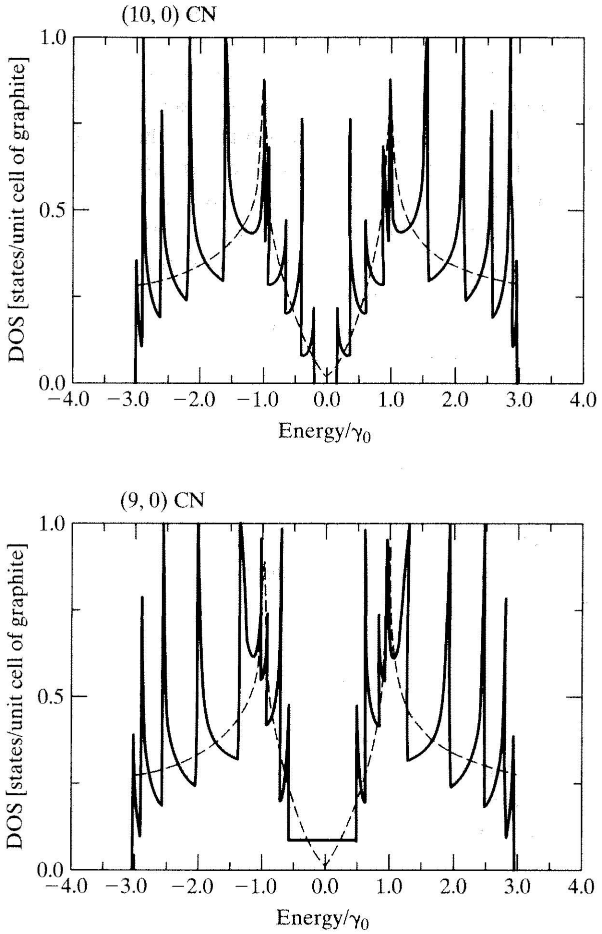


Figure 9.12 Density of states per unit cell for a (10, 0) (top) and (9, 0) (bottom) single-wall carbon nanotube. The (10, 0) tube is semiconducting, and the (9, 0) tube is metallic. $\gamma_0 = 2.5$ eV. The dashed lines are the density of states for graphene. (Based on a figure from Saito, R., et al., "Electronic Structure of Chiral Graphene Tubules," *Appl. Phys. Lett.* 60 (1992): 2204. © 1992, American Institute of Physics.)

[†] See, e.g., [21] and [22].

it seems clear that, especially for semiconducting nanotubes, excitonic effects are very important.

Last, it should be noted that for the quantum wire to act in a truly one-dimensional manner, the width (or radius for a round cross-section wire) should be on the order of the Fermi wavelength, or smaller. For the situation where the wire is thin but relatively wide, electrical transport will typically involve electrons in a single subband in the coordinate normal to the interface (i.e., in the "thin" direction, where energy level spacing is large), yet multiple subbands in the lateral (wide) direction due to more closely spaced energy levels. More generally, a wire that has radius values on the order of nanometers or tens of nanometers, but much larger than λ_F, might be called a *nanowire*, rather than a quantum wire. Quantum wires and nanowires will be further discussed in Chapter 10, where the concept of ballistic transport is introduced.

9.3 QUANTUM DOTS AND NANOPARTICLES

If we confine electrons in all three coordinates, forming a quantum dot (Section 4.7.3), electrons do not have a plane wave dependence in any direction. The three-dimensional Schrödinger's equation

$$\left(-\frac{\hbar^2}{2m^*}\nabla^2 + V(x, y, z)\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{9.55}$$

must be solved, and, not unexpectedly, the resulting energy will be fully quantized. For example, for the three-dimensional infinite potential well considered in Section 4.3.2, energy levels were obtained as (4.53),

$$E = E_{n_x,n_y,n_z} = \frac{\hbar^2\pi^2}{2m^*}\left(\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2\right). \tag{9.56}$$

In general, regardless of the specific potential $V(\mathbf{r})$, energy will take the form

$$E = E_{n_x,n_y,n_z} = \frac{\hbar^2}{2m^*}\left(k_{x,n_x}^2 + k_{y,n_y}^2 + k_{z,n_z}^2\right), \tag{9.57}$$

where k_{x,n_y}, k_{y,n_y}, and k_{z,n_z} will be discrete, and must be found from the specific boundary conditions of the problem.

As described previously, the density of states for a quantum dot is a series of delta functions (8.18). This discrete density of states leads to several fundamental differences compared with higher (than zero) dimensional systems, which have at least a piece wise continuous density of states. (A comparison of the density of states in various dimensions is shown in Fig. 9.13.)

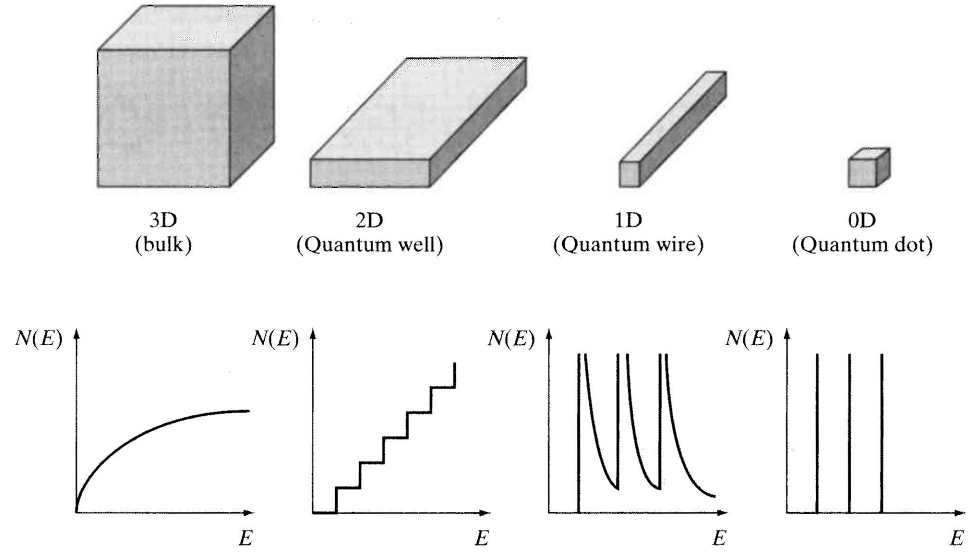


Figure 9.13 A comparison of the density of states in various dimensions.

For example, electron dynamics are obviously quite different in a quantum dot, since current cannot "flow." Thermal effects will be different in quantum dots than in bulk materials (or even quantum wells), since thermal energy can only excite electrons to a limited number of widely separated states. For this same reason, the frequency spectrum of emitted optical radiation from energy level transitions (called the *luminescence linewidth*) in quantum dots is very narrow even at relatively high temperatures. This makes quantum dots attractive in laser applications, as biological markers (as discussed next), and in a host of other optical applications.

9.3.1 Applications of Semiconducting Quantum Dots

Some applications of quantum dots in electronics are discussed in Chapter 7, in particular, their use as charge islands in Coulomb blockade circuits. Here we will discuss an important application of quantum dots in biology and medicine: the use of quantum dots as *biological markers*. In this application, the main idea is to coat the dot with a material that causes it to bind selectively to a certain biological structure, such as a cancer cell, and then to reveal its presence by absorbing and emitting light (in a process known as *fluorescence*). Light is absorbed by the dot, raising electrons from lower to higher energy states. In general, there is some energy dissipation, and more importantly, the electrons will tend to fall back down to lower states, emitting a photon. The ability to tune precisely the energy levels in a quantum dot is of great importance.

As a rough approximation, the emission spectra of a quantum dot can be determined by the simple three-dimensional hard-wall particle-in-a-box model discussed in Section 4.3. For example, energy levels for a cubic dot having side length L are, from (4.54),

$$E_n = \frac{\hbar^2\pi^2 n^2}{2m^* L^2}, \tag{9.58}$$

where $n^2 = n_x^2 + n_y^2 + n_z^2$ represents the triplet of integers $n = (n_x, n_y, n_z)$. Energy transitions from, say, the $n = (2, 1, 1)$ state to the $n = (1, 1, 1)$ state release the energy quantum

$$\Delta E_{2,1} = E_{2,1,1} - E_{1,1,1} = \frac{\hbar^2 \pi^2 (6)}{2m^* L^2} - \frac{\hbar^2 \pi^2 (3)}{2m^* L^2} = 3\frac{\hbar^2 \pi^2}{2m^* L^2}. \tag{9.59}$$

As an example, for a $10 \times 10 \times 10$ nm^3 GaAs dot ($m^* = 0.067 m_e$), $\Delta E_{2,1} = 0.168$ eV.

For a spherical dot of radius R, the details of obtaining the energy levels are a bit more involved than for the rectangular dot, although the resulting expression for the energy levels is very similar,

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2m^* R^2}. \tag{9.60}$$

For an $R = 6.2$ nm radius sphere, which has the same volume as the $10 \times 10 \times 10$ nm^3 box considered previously, $E_{2,1} = 0.438$ eV.

We can see that the electron mass is important in determining energy spacings. In a typical metal, the electron mass is the familiar quantity $m_e = 9.1095 \times 10^{-31}$ kg, which would lead to $\Delta E_{2,1} = 0.0293$ eV for the $R = 6.2$ nm radius sphere. Therefore, it is much easier to distinguish the discrete nature of the energy levels in a semiconducting quantum dot than in a metallic dot.[†]

In light of the discussion in Section 9.1.2, it can be seen that (9.58) and (9.60) actually relate to intersubband transitions. Referring to Fig. 9.6, we find that direct-gap interband transitions in spherical dots from the n^{th} level in the valence band well to the m^{th} level in the conduction band well can occur for incident photon energies

$$\hbar\omega = E_g + E_n + E_m \tag{9.61}$$

$$= E_g + \frac{\hbar^2 \pi^2 n^2}{2m_h^* R^2} + \frac{\hbar^2 \pi^2 m^2}{2m_e^* R^2}, \tag{9.62}$$

where E_g is the band gap energy of the bulk semiconductor.

For transitions among the lowest states ($n = m = 1$), we have

$$\hbar\omega = E_g + \frac{\hbar^2 \pi^2}{2R^2} \left(\frac{1}{m_h^*} + \frac{1}{m_h^*} \right) \tag{9.63}$$

$$= E_g + \frac{\hbar^2 \pi^2}{2m_r^* R^2}, \tag{9.64}$$

where m_r^* is the reduced mass.

Furthermore, in contrast to bulk semiconductors, in quantum dots exciton effects often play a dominant role in determining optical properties of the dot at room temperature. In fact, optical transitions in quantum dots are usually associated with excitons (Section 5.4.5), and an approximation expression known as the *Brus equation* models the transition energy in spherical dots,

$$E_g^{\text{dot}} = E_g + \frac{\hbar^2 \pi^2}{2m_r^* R^2} - \frac{1.8 q_e^2}{4\pi\varepsilon_r \varepsilon_0 R}. \tag{9.65}$$

The third term in (9.65) is related to the binding energy of the exciton (i.e., it is due to the Coulomb attraction between the electron and the hole), which is modified from the bulk case by the size of the dot.

In bulk semiconductors, the exciton radius is given by (5.56),

$$a_{ex} = \frac{4\pi\varepsilon_r \varepsilon_0 \hbar^2}{m_r^* q_e^2} = \frac{\varepsilon_r m_e}{m_r^*} a_0 = \frac{\varepsilon_r m_e}{m_r^*} (0.53 \text{ Å}), \tag{9.66}$$

where a_0 is the Bohr radius. For quantum dots, the actual (highest probability) separation between the electron and hole is influenced by the size of the dot. In this case, we will consider a_{ex} to be the excition Bohr radius, which is often taken as the measure of quantum confinement in quantum dots. In particular,

- if $R \gg a_{ex}$, then confinement effects will generally not be very important. Otherwise,
- if $R > a_{ex}$, then we have the *weak confinement regime*, and
- If $R < a_{ex}$, then we have the *strong confinement regime*. In this case, the hydrogen model tends to break down, and the exciton is delocalized over the entire dot.

For semiconducting dots, Group II–VI semiconductors such as ZnSe, ZnS, and CdSe are often used, since these materials tend to have relatively large bandgaps and can be fabricated using a variety of methods. For self-assembly, in particular, CdSe/ZnSe systems naturally lead to quantum dots because of the large lattice mismatch between ZnSe and CdSe (about 7 percent).

As an example, consider a cadmium selenide (CdSe) quantum dot. Using $E_g^{\text{bulk}} = 1.74$ eV, $m_e^* = 0.13 m_e$, $m_h^* = 0.45 m_e$ ($m_r^* = 0.101 m_e$), and $\varepsilon_r = 9.4$ (choosing reasonable values from measurements), for transitions from the conduction to valence bandedges through the band gap, we obtain from (9.63), for an $R = 2.9$ nm dot, $E_g = 2.092$ eV. Using $E = hf$ and $c = \lambda f$, this energy corresponds to $\lambda = 593$ nm, indicating that yellow light (Table 9.1)

TABLE 9.1 COLORS AND ASSOCIATED ELECTROMAGNETIC WAVELENGTHS.

Color	λ (nm)
Red	780–622
Orange	622–597
Yellow	597–577
Green	577–492
Blue	492–455
Violet	455–390

[†] We could say, equivalently, that this effect is due to the small Fermi wavelength in metals, and the relatively larger Fermi wavelength in semiconductors.

would be absorbed by the dot. The experimental absorption peak is closer to 580 nm, although the Brus formula gives reasonably accurate (at least, better than order-of-magnitude) results. In this case, the exciton Bohr radius is

$$a_{ex} = \frac{\varepsilon_r m_e}{m_r^*}\left(0.53\ \text{Å}\right) = 4.93\ \text{nm}, \tag{9.67}$$

which would correspond to strong confinement.

The energy reradiated by the dot is less than that which excites the dot, so that radiated wavelengths of the fluorescence are longer. This difference is called the *Stokes shift*, which relates to relaxation of angular momentum in the dot, although the details won't be presented here. However, experiments show that an $R = 2.9$ nm dot emits at approximately $\lambda = 592$ nm, or yellow (and so, using the previously given experimental absorption peak of 580 nm, we find that the Stoke's shift is approximately $\Delta\lambda = 592\ \text{nm} - 580\ \text{nm} = 12\ \text{nm}$). Note that having a Stokes shift is generally considered a positive attribute of a biological marker, since the illuminating energy and the re-radiated energy can be separated (filtered). This provides an advantage for quantum dots compared with fluorescent dyes, which radiate at nearly the same wavelength as their excitation. Illumination is often provided in the UV range, since the absorbance spectrum is relatively broad, whereupon the dot reradiates at the wavelength corresponding to its size and material composition. Thus, if different-sized dots are coated to bind with different biological features, with a common illumination, the various dots will fluoresce in different colors. As another example, if one breaks off a nanoscopic piece of a material (say, a rock) of a certain color, that piece will often be a different color from the original, due to the described quantum confinement effects. This has been used (unknowingly) in making stained glass for centuries.

Size Effects. As described previously, when material regions shrink to the nanoscale, the properties of the material change. This is due to a variety of reasons, such as the large ratio of surface-to-interior atoms, and the change in bandstructure due to having the periodic crystalline geometry of a material, which is effectively infinite in the bulk case, significantly truncated. The effect on bandgap energy is described previously.

Furthermore, it should be noted that even the lattice structure of a material may change for nanoscopic material samples. For example, both gold and silver have the fcc lattice in the bulk state, but for particles having a diameter of less than 5–8 nm, significant deviations from the fcc structure have been observed, which will obviously alter the material's properties.

Lastly, it can be appreciated that the role of semiconductor doping is also strongly affected when considering nanoscale material regions. Typical semiconductor doping levels for traditional device applications are on the order of 10^{16}–10^{18} impurity atoms per cm^3. This would result in, on average, 0.00001–0.001 atoms per nm^3. Thus, a quantum dot 10 nm on a side would occupy 10^3 nm^3, and would, therefore, contain approximately one extra electron or hole available for conduction.

Blinking and Spectral Diffusion. It has been observed that the emission spectrum of colloidal dots shifts randomly over time, which is known as *spectral diffusion*. This shift can occur over times as long as seconds, or even minutes. Although this phenomenon is not completely understood, it is thought to arise from the local environment of the dot, which

can admit fluctuating electric fields. Spectral diffusion is not observed in self-assembled dots embedded in a host matrix.

Another phenomenon observed in colloidal quantum dots is known as *blinking*, which is, as its name suggests, a turning on and off of the dot's emissions. This phenomenon is also not completely understood (although it seems to be related to charging of the dot), and along with spectral diffusion, is generally considered a detrimental aspect. The off state of the dot can last from milliseconds to minutes.

Coated and Functionalized Quantum Dots. In addition to simple dot structures, often dots are coated with multilayered shells to tailor their electronic, chemical, or biological properties. A core material (often cadmium sulfide (CdS), CdSe, or cadmium telluride (CdTe)) and a core size/shape is chosen based on the desired emission wavelength; CdSe dots are currently often chosen for most of the visible spectrum. To the core is added a shell, consisting of a transparent material or layering of materials that can be attached to the core. The shell material (e.g., ZnS) serves as a protective coating for the core. Multilayer coatings can be added to stabilize the dots' emission spectra, to render the dots chemically inert,† and to provide a surface to attach biological structures for, e.g., selective binding.

For example, Fig. 9.14 shows cultured HeLa cells‡ labeled with two different quantum dots, one that fluoresces red, and the other, green. The red and green dots are conjugated (bound) to different antibodies and label two different proteins of interest in cell biology.

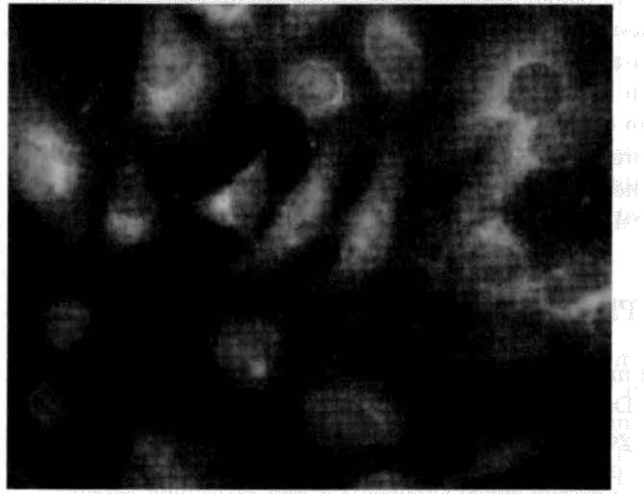


Figure 9.14 Cultured HeLa cells labeled with two different quantum dots, one that fluoresces red, and the other, green. The dots are conjugated (bound) to different antibodies, and label two different proteins of interest in cell biology. (Reprinted with permission from Quantum Dot Corporation.)

†Cadmium itself is a toxic heavy metal.

‡HeLa cells are cultured cervical cancer cells that are routinely used in the study of cancer.

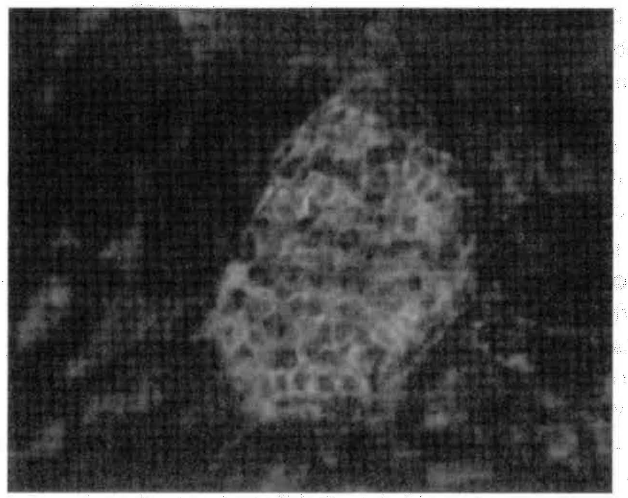


Figure 9.15 Two different quantum dots that fluoresce red and green are shown labeling proteins in a tissue section of mouse kidney. (Reprinted with permission from Quantum Dot Corporation.)

The intensity of the fluorescence and the distribution in the cells are the key features of this technique in cell biological research.

In a similar manner in Fig. 9.15, two quantum dots which fluoresce red and green are shown labeling proteins in a tissue section of mouse kidney. This application is typically used in clinical pathology and medicine for the diagnosis of disease states. An understanding of the amount and distribution of protein expression helps define the cancerous or noncancerous state of the cells in the tissue section.

In Fig. 9.16, quantum dots injected into a live mouse mark the location of a tumor.

In addition to quantum dots, quantum wires can also be functionalized. For example, nanowires can be coated with substances that will bind with certain molecules, which will, in turn, alter the conductance of the wire. Thus, such nanowires can be used as, for example, chemical sensors.

9.3.2 Plasmon Resonance and Metallic Nanoparticles

Quantum confinement effects can be seen in metallic nanoparticles if their size is extremely small. Due to the small Fermi wavelength of electrons in metals, quantum confinement effects generally only become important for metallic spheres having radius values far below 10 nm; perhaps on the scale of 1–2 nm or less. For metallic nanoparticles above this size range, optical properties are governed by *plasmon resonances*, which are collective modes of oscillation of the electron gas in the metal. This is a classical, rather than quantum, effect, and can be described by classical Maxwell's equations.

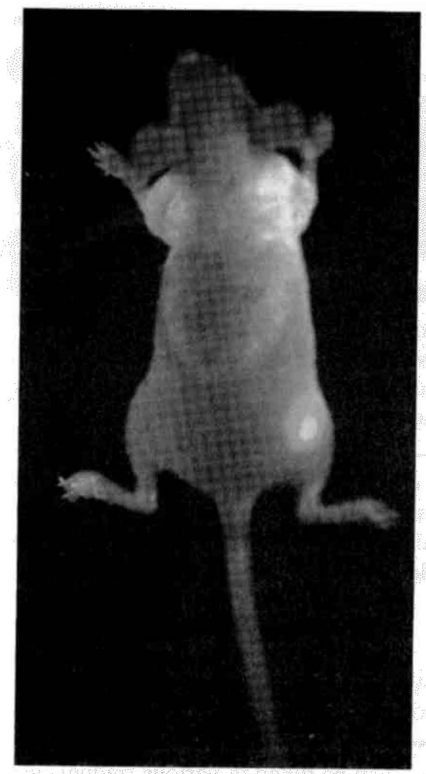


Figure 9.16 Quantum dots injected into a live mouse mark the location of a tumor. (From Seydel C., "Quantum Dots Get Wet," *Science* 4 April 2003, 300: 80 [DOI: 10.1126/Science.300.5616.80]. Courtesy Xiaohu Gao and Shuming Nie, Emory University School of Medicine.)

At the frequencies of plasmon resonances, the response of the material to illumination is particularly strong and easy to detect. It is important to note that these oscillations are not very sensitive to the size of the nanoparticle, contrasting the extreme size dependence of confinement effects in semiconductor dots and very small metallic dots. For example, consider applying a static (non-time-varying) electric field to a small dielectric sphere of radius a in free space. It can be shown that the ratio of the field inside the sphere to the applied field is

$$\frac{3}{\varepsilon_r + 2}, \tag{9.68}$$

where ε_r is the relative permittivity of the dielectric. For low-frequency, time-varying fields where $\lambda \gg a$, with λ the electromagnetic wavelength, the same result holds.[†] Thus, if $\varepsilon_r = -2$, the field inside the sphere "blows-up," (i.e., in a real material the interior field takes on a very large value[‡]), independent of the specific radius of the sphere. That is, the effect is not due to fitting an integer number of half wavelengths inside the sphere, as is the case for most resonance phenomena. The effect is actually due to interactions between interior charge and charge induced on the surface of the metallic particle. Furthermore, a similar result holds for other structures, not only spheres. Metals in the optical range have negative

---

[†] Note that since a is on the order of nanometers, "low frequency" certainly includes optical frequencies.

[‡] All real materials have some damping, which prevents perfect oscillations from occurring.

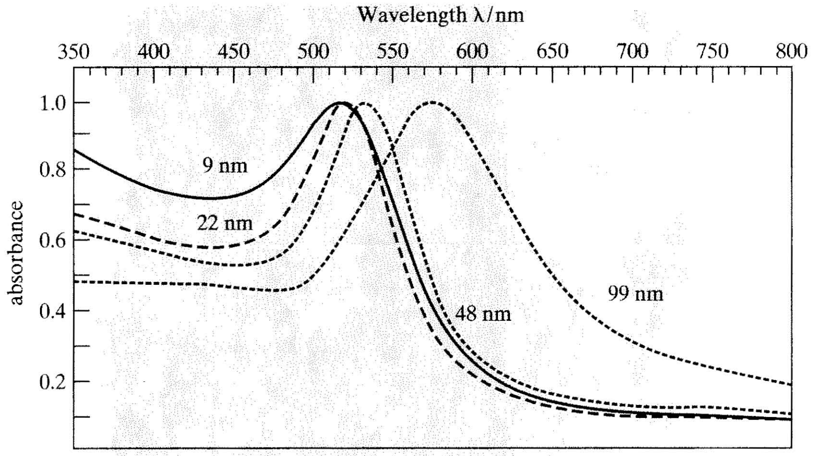


Figure 9.17 Size effects on the surface plasmon absorption of colloidal solutions of spherical gold nanoparticles with diameters varying between 9 and 99 nm. (Based on a figure from Link, S. and M. A. El-Sayed, "Spectral Properties and Relaxation Dynamics of Surface Plasmon Electronic Oscillations in Gold and Silver Nanodots and Nanorods," *J. Phys Chem B* 103 (1999): 8410–8426.)

relative permittivities, and therefore small metal particles tend to exhibit relatively size-insensitive plasmon resonances, rather than size-specific geometrical resonances, although the latter will also occur. In addition, particularly in hollow nanoshells of metal, these plasmon resonances can be tuned to various frequencies by controlling the relative thickness of the core and shell layers.

To demonstrate the relative size insensitivity of plasmon resonances, Fig. 9.17 shows the measured absorbance of a colloidal solution of spherical gold nanoparticles with diameters varying between 9 and 99 nm. It can be appreciated that as the diameter changes from 9 nm to 48 nm (a 433 percent change), the plasmon resonance energy only changes by approximately 0.06 eV (from 2.34 eV to 2.40 eV; a 2 percent change).

9.3.3 Functionalized Metallic Nanoparticles

In addition to semiconducting dots, metallic nanoparticles can also be functionalized. For example, Fig. 9.18(a) shows silver nanoparticles, and Fig. 9.18(b) shows gold nanoshells made by coating the silver particles and dissolving the silver. These metallic nanoparticles can be functionalized for diagnostic, or even therapeutic purposes. For example, cancer cells tend to have certain proteins on their surfaces that are not found, at least in the same concentration, in healthy cells. By conjugating (binding) a certain antibody to the gold nanoparticles, the nanoparticles will attach themselves preferentially to the cancer cells. Since gold nanoparticles can be detected using visible light (as mentioned previously, due to plasmon resonances), the cancer cells can be imaged. Furthermore, significant amounts of heat can be generated by a resonating nanoshell, and, by tuning the nanoshell to resonate in response to a particular frequency,[†] nanoshells embedded in cancer cells can deliver a therapeutic dose of heat, pinpointed to the cell itself. This only requires moderate amounts of energy illuminating the subject, such that excessive heating of nontargeted (i.e., healthy) cells can be avoided.[‡]

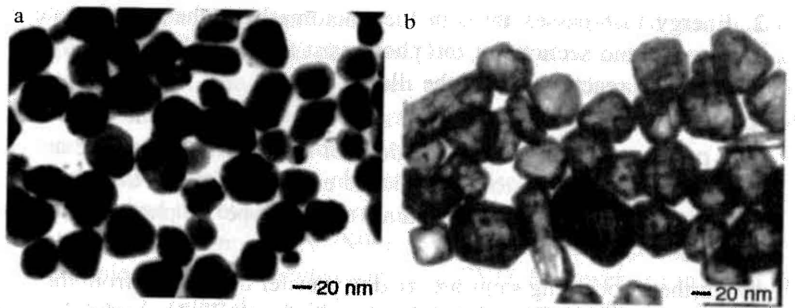


Figure 9.18 (a) TEM of silver nanoparticles. (b) TEM of gold nanoshells formed by reacting the silver nanoparticles of (a) with an aqueous HAuCl4 solution. The silver dissolves, leaving the gold shell remaining. (From Sun, Y., B. T. Mayers, and Y. Xia, "Template-Engaged Replacement Reaction: A One-Step Approach to the Large-Scale Synthesis of Metal Nanostructures with Hollow Interiors," *Nano Lett.* 2 (2002): 481. © 2002, American Chemical Society.)

9.4 FABRICATION TECHNIQUES FOR NANOSTRUCTURES

The mainstay of commercial silicon technology fabrication is optical lithography. As discussed briefly in Chapter 1, it seems that optical lithography can be pushed to resolutions on the order of several tens of nanometers. Other technologies, like extreme UV lithography, immersion lithography, or electron beam lithography may become sufficiently mature for commercial applications in the near future. A very brief discussion of several important techniques will be provided next.

9.4.1 Lithography

The main idea of lithography was described in chapter 1, and here a brief recapitulation of electromagnetic lithography, which includes optical, EUV, and X-ray techniques, is provided.

1. Electromagnetic energy is directed at a photomask containing opaque and transparent regions that correspond to the desired pattern.

[†] Infrared frequencies are typically used, where transmission through tissue is adequate such that the illuminating field can penetrate to the targeted cell.

[‡] Heating of metal nanoparticles for medical applications has even been observed at frequencies far below the range of plasmon resonances. For example, in [23], a solution consisting of 10 nm gold nanoparticles and protein aggregates associated with Alzheimer's disease was heated using low-power (100 mW) 12 GHz electromagnetic radiation. It was estimated that each nanoparticle dissipated 10^{-14} J/s, which was more than enough to dissolve the protein aggregates.

2. Energy that passes through the photomask reaches a substrate coated with a photoresist, and sections of the photoresist that are illuminated by the energy undergo a chemical reaction due to the illumination.
3. Upon washing the structure in a solvent, sections of the resist are dissolved, forming a pattern on the resist. The regions of the resist that dissolve are either those sections that were exposed to the incident energy, or those that were shielded by opaque regions of the mask, depending on the type of photoresist (positive or negative) used.
4. Further processing steps are used to transfer the pattern from the resist to the substrate, or otherwise to change the substrate in the desired manner. These include
   (a) *Wet Etching*: Wet etching is the removal of material using acids, or any liquid solution that will dissolve the material in question. The patterned resist layer protects areas of material that are not to be removed. The depth of the etch can be controlled by the choice of etchant, and by limiting the etching time. However, undercutting of the resist is common in wet etching.
   (b) *Dry Etching*: Although there are several forms of dry etching, the easiest to understand is a method of bombarding the material by energetic ions inside a vacuum chamber. The ions collide with atoms in the material (called the target), resulting in momentum transfer, and causing these atoms to be ejected from the surface of the target. Again, the resist protects areas of the target that are not to be removed. In general, dry etching is more expensive than wet etching, although it has better resolution, and undercutting can be generally avoided. Another form of dry etching uses the chemical reactions between a gas and the substrate to remove material.
   (c) *Doping*: For semiconducting substrates, it is often necessary to dope the material. This can be accomplished, for instance, by accelerating a beam of dopant ions towards the substrate. The resist blocks the ions from reaching those regions of the substrate covered by the resist, and, thus, creates regions of doping in areas not covered by the resist. This is known as *ion implantation*.
   (d) *Material Deposition*: Material may be deposited (for example, metal) onto the wafer. There are several methods of deposition:
      i. *Sputtering* : Sputtering is basically the bombardment of a material (the target) by energetic ions inside a vacuum chamber. The ions collide with atoms in the target, resulting in momentum transfer, causing these atoms to be ejected from the surface of the target. These atoms then become deposited on the adjacent substrate, forming a thin film. Since sputtering is essentially a mechanical (rather than chemical) process, a wide range of materials can be deposited onto the substrate.
      ii. *Chemical vapor deposition* (CVD): In CVD, a substrate is placed inside a chamber with a number of gases. A chemical reaction between the gases produces a solid material that condenses on the substrate (and everywhere else), forming a thin film.

      iii. *Evaporation*: In evaporation, a metal to be deposited and a substrate are placed inside a vacuum chamber. The metal is heated till it melts, and begins to evaporate. The evaporate then condenses on the substrate, forming a thin film. Heating of the metal occurs either by resistive effects in a sample holder, or by directing an electron beam at the metal.
      iv. *Laser ablation*: In laser ablation, a pulsed laser beam vaporizes part of a target, and the evaporate forms a thin layer on a nearby substrate.
      v. *Molecular beam epitaxy* (MBE): Molecular beam epitaxy[†] is an important technique that is somewhat similar to CVD. However, it can result in epitaxial layers of materials, including compound semiconductors. In MBE, elemental materials in separate chambers are evaporated in a high vacuum, and beams of the resulting molecules or atoms deposit material on a heated substrate. The deposited material builds up on the substrate with the same crystallographic orientation as the preceding atomic layer. This can form an atomically sharp interface between, say, GaAs and AlGaAs. Therefore, MBE in particular is an important technique in forming two-dimensional electron gases.
   (e) After processing, the photoresist is removed (along with any material deposited on top of the photoresist, which is known as *lift-off*), leaving the desired structure.

For fabricating nanostructures, electron beam lithography, because of its nanometer scale resolution, has been a mainstay in laboratory settings. In this technique, electron beams serve as the illumination source, rather then electromagnetic waves, although the idea of having a resist, and doing various subsequent processing steps, is the same as for electromagnetic lithography. For research applications, it is common to perform electron beam lithography using an electron microscope. A figure of a quantum wire formed by electron beam lithography and ion etching is shown on page 125.

Many other variations of lithography exist, all involving forming some sort of template, and then removing or depositing material on the structure, after which the template is removed.

9.4.2 Nanoimprint Lithography

Nanoimprint lithography (NIL) consists of pressing a stamp or mold conforming to the desired pattern onto a thin film (perhaps a photocurable resist) on top of a wafer. The resist is heated to a temperature at which it is thermoplastic, becoming a viscous liquid. Thus, the film can flow, and be deformed into the shape of the mold, as depicted in Fig. 9.19. Various processing steps, including etching and material deposition, are then performed, as is done in the previously described lithography technologies. The mold itself is often produced by electron beam lithography.

Fig. 9.20 shows a T-structure made from the NIL process depicted in Fig. 9.19.

[†] Epitaxy is the growth of one material on another such that there is a crystallographic relationship between the two materials.

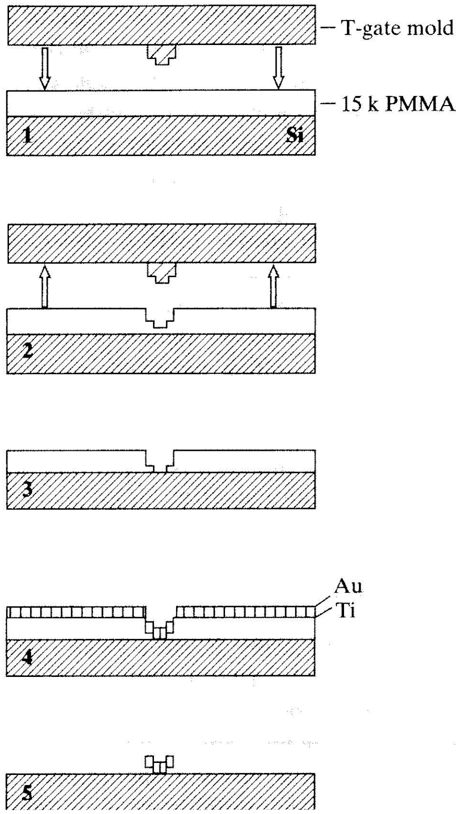


Figure 9.19 Depiction of NIL lithography to make a T-structure: the mold is pressed into the substrate, then retracted; metal is deposited; the substrate is dissolved and the extra metal is lifted off, leaving the desired shape. (Based on an image from Li, M., L. Chen, and S. Y. Chou, "Direct Three-Dimensional Patterning Using Nanoimprint Lithography," *Appl. Phys. Lett.* 78 (2001): 3322. © 2001, American Institute of Physics.)

9.4.3 Split-Gate Technology

Another method to form laterally confined nanostructures is the use of what are called *split gates*. This forms an electrically controllable structure that can be adjusted by an applied voltage. The idea is schematically depicted in Figs. 9.21 and 9.22.

One starts with a two-dimensional electron gas, formed, for instance, by a semiconductor heterostructure. As shown in Fig. 9.21, Schottky gate electrodes are deposited on top of the heterostructure. Left unbiased, the electron gas remains unperturbed. However, application of a negative potential to the electrodes depletes the electron gas below the gate, and so the gate electrodes form a narrow electron gas channel under the split, resulting in a quantum wire (into and out of the plane of the page), as shown in Fig. 9.22. Often the lateral confinement forming the wire is somewhat weak (this can also be true of wires formed by etching, due to, for example, undercutting), resulting in a "soft" barrier. This results in wire energies smaller than those given by (9.50).

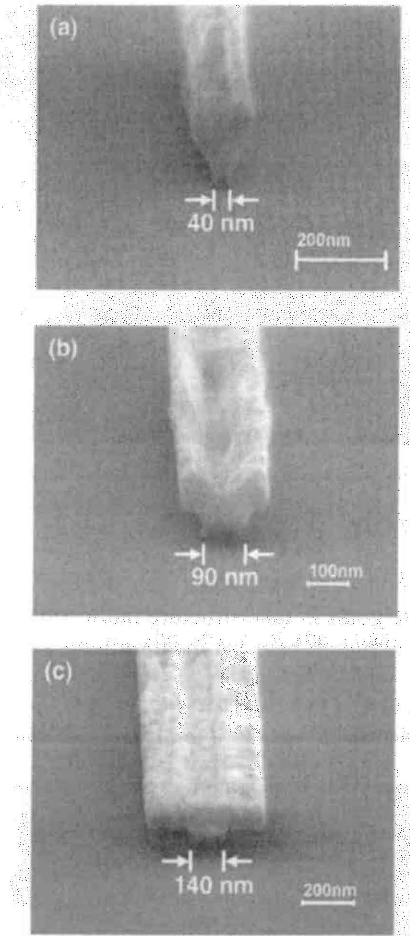


Figure 9.20 Resulting T-structure made from the NIL process depicted in Fig. 9.19. (Reused with permission from Li, M., L. Chen, and S. Y. Chou, "Direct Three-Dimensional Patterning Using Nanoimprint Lithography," *Appl. Phys. Lett.* 78 (2001): 3322. © 2001, American Institute of Physics.)

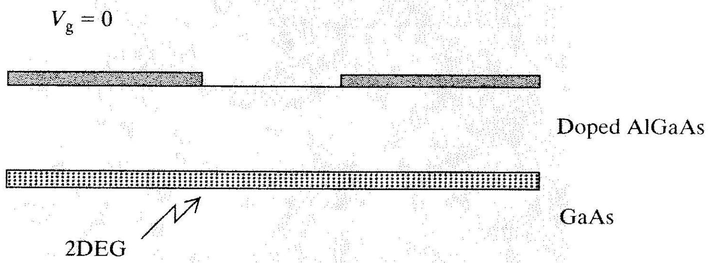


Figure 9.21 Split-gate electrodes over a two-dimensional electron gas.

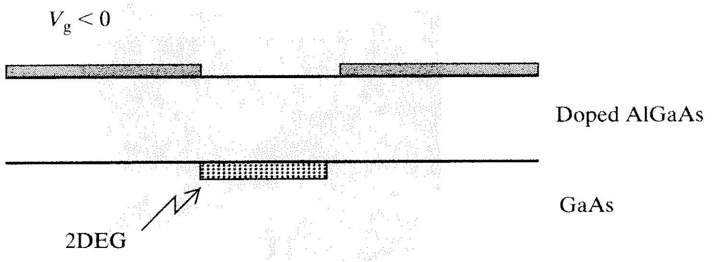


Figure 9.22 Split-gate electrodes over a two-dimensional electron gas. Negative gate bias depletes electrons below the gates, constraining the electron gas to a narrow region, forming a quantum wire.

Quantum dots can also be fabricated in a similar manner. In the case of a split-gate quantum dot, the electrode pattern constrains the 2DEG formed by a heterojunction in the shape of a hole in the electrode, as shown in Fig. 9.23.

9.4.4 Self-Assembly

Self-assembly techniques can be used to form a variety of nanostructures, and self-assembly is one of the ultimate goals in nanostructure fabrication. Many new techniques will undoubtedly be developed, although here we mention three common methods.

1. **Lattice mismatch:** The different crystallographic features of different materials can be used to advantage in self-assembly, especially in the formation of quantum dots and related nanostructures. For example, the lattice mismatch (and associated strain) between Ge and Si results in the formation of three-dimensional islands when Ge is deposited epitaxially on Si. The shape of the islands depend on the number of monolayers of Ge deposited on the Si surface—for six monolayers characteristic square pyramids tend to form, as shown in Fig. 9.24.

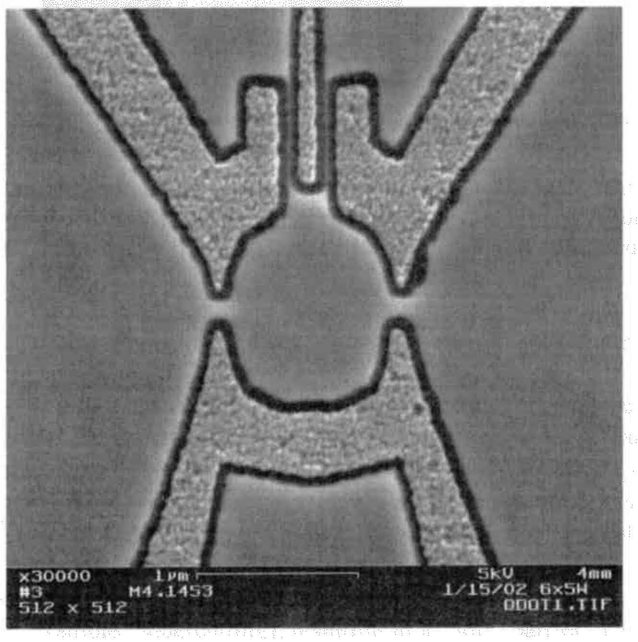


Figure 9.23 Quantum dot formed from split gates. (Courtesy of Jonathan Bird and Yuki Takagaki.)

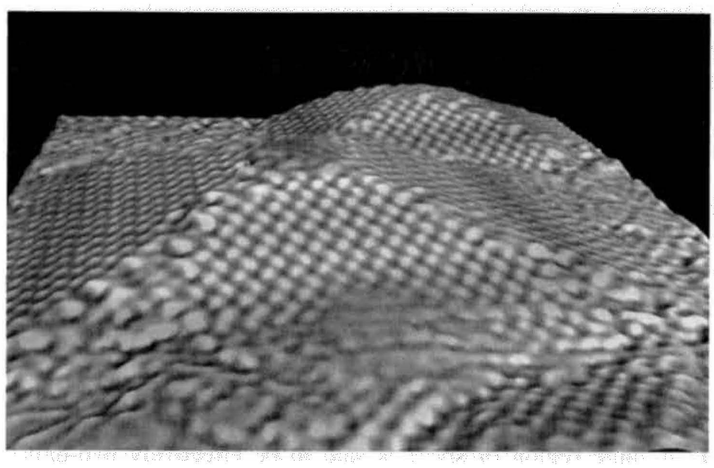


Figure 9.24 STM image of a self-assembled Ge islands on Si. (Reprinted with permission from Gilberto Medeiros-Ribeiro, HPL Palo Alto.)

2. **Wet chemistry methods:** Quantum dots can also be synthesized by forming nano-sized precipitates from chemical solutions. Quantum dots so formed are called *colloidal quantum dots*.[†] This method has the advantage of producing large quantities of quantum dots in solution, which is often useful for biological or medical applications.
3. **Molecular self-assembly:** Molecular self-assembly is, as the name implies, the process of molecules organizing themselves into a desired structure. Often this takes place in liquids, but it can occur in a gas and other host mediums. Molecular self-assembly is obviously the goal in molecular electronics.

9.5 MAIN POINTS

In this chapter, we have considered models of quantum wells, quantum wires, and quantum dots, especially those structures fabricated from semiconducting materials. In particular, after studying this chapter, you should understand

- the idea of Fermi wavelength and exciton Bohr radius as important length scales in deciding when quantum confinement effects are important;
- the idea of a semiconductor heterostructure;
- hard-wall and other confinement models for forming a quantum well;

[†] A colloidal system (or colloid) is a collection of small particles, typically having sizes on the order of one to several hundred nanometers, dispersed in a host medium (often a liquid). The particles are generally small enough to be moved about by molecular collisions, and tend not to settle under gravity.

- energy subband structures in quantum wells and quantum wires;
- interband, intersubband, and exciton transitions in quantum wells, and the existence of selection rules;
- simple models of quantum dots, including the importance of exciton effects and the exciton radius;
- applications of quantum dots, including their use as biological markers;
- the idea of metal nanoparticles and plasmon resonance;
- fabrication methods, including various lithography techniques, split-gate techniques, and self-assembly.

9.6 PROBLEMS

1. A finite region of space is said to be effectively two-dimensional if $L_x \leq \lambda_F \ll L_y, L_z$. However, λ_F was computed from the three-dimensional result (9.3)

$$\lambda_F = \frac{2\pi}{\left(3N_e\pi^2\right)^{\frac{1}{3}}}, \tag{9.69}$$

where N_e is the electron density m^{-3}. Is it then permissible to use this formula to conclude that a structure is two or one dimensional? Shouldn't the two- or one-dimensional λ_F be used? Why or why not?

2. At the beginning of this chapter, it was stated that the Fermi wavelength is the important parameter in deciding if quantum confinement effects are important, and for determining the effective dimensionality of a system. Another method to characterize when quantum confinement effects are important is to say that if the difference in adjacent energy levels for a finite space is large compared with other energies in the system (thermal, etc.), then confinement effects are important. Considering a one-dimensional infinite-height potential well of width L, as considered in Section 4.3.1, show that confinement effects will be important compared with thermal energy when

$$L < \sqrt{\frac{3\hbar^2\pi^2}{m_e^* k_B T}}. \tag{9.70}$$

However, note that this result is independent of doping, and is, therefore, quite a rough approximation for semiconductors. For GaAs and Si, estimate the preceding relationship at room temperature.

3. As noted in Section 9.3.1, in quantum dots, one can take the exciton Bohr as the important length scale for deciding if quantum confinement effects are important. In what way is this a qualitatively similar idea to the use of the Fermi wavelength for the same purpose?

4. On page 288, it was shown that for GaAs at room temperature, one would need $L_x \ll 15$ nm for two-dimensional subbands to be resolved. Determine the corresponding value for copper.

5. In a 2DEG, in order for energy subbands to be evident, $k_B T$ should be much less than the difference between subbands, leading to (9.35). Consider the situation from the perspective of the de Broglie wavelength associated with thermal energy,

$$\lambda_e = \frac{h}{\sqrt{2m^* E}} = \frac{h}{\sqrt{2m^* k_B T}}, \tag{9.71}$$

and determine the size constraints on L_x to observe subbands at room temperature, and at $T = 4$ K, for GaAs. Comment on the relationship between the results of (9.35) and the results from (9.71).

6. Assume the well structure depicted in Fig. 9.3 on page 283. Since GaAs has a bandgap of $E_g \simeq 1.5$ eV, and $\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}$ has a band gap of

$$E_g \simeq 1.426 + 1.247\alpha \tag{9.72}$$

for $\alpha < 0.45$, determine α if the height of the confinement barrier $\Delta E_C = E_C^{\mathrm{AlGaAs}} - E_C^{\mathrm{GaAs}}$ is 0.238 eV.

7. In Section 4.5.1, the quantum states in a finite-height potential well were determined, leading to the eigenvalue equation (4.83) for symmetric states, and (4.85) for anti-symmetric states. In that example, the effective mass in each region was the same. Redo this problem assuming effective masses m_w^* and m_b^* for the well and barrier regions, respectively, and show that the resulting eigenvalue equations are

$$k_2 \tan(k_2 L) = \frac{m_w^*}{m_b^*} k_1 \tag{9.73}$$

for the symmetric states, and

$$-k_2 \cot(k_2 L) = \frac{m_w^*}{m_b^*} k_1 \tag{9.74}$$

for the antisymmetric states, where

$$k_2^2 = \frac{2m_w^* E}{\hbar^2}, \quad k_1^2 = \frac{2m_b^* (V_0 - E)}{\hbar^2}. \tag{9.75}$$

8. Comparing energy levels in one-dimensional quantum wells, for an infinite-height well of width $2L$ and effective mass m^*, energy states are given by (4.35) with L replaced by $2L$ and m_e replaced by m^*,

$$E_n = \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{2L}\right)^2, \tag{9.76}$$

and for a finite-height well of width $2L$, the symmetric states are given by a numerical solution of

$$k_2 \tan(k_2 L) = \frac{m_w^*}{m_b^*} k_1, \tag{9.77}$$

as described in Problem 9.7. Assume that $L = 5$ nm, $m_w^* = 0.067 m_e$ (GaAs), and $m_b^* = 0.092 m_e$ ($\mathrm{Al}_{0.3}\mathrm{Ga}_{0.7}\mathrm{As}$; see Table VI in Appendix B). Compute E_1 and E_3 (the second symmetric state) for the infinite-height well, and compare it with the corresponding values obtained from the numerical solution for the finite-height well. (You will need to use a numerical root solver.) For the finite-height well, assume barrier height $V_0 = 0.3$ eV, which is $\Delta E_C = E_C^{\mathrm{Al}_\alpha\mathrm{Ga}_{1-\alpha}\mathrm{As}} - E_C^{\mathrm{GaAs}}$ for $\alpha = 0.3$. Comment on the appropriateness, at least for low energy states, of the much simpler infinite well model.

9. Assume a hard-wall model of a rectangular cross-section metallic quantum wire. If the wire is 1 nm thick and 10 nm wide, determine how many subbands are filled at 1 eV.

10. In a quantum wire, the available energy gets partitioned into various channels (subbands). Is this partitioning unique?

11. Calculate the energy levels in a quantum dot in the form of a cube, 5 nm on a side. Assume zero potential energy in the dot, and an infinitely high potential bounding the dot. Assume that $m^* = 0.045m_e$ in the dot material.

12. How small must a metal nanosphere be in order for the ground state energy to be 1 eV? Assume a hard-wall model.

13. Consider a hard-wall model of a cubical metal quantum dot, having side length $L = 3$ nm. What is the energy of a transition from the $(1, 1, 1)$ state to the $(2, 2, 1)$ state?

14. Consider a CdS quantum dot, having $E_g^{\mathrm{bulk}} = 2.4$ eV, $m_e^* = 0.21m_e$, $m_h^* = 0.8m_e$, and $\varepsilon_r = 5.6$. What is the expected peak absorbance energy (E_g^{dot}) and wavelength for a $R = 2$ nm dot? Repeat for a $R = 3$ nm dot.

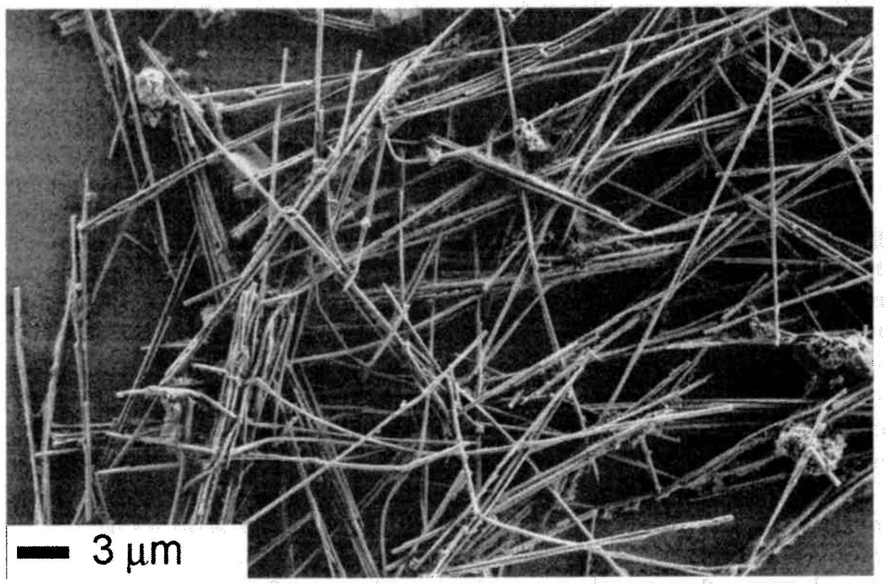
15. Using the Internet, find a company that sells quantum dots for biological and/or medical applications, and write a one-half to one page summary of one of their products, and its applications.

16. There is significant debate about the possible adverse health effects of quantum dots in biological bodies. Much, but not all, of the concern revolves around the use of cadmium in quantum dot structures. Write a one page summary of the debate, including the benefits and drawbacks of using quantum dots in living bodies.

17. Determine quantum dot applications to lasers, and write a one-half to one page summary of the uses, pros, and cons, of using quantum dots as lasing materials.

Chapter 10

NANOWIRES, BALLISTIC TRANSPORT, AND SPIN TRANSPORT



SEM image of 200 nm diameter nickel nanowires. Ann Bentley & Professor Art Ellis, Materials Science and Engineering Department, University of Wisconsin-Madison. Image courtesy the Nickel Institute.)

In an ordinary electrical circuit, conducting wires are used to interconnect electrical devices. For example, common residential wiring typically consists of 12 and 14 gauge[†] round cross-section wire, having diameters of 2.05 mm and 1.63 mm, respectively. Discrete electronic circuits built on a breadboard typically use 22 gauge wire, having a diameter of 0.64 mm. Integrated circuits use printed interconnects, having roughly rectangular cross sections, with typical widths on the order of 0.16 μm, and thicknesses on the order

---

[†] Americal Wire Gauge, AWG.