Chapter 1
Fundamentals on Bionanotechnologies

Abstract This is the introductory chapter of the book. The basic theoretical and experimental facts regarding the application of electronics at the nanoscale and for biological systems are developed here. Transport phenomena at the nanoscale, the principles of nanotechnologies, the physical properties of biological materials, and micro/nanofluidics are reviewed and explained in this chapter. The knowledge gained in this chapter will then be used in the entire book.

1.1 Transport Phenomena at the Nanoscale

When electronic devices are scaled down from few microns up to nanoscale, they become comparable with living organisms, such as bacteria, viruses, or the dimensions of DNA bases. The nanoscale is represented in Fig. 1.1. This fact is of paramount importance for sensing, detecting, or manipulating microorganisms or biomolecules.

The reduced nanometer dimensions of electronic devices changes completely the transport properties. A nanoscale device is an electron device where one, two, or even all three spatial dimensions have few nm. If at a scale of few microns any electronic device can be described by macroscopic physical equations such as Ohm's law, at the nanoscale, microscopic equations are replaced by equations based on quantum mechanics. Quantum mechanical effects manifest at the nanoscale even at room temperature.

A homogenous semiconductor has a conduction band (the first empty band), a valence band (the last occupied band), and a bandgap that separates them. The distribution function of charge carriers in these bands is described by the Fermi–Dirac function

$$f(E) = 1/\{1 + \exp[(E - E_F)/k_B T]\}, \tag{1.1}$$

where E_F is the Fermi energy level. In semiconductors, the Fermi level is located inside the energy bandgap. In Fig. 1.2, we have displayed the Fermi function at two temperatures.

D. Dragoman and M. Dragoman, *Bionanoelectronics*, NanoScience and Technology,
DOI 10.1007/978-3-642-25572-4_1, © Springer-Verlag Berlin Heidelberg 2012

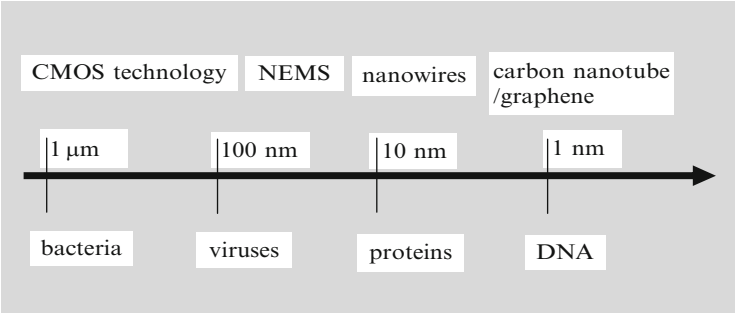
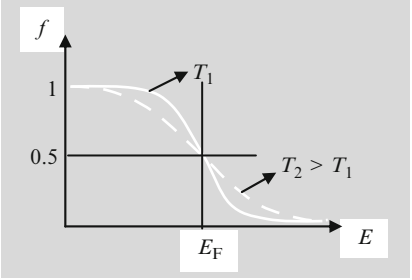


Fig. 1.1 Dimension scale of biological systems and electron devices, where NEMS stands for nanoelectromechanical systems

Fig. 1.2 The Fermi–Dirac distribution function



In the case of nanoscale devices, the confinement of carrier wavefunctions produces a discretization of the energy spectrum of charge carriers as well as discontinuities in the density of states. These effects cause further important changes in the transport properties of charge carriers depending on the number of dimensions along which the motion of carriers is restricted.

In bulk materials with dimensions of few millimeters, the transported carriers move randomly due to repeated scatterings with impurities and phonons. The carrier transport is thus of a diffusive type, which is modeled in general by a stochastic Boltzmann equation. The Boltzmann equation loses its validity as soon as the dimensions of the material shrink to nanoscale. The nanoscale is often termed as mesoscale since it is intermediate between the macroscopic scale and the atomic scale, where the atoms and molecules with sizes of the order of 1 Å $= 10^{-10}$ m are described by quantum mechanical laws.

At the nanoscale, the electron transport is dictated by the relation between the dimensions of the sample and three parameters (Datta 1997):

1. The mean-free path L_{fp}, which is the average distance between two electron collisions with phonons or impurities that cancel the initial momentum of a charge carrier.
2. The phase relaxation length L_{ph}, which represents the propagation distance after which the electron coherence, i.e., the phase memory of electrons, vanishes as a result of time-reversal breaking. Examples of such processes are

electron–electron collisions, dynamic scatterings, or certain impurity scattering processes in which an internal degree of freedom changes; the phase relaxation length is often called the coherence length.

3. The electron Fermi wavelength, denoted as λ_F.

When one or more dimensions of a device are smaller than the mean-free path and the phase relaxation length, the number of scatterings reduces dramatically, and the transport in the device is termed ballistic. In this case, the electrons behave no longer as particles but as waves that follow all the reflection and refraction rules of common light or acoustic waves. As will be seen later, the ballistic transport manifests over distances of few hundreds of nanometers in carbon nanotubes (CNTs), graphene, or high-mobility transistors at room temperature.

In Fig. 1.3, we have schematically displayed a transistor with scaled down dimensions. The transport is diffusive when the transistor has gate lengths of 1 μm or greater and ballistic as soon as the gate channel shrinks to tens of nanometers. In the ballistic transport regime, the carriers traverse the gate channel in a much shorter time and with higher speeds.

The transport of ballistic charge carriers with electron effective mass m and constant energy E can be modeled by the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2}\{m^\alpha \nabla[m^\beta \nabla(m^\alpha \Psi)]\} + V\Psi = E\Psi, \tag{1.2}$$

when the coupling phenomena between different electron bands can be neglected (Dragoman and Dragoman 1999). In (1.2), Ψ denotes the envelope electron wavefunction, which has a slow variation over the unit cell of the crystalline lattice and V is the potential energy. The material-dependent parameters α and β are

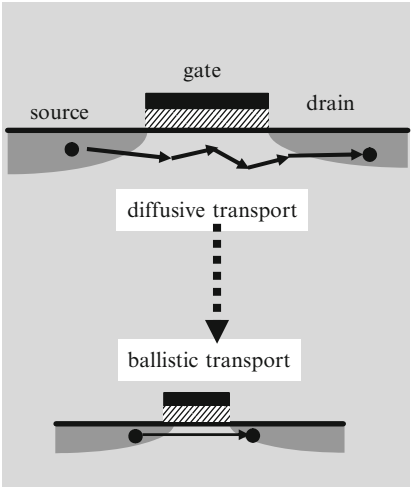


Fig. 1.3 The scaling down of a transistor

related by $2\alpha+\beta=-1$ and are equal to $\alpha=0$ and $\beta=-1$ in AlGaAs compounds, which were the first semiconductors that displayed ballistic transport.

The spatial restrictions on electron motion are expressed in the specific form of the boundary conditions imposed on the Schrödinger equation. A structure in which electrons are confined at the nanoscale by potential barriers along the, say, z direction but are free to travel along the transverse x and y directions is referred to as quantum well (QW). In a quantum well with infinite-height potential barriers, the Schrödinger equation is accompanied by the boundary conditions $\Psi(x,y,0)=\Psi(x,y,L_z)=0$, where L_z is the width of the quantum well.

If $V=0$, the solution of the Schrödinger equation can be written as $\Psi(x,y,z)=(2/L_zL_xL_y)^{1/2}\sin(k_zz)\exp(ik_xx)\exp(ik_yy)$, where L_x and L_y are, respectively, the dimensions of the structure along x and y. In ballistic devices, L_z is comparable to the Fermi wavelength λ_F and $L_z<L_x,L_y\ll L_{fp},L_{ph}$. Another effect of the boundary conditions is a discrete spectrum for the z component of the electron momentum $k_z=p\pi/L_z$, which induces a discretization of the energy levels along the direction of spatial restriction. The energy dispersion relation in the quantum well in which the bottom of the conduction band E_c is considered as reference is given by

$$E(k_x,k_y,k_z)=E_c+\frac{\hbar^2}{2m}\left(\frac{p\pi}{L_z}\right)^2+\frac{\hbar^2}{2m}(k_x^2+k_y^2)=E_{s,p}+\frac{\hbar^2}{2m}(k_x^2+k_y^2), \quad (1.3)$$

where $E_{s,p}$ is the cutoff energy of the discrete subband labeled by the integer p; the subbands are also referred to as transverse modes. The difference in energy between adjacent subbands is greater for more confined electrons, i.e., for smaller L_z.

For an arbitrary energy distribution in the $\mathbf{k}$ space, which takes $E(\mathbf{k})$ constant values on a $\mathbf{k}$-space surface Σ, a spin-degenerate density of states (DOS) can be defined as

$$\rho(E)=(2\pi)^{-3}\int_{\Sigma}\frac{\mathrm{d}\,S}{|\nabla_{\mathbf{k}}E|_{E=\text{const.}}}. \quad (1.4)$$

Then, in the quantum well case, the DOS particularizes to

$$\rho_{QW}(E)=\frac{m}{\pi\hbar^2L_z}\sum\nolimits_p\vartheta(E-E_{s,p}), \quad (1.5)$$

where ϑ denotes the unit step function. As follows from (1.5), and as illustrated in Fig. 1.4, the DOS in the quantum well is discontinuous, in contrast to the case of bulk semiconductors, where the absence of spatial constraints leads to a continuous DOS.

At equilibrium conditions at temperature T, the electrons occupy the discrete energy levels of the quantum well according to the Fermi–Dirac distribution function (1.1), so that the electron density per unit area at equilibrium is given by (Ferry and Goodnick 2009)

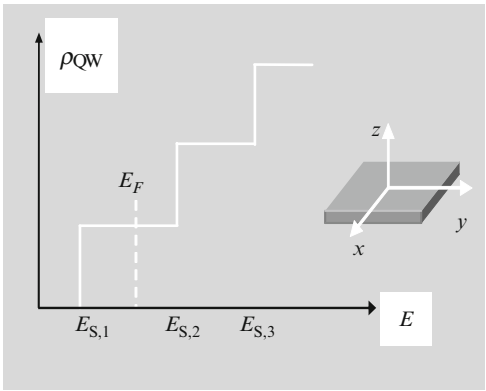


Fig. 1.4 Density of states in a quantum well

$$n = L_z \int_0^\infty \rho_{QW}(E) f(E) \mathrm{d}E = k_B T \frac{m}{\pi \hbar^2} \sum_p \ln[1 + \exp(E_F - E_{s,p})/k_B T]. \quad (1.6)$$

In the degenerate limit or at low temperatures, when $k_B T \ll E_F$, the Fermi–Dirac distribution function is proportional to $\vartheta(E_F - E)$, so that all electron subbands below the Fermi energy are filled with electrons, and all subbands above it are empty. At low temperatures, the electrons with energy E reside in a number of subbands $M(E)$, which can be determined by counting the transverse modes with cutoff energies below E.

When the Fermi energy level in a quantum well is positioned between the first and the second energy subband, as displayed in Fig. 1.4, we have a two-dimensional electron gas (2DEG), which has a metallic behavior because E_F is inside the conduction band. In this case, the Fermi wavenumber k_F, determined from the electron kinetic energy as $E_{kin} = E_F - E_{s,1} = \hbar^2 k_F^2/2m$, is correlated to the electron density per unit area $n = (m/\pi\hbar^2)(E_F - E_{s,1})$ through the formula (Ferry and Goodnick 2009)

$$k_F = (2\pi n)^{1/2}. \quad (1.7)$$

The Fermi wavelength is defined as $\lambda_F = 2\pi/k_F$.

A nanoscale structure is called quantum wire (QWR) if the electron motion is spatially restricted by energy potentials in regions of widths L_y and L_z along two directions: y and z, but the electron can move freely along x. If the constraining potentials have infinite heights, the electron wavefuntion has the expression $\Psi(x, y, z) = [2/(L_y L_z L_x)^{1/2}] \sin(k_y L_y) \sin(k_z L_z) \exp(i k_x x)$, and boundary conditions similar to those in the quantum well case imply that $k_y = p\pi/L_y, k_z = q\pi/L_z$, with p, q integer numbers. The energy dispersion relation in QWR is then given by

$$E(k_x, k_y, k_z) = E_c + \frac{\hbar^2}{2m}\left(\frac{p\pi}{L_y}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{q\pi}{L_z}\right)^2 + \frac{\hbar^2 k_x^2}{2m} = E_{s,pq} + \frac{\hbar^2 k_x^2}{2m}, \quad (1.8)$$

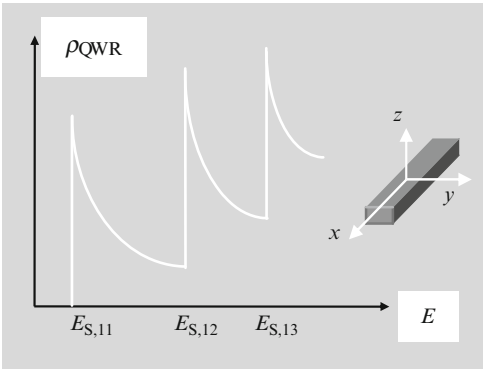


Fig. 1.5 The DOS of QWR

so that the DOS becomes

$$\rho_{\mathrm{QWR}}(E) = \frac{(2m)^{1/2}}{\pi\hbar L_y L_z} \sum_{p,q} (E - E_{\mathrm{s},pq})^{-1/2}. \tag{1.9}$$

The DOS of QWR is represented in Fig. 1.5.

Similar to the optical waveguides, the QWRs with $L_y, L_z < L_x \ll L_{\mathrm{fp}}, L_{\mathrm{ph}}$ and $L_y, L_z \cong \lambda_{\mathrm{F}}$ are called electron waveguides if the energy difference between adjacent subbands is higher than the thermal energy $k_{\mathrm{B}}T$ and the possible potential drop along the waveguide, eV, where V is the applied bias (Dragoman and Dragoman 2004). At temperatures around 4 K, this condition is satisfied in modulation-doped AlGaAs/GaAs heterostructures for $L_y, L_z \cong 0.1$–0.5 μm and $V < 1$ mV.

In a quantum dot (QD), the motion of charge carriers is spatially constrained along all three directions in regions much smaller than the mean-free path and the phase relaxation length. In this case, the discrete energy dispersion is given by

$$E(k_x, k_y, k_z) = E_{\mathrm{c}} + \frac{\hbar^2}{2m}\left(\frac{p\pi}{L_x}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{q\pi}{L_y}\right)^2 + \frac{\hbar^2}{2m}\left(\frac{r\pi}{L_z}\right)^2 = E_{\mathrm{s},pqr}, \tag{1.10}$$

and the DOS is proportional to the Dirac function

$$\rho_{\mathrm{QD}} \propto \delta(E - E_{\mathrm{s},pqr}). \tag{1.11}$$

The DOS of QD is shown in Fig. 1.6. The discrete energy spectrum of quantum dots is similar to that of atoms or molecules, and therefore, sometimes, QDs are referred to as artificial atoms.

In the ballistic regime, only electrons with energies around E_{F} take part at transport, whereas in the diffusive regime, electrons with a wide energy spectrum contribute to electrical transport.

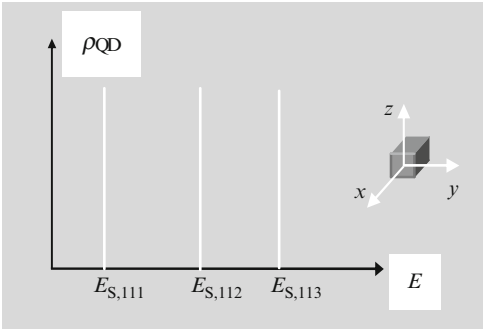


Fig. 1.6 The DOS of QWR

When a one-dimensional (1D) ballistic conductor is situated between two contacts with the role of electron reservoirs, an external bias V induces the electron transfer and imposes a nonequilibrium regime, with no common Fermi energy level across the structure. However, it is possible to define a spatially varying local quasi-Fermi level, with values E_{FL} and E_{FR}, respectively, in the left and right contacts. Then, in a ballistic quantum wire conductor, supposing that $E_{\mathrm{FL}} > E_{\mathrm{FR}}$ and that the contacts are reflectionless, we have $eV = E_{\mathrm{FL}} - E_{\mathrm{FR}}$ if the bias is not very large. At zero temperature, the electrons participating at current flow have energies only in the $E_{\mathrm{FR}} < E < E_{\mathrm{FL}}$ interval. Moreover, for a ballistic conductor with unchanging cross section, in which no electron scattering between different subbands occurs, each occupied subband adds a term of $I = ev\delta n$ to the total net current, where $\delta n = (\mathrm{d}n/\mathrm{d}E)eV$ denotes the additional electron density in the left contact and $v = \hbar^{-1}(\mathrm{d}E/\mathrm{d}k)$ is the velocity of electrons along the direction of current flow. The total current is then $I = (2e^2/h)MV$ if the number of subbands $M(E)$ does not change across the energy range $E_{\mathrm{FR}} < E < E_{\mathrm{FL}}$. In this case, the conductance is given by

$$G = I/V = 2e^2M/h \tag{1.12}$$

and is an integer multiple of $G_0 = 2e^2/h$, called quantum conductance. In the ballistic regime, the resistance $R = 1/G = 1/(MG_0) \cong 12.9\,\mathrm{k\Omega/M}$ originates in the difference at the conductor/contact interface between the infinite number of subbands in the contacts and the finite number of transverse modes in the conductor. So R is termed contact resistance, and its value becomes increasingly smaller than the quantum value $R_0 = 12.9\,\mathrm{k\Omega}$ as the number of occupied energy subbands in the conductor raises.

In bulk materials, Ohm's law states that the conductance G is inversely proportional to the length of the sample. In deep contrast, in ballistic structures, the conductance does not depend on the conductor length but only on its width W since the number of subbands that are occupied by electrons with the Fermi wavenumber k_{F} is given by $M \cong \mathrm{Int}[k_{\mathrm{F}}W/\pi]$, where $\mathrm{Int}[x]$ symbolizes the integer value of the argument x.

The step-like conductance dependence on the number of occupied subbands in (1.12) was demonstrated experimentally by low-temperature measurements on a

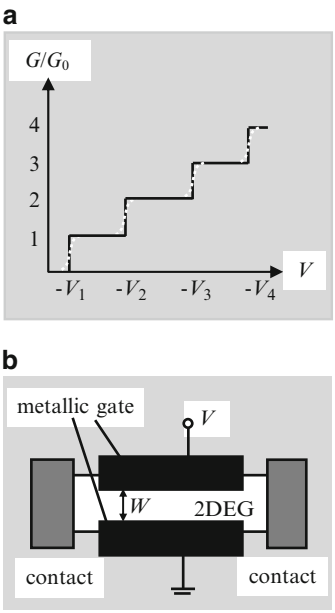


Fig. 1.7 (**a**) The gate voltage dependence of conductance at zero temperature (*solid line*) and finite temperatures (*dotted line*) in (**b**) a 1D ballistic conductor with tunable width

ballistic wire with a variable width, as illustrated in Fig. 1.7a (del Alamo et al. 1998). The conductance rises in steps of G_0 for any increases with one unity of M, this stair-like shape being "smoothed" as the temperature increases due to thermal vibrations. A split-gate geometry as that displayed in Fig. 1.7b must be used to obtain a 1D ballistic quantum wire from a 2DEG. In this geometry, a narrow slit, which has a width W on the order of λ_F, is cut in a depleting gate patterned above the 2DEG. The effective width W of the conductor can be decreased by applying a gradually increasing negative gate voltage V, so that the number of transverse modes M is modified in a stepwise manner. The split-gate configuration is termed quantum point contact if the constriction length is small enough (in fact, is comparable to its width) and is referred to as electron waveguide if the constriction is much longer than W.

The hypothesis made in deriving (1.12), i.e., that all electrons originating in the left contact reach the right contact, is not always valid. For instance, if the ballistic conductor is composed of several sections with different potential energies or widths, the electrons from the left contact are only partially transmitted to the right contact. If T designates the transmission probability of the ballistic conductor attached to reflectionless contacts via ballistic leads, as illustrated in Fig. 1.8, the zero-temperature conductance between the contacts is given in this situation by the Landauer formula (Datta 1997)

$$G = \frac{2e^2}{h} MT, \tag{1.13}$$

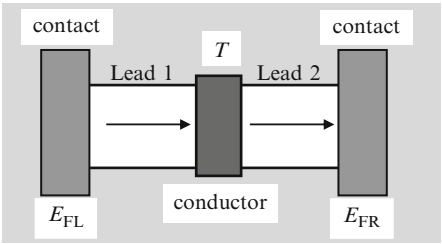


Fig. 1.8 The model of a ballistic conductor

where M is the number of subbands in the leads. The current that flows between contacts at low temperatures is then $I = (2e^2/h)MTV$.

In this case, the total resistance between contacts, $R = h/(2e^2MT)$, can be written as a sum between $h/(2e^2M)$, which has the meaning of contact resistance, and $R_s = h(1-T)/(2e^2MT)$, which is the resistance of a scatterer with transmission T. In the same way, the resistance of a succession of scatterers with transmissions T_i is $R_s = \sum_i R_{s,i}$, where $R_{s,i} = h(1 - T_i)/(2e^2MT_i)$. This formula suggests that a series of scatterers is equivalent to a single scatterer with total transmission probability given by $(1 - T)/T = \sum_i (1 - T_i)/T_i$, expression that results from adding all partially transmitted waves.

We have assumed up to now that the Fermi–Dirac distribution $f(E)$ can be approximated with a step function, but this assumption does not hold at higher temperatures (see Fig. 1.2), case in which the electrons contributing at electrical conduction have energies in the range $E_{FR} - \Delta E < E < E_{FL} + \Delta E$, where ΔE is a few $k_B T$. Then, the current flowing between the left and right contacts, which have respective Fermi–Dirac quasi-distribution functions $f_L(E)$ and $f_R(E)$, is given by Datta (1997)

$$I = \frac{2e}{h} \int M(E)T(E)[f_L(E) - f_R(E)]\mathrm{d}E. \tag{1.14}$$

From (1.14), it follows that the calculation of the transmission probability is essential for current estimation using the Landuaer formula. The transmission probability can be computed using the transfer Hamiltonian formalism, the Green's function approach, or the Kubo formalism (Datta 1997; Ferry and Goodnick 2009), but the easiest method to determine it is the matrix formalism, described below.

The transmission probability can be calculated once the electron wavefunctions are known. In the simplest case, when the 1D ballistic conductor is composed of a succession of several regions with constant but different electron effective masses and potential energies, which extend along the x direction, the solution of the Schrödinger equation (1.2) in the ith region, $\Psi_i(x) = A_i \exp(ik_i x) + B_i \exp(-ik_i x)$, can be regarded as a superposition of waves that propagate forward and backward with wavenumbers $k_i = \hbar^{-1}\sqrt{2m_i(E - V_i)}$. Continuity conditions require that at each interface between layers i and $i + 1$, situated at $x = x_i$, as shown in Fig. 1.9, the wavefunction and $(\partial\Psi/\partial x)/m^{\alpha+1}$ are constant.

If $\alpha = 0$ in (1.2), these requirements connect the wavefunction components on each side of the interface via a transfer matrix

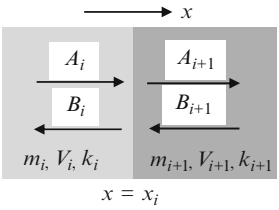


Fig. 1.9 Forward- and backward-propagating components of the electron wavefunction at an interface between adjacent layers i and $i + 1$

$$\begin{pmatrix} A_i \exp(ik_i x_i) \\ B_i \exp(-ik_i x_i) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} (1+v_{i+1}/v_i) & (1-v_{i+1}/v_i) \\ (1-v_{i+1}/v_i) & (1+v_{i+1}/v_i) \end{pmatrix} \begin{pmatrix} A_{i+1} \exp(ik_{i+1} x_i) \\ B_{i+1} \exp(-ik_{i+1} x_i) \end{pmatrix}, \tag{1.15}$$

where $v_i = \hbar k_i / m_i$ is the electron velocity in layer i. Analogously, the 2×2 transfer matrix for free propagation across the ith layer between planes situated at $x = x_{i-1}$ and $x = x_i$ is diagonal, with nonvanishing elements $\exp[ik_i(x_i - x_{i-1})]$ and $\exp[-ik_i(x_i - x_{i-1})]$. As a result, the transmission probability for a structure composed of a succession of N layers is given by

$$T = v_N |A_N|^2 / (v_1 |A_1|^2) = v_N / (v_1 |M_{11}|^2) \tag{1.16}$$

and is directly dependent on the element M_{11} of the total transfer matrix with elements M_{pq}, $p,q = 1, 2$, determined by multiplying the matrices corresponding to each interface and each layer. The transmission probability is the ratio between transmitted and incident electron probability currents, defined in any layer j as $J_j = (\hbar / 2m_j i)(\Psi_j^* \partial_x \Psi_j - \Psi_j \partial_x \Psi_j^*)$, where * signifies complex conjugation and ∂_x is a shorthand notation for partial derivation with respect to x.

When some of the wavenumbers k_i are imaginary, we encounter the tunneling phenomenon, widely used in nanodevices, and which can be modeled using the matrix formalism described above. The wavenumber is imaginary only if the electron energy is smaller than the potential energy, situation in which electron propagation is not allowed from a classical point of view. Therefore, tunneling is a quantum phenomenon. The layer with an imaginary wavenumber is a barrier for electron propagation, and the electron wavefunction decays exponentially inside it, analogous to evanescent electromagnetic waves. Because the transmission probability across a barrier layer becomes zero unless it is narrow enough, tunneling of electrons propagating with constant energy E occurs only through thin potential barriers or a succession of such barriers surrounded by quantum wells, i.e., regions with real wavenumbers.

However, classical transport across a potential barrier is permitted if the electron acquires extra energy. When this additional energy is thermal, the process is called thermionic emission. At finite temperatures, thermionic emission accompanies the tunneling process and becomes the major electron transport mechanism at high temperatures (Appenzeller et al. 2004). The thermionic emission contributes to the net current across a rectangular barrier with a temperature-dependent term $I \propto T^2 \exp(-\phi / k_B T)$, where ϕ is the barrier height.

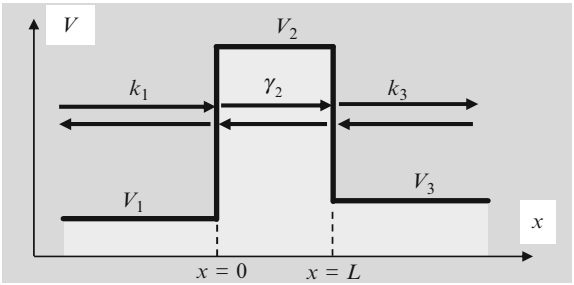


Fig. 1.10 The geometry of a tunneling structure consisting of a single barrier

In tunneling devices, the direction of electron propagation, denoted by x in the expressions above, is determined by the direction of the applied electric field, and the transfer matrix method remains applicable as long as Ψ denotes the x-dependent part of the envelope wavefunction of electrons and the electron motion along the longitudinal x direction can be separated from that along the transverse y and z directions. The transmission probability through a single barrier, which is the simplest tunneling structure displayed in Fig. 1.10, is given by

$$T = \frac{4v_1v_3}{(v_1+v_3)^2 + [(v_1^2+v_2^2)(v_2^2+v_3^2)/v_2^2]\sinh^2(\gamma_2 L)}, \tag{1.17}$$

where $v_2 = \hbar\gamma_2/m_2$ and $k_2 = -i\gamma_2$ is imaginary in the barrier region labeled by 2 and surrounded by regions 1 and 3 with real wavenumbers.

From (1.17) it follows that, if $\gamma_2 L \gg 1$, the transmission probability decreases exponentially with the barrier width L: $T \propto \exp(-2\gamma_2 L)$. At low temperatures, when all electrons contributing to transport have energies around E_F, this expression can be rewritten as $T \propto \exp[-2L\sqrt{2m_2(V_2 - E_F)}/\hbar]$ (Zhu 2001). In this case, the current density through the structure is proportional to T and therefore decreases also exponentially with L, while the conductance is much smaller than G_0.

The example of electron transmission through a single barrier is relevant for many applications, such as scanning tunneling microscopy or vacuum microelectronic devices, which contain flat panel field emission displays and electron sources for microscopes. In these devices, electrons tunnel from a solid surface into vacuum under the influence of a high electric field (Zhu 2001).

The potential barrier in the vacuum region, given by the electron affinity ϕ, takes a triangular shape in the presence of an applied electric field F, acquiring a spatial dependence of $\phi - eFx$ (see Fig. 1.11). The wavefunction of the electron in vacuum with mass m_0 is then a solution of the Schrödinger equation with a triangular barrier, which, if $\alpha = 0$, is a superposition of Ai and Bi Airy functions: $\Psi_{vac}(\xi) = A_{vac}\mathrm{Ai}(\xi) + B_{vac}\mathrm{Bi}(\xi)$, with $\xi = [2m_0/(eF\hbar)^2]^{1/3}(\phi - eFx - E)$. The condition of current flow along the positive x direction imposes the following values for the constant coefficients: $A_{vac} = 1$, $B_{vac} = i$, so that the transmission coefficient T across the barrier has the expression

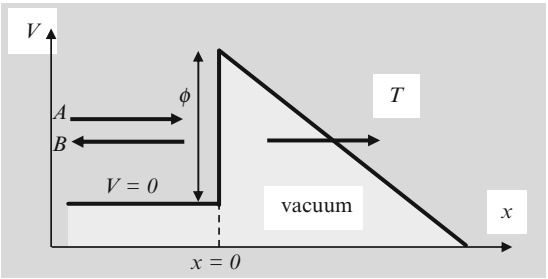


Fig. 1.11 Geometry of electron emission in vacuum by tunneling through a triangular potential barrier

$$T = \left(\frac{2eF}{\hbar^2 m_0^2}\right)^{1/3} \frac{m}{\pi} \frac{|A|^2}{k} \tag{1.18}$$

for incident electrons with energy E, mass m, and wavefunction $\Psi(x) = A\exp(ikx) + B\exp(-ikx)$. In (1.18), $k = (2mE/\hbar^2)^{1/2}$ and the coefficients A and B are obtained from the continuity conditions imposed on the wavefunction and its derivative at $x = 0$.

At low temperatures and high electric fields, and for $m = m_0$, the current density associated to the transmission coefficient in (1.18) is

$$J = \frac{em_0 k_B T}{\pi^2 \hbar^3} \int T(E) \ln\left[1 + \exp\left(\frac{E_F - E}{k_B T}\right)\right] dE \propto F^2 \exp\left(-\frac{4}{3\hbar F}\sqrt{2m_0\varphi^3}\right), \tag{1.19}$$

where the logarithmic term describes the effect of the transverse degrees of freedom on the Fermi–Dirac distribution function. Formula (1.19) is the so-called Fowler–Nordheim equation, which models the current–voltage characteristic of devices based on field emission.

Let us consider now the case when electrons tunnel through a structure consisting of two (or several) barriers that alternate with quantum well regions. Because ballistic electrons propagate coherently, constructive or destructive interferences appear between quantum waves that are only partially reflected and transmitted at different interfaces. So we expect high and low transmission probability values in a typical geometry such as that displayed in Fig. 1.12, containing a quantum well surrounded by two thin barriers. This phenomenon is analogous to the occurrence of high- and low-intensity values associated with interference between coherent light beams. A large transmission probability through a structure consisting of two (or several) barriers, each of them with a low transmission probability, defines the phenomenon of resonant tunneling.

This phenomenon can be described also with the transfer matrix method developed above. For the case of a structure containing two barriers that surround a quantum well with width L in which electrons propagate with a wavevector k, the total transmission probability T is determined by the M_{11} element of the transmission matrix M through the whole structure, found by first multiplying the transmission matrix M_L of the left barrier with the diagonal matrix with elements $\exp(-ikL)$ and $\exp(ikL)$, corresponding to free propagation across the quantum well,

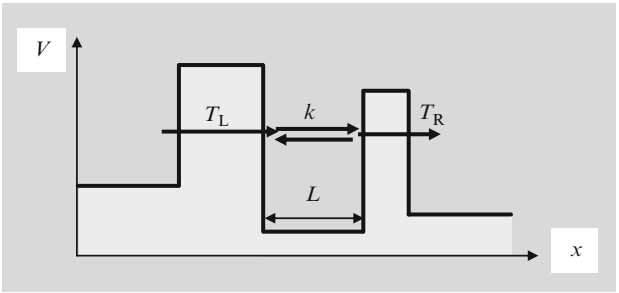


Fig. 1.12 Geometry of the double-barrier tunneling structure

followed by multiplication with the transmission matrix M_R of the right barrier. The final result is $M_{11} = M_{L,11}M_{R,11}\exp(-ikL) + M_{L,12}M_{R,21}\exp(ikL)$, where $M_{L,11}$ stands for the 11-element of the M_L matrix, which gives

$$T = \frac{T_L T_R}{(1-\sqrt{R_L R_R})^2 + 4\sqrt{R_L R_R}\cos^2\theta}. \tag{1.20}$$

In (1.12), T_L and T_R are, respectively, the transmission probabilities through the left and right barriers, $R_L = 1 - T_L$ and $R_R = 1 - T_R$ denote the related reflection probabilities, and

$$\theta = kL + (\arg M_{L,12} + \arg M_{R,21} - \arg M_{L,11} - \arg M_{R,11})/2.$$

Equation (1.20) suggests that T can be close to unity if the energy resonance requirement $\theta = (2n+1)\pi/2$ holds, with n integer, even if T_L and T_R have small values, for which $1 - \sqrt{R_L R_R} \cong (T_L + T_R)/2$. The resonance transmission probability, $T_{res} = 4T_L T_R/(T_L + T_R)^2$, can become unity if $T_L = T_R$, irrespective of the values of T_L and T_R, or can be approximated with $4\min(T_L, T_R)/\max(T_L, T_R)$ if T_L and T_R differ significantly. At resonance, when the energy of incident electrons E equals a resonant value E_{res}, besides being transmitted with high probability, the electron travels through the structure faster than in off-resonance conditions. In general, the devices based on resonant tunneling are ultrafast. Around resonance, we have

$$T(E) \cong \frac{\Gamma_L \Gamma_R}{(\Gamma_L + \Gamma_R)^2/4 + (E - E_{res})^2}, \tag{1.21}$$

where $\Gamma_L = (\mathrm{d}E/\mathrm{d}\theta)T_L/2$ and $\Gamma_R = (\mathrm{d}E/\mathrm{d}\theta)T_R/2$ (divided by $\hbar$) are, respectively, the rates at which an electron situated in the well leaks out of it through the left and right barriers. Close to resonance, T is very sensitive to the values of E and E_{res}. The last parameter can be easily tuned by applying a bias across the structure.

It is quite remarkable that, although the coherent nature of the electron transport is explicitly used in deriving (1.20), which implies that electrons are transmitted in

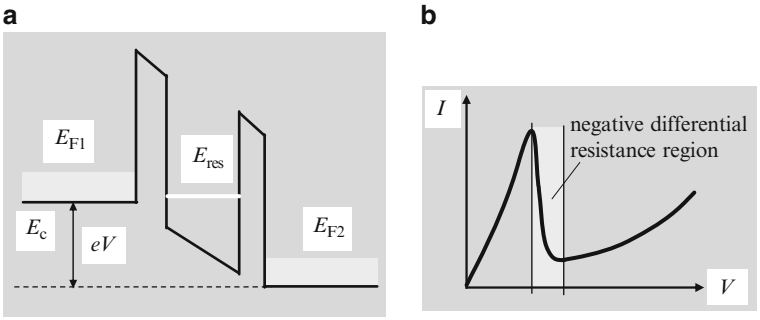


Fig. 1.13 (**a**) Energy band representation of a resonant tunneling diode, and (**b**) a typical current–voltage characteristic

a single quantum process through the whole structure, this expression can be used also if sequential tunneling occurs.

The most notorious device based on the resonant tunneling phenomenon is the resonant tunneling diode (RTD), displayed in Fig. 1.13. The RTD is a biased double-barrier structure, which has a single resonant energy level E_{res} in the quantum well region that filters the electrons from the left contact (emitter). In consequence, only the electrons with energy E_{res} from all electrons in the emitter with energies between the Fermi energy E_{F1} and the bottom of the conduction band E_c can tunnel toward the contact on the right (collector) with Fermi energy E_{F2}. As a result, the $I - V$ characteristic of the RTD displays a negative differential resistance (NDR) region since E_{res} drops below E_c at a sufficiently high bias and electron tunneling into the collector is no longer possible. Consequently, the current decreases drastically. The NDR of the RTD occurs even at room temperature. RTDs have numerous applications as bistable elements, oscillators, or logic circuit components (Dragoman and Dragoman 2009a).

When a constrained structure such as a quantum well, quantum wire, or quantum dot is positioned in the immediate neighborhood of another constrained structure of the same type, the interaction strength between electrons in the two structures depends on the height and width of the barrier separating the confined structures. The quantum wells (quantum wires or quantum dots) are said to be coupled when the electrons interact and the electron wavefunction expands over the whole structure, whereas when the electrons in neighboring structures do not interact, we have a series of noninteracting structures, called multiple quantum wells (quantum wires or quantum dots). In the last case, the electrons localized in one quantum well (quantum wire or quantum dot) are transferred to the other confined structure only through sequential tunneling. Nanoscale structures separated by finite barrier regions are coupled by the exponentially decaying electron wavefunction in the barriers, which allows overlapping of the envelope electron wavefunctions in neighboring wells if the potential barriers are thin enough.

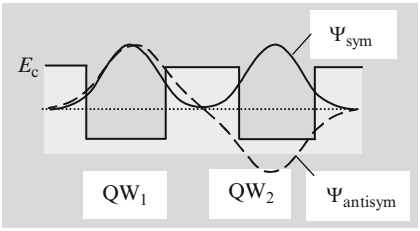


Fig. 1.14 Electron wavefunction splitting into an antisymmetric (*dashed-dotted line*) and a symmetric (*solid line*) part when two identical quantum wells are brought in close proximity. The dashed line represents the coordinate axis

In particular, if two identical quantum wells, denoted by QW_1 and QW_2 in Fig. 1.14, are coupled, the resonant energy levels in individual quantum wells are no longer degenerated, and the electron wavefunction breaks up into a symmetric and an antisymmetric part, Ψ_{sym} and $\Psi_{antisym}$, respectively, which expand over the entire structure. On the contrary, in two nonidentical coupled quantum wells, the electron wavefunction is still confined in one of the two wells except when the different individual resonant energy levels in the wells align in the presence of an applied bias that produces a resonance in the structure. Only in this last situation the electron wavefunction delocalizes and breaks up into a symmetric part and an antisymmetric part.

A generalization of the previous example involves several identical quantum wells, which form a periodic structure with period Λ and are brought in close proximity. Then, they become coupled, and the delocalized electrons experience a periodical energy potential distribution with the same period Λ. In particular, in a quantum structure consisting of N periods, each N-fold degenerate energy level in the quantum wells splits, so that each well provides one state to the resulting energy bands containing discrete levels.

In the limiting case of quantum structures with a large number of periods and sufficiently narrow barriers, the allowed and forbidden electron energy bands become almost continuous, analogous to bulk materials, in which electrons feel the periodicity of the crystalline lattice; these energy bands of a periodic structure are illustrated in Fig. 1.15. Such a quantum structure that can be assimilated with an artificial lattice with a controllable unit cell is a superlattice. The widths and positions of the allowed and forbidden electron energy bands depend on the spatial distribution of the periodic potential, which can be modulated at will using advanced semiconductor techniques.

The prospect of tuning the energy dispersion of electrons in a semiconductor superlattice is especially useful in optoelectronics when the period of the superlattice is equal to the wavelength of the electromagnetic field. In this situation, allowed and forbidden frequency bands develop also for the electromagnetic radiation, analogous to the case of electrons in superlattices or bulk materials; the position and width of these bands are controllable. A comprehensive theory of electron

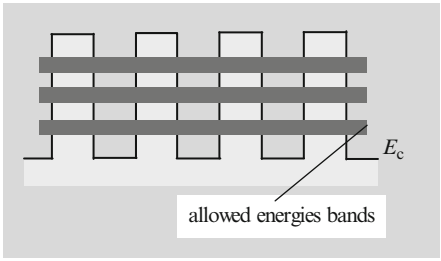


Fig. 1.15 Allowed and forbidden electron energy bands in the periodic potential of a superlattice

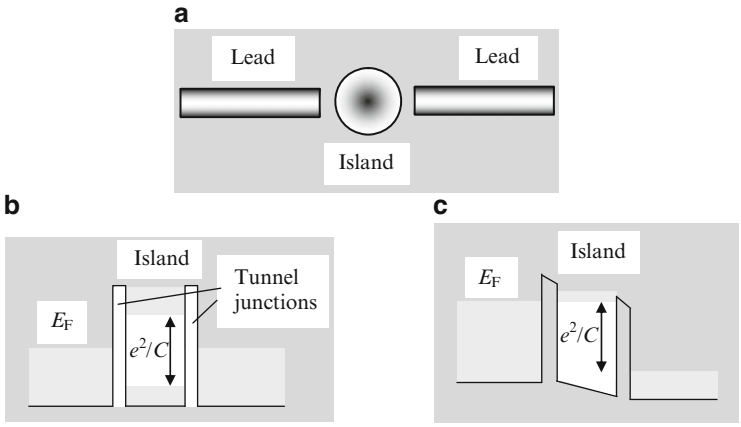


Fig. 1.16 (**a**) Schematic representation of an island connected to metallic leads through tunneling junctions. (**b**) Energy gap opening at the Fermi level due to the Coulomb energy needed to tunnel in or out of the island, (**c**) which is overcome by a bias $V = e/C$

transport in superlattices that involves nonequilibrium Green functions can be found in Wacker and Jauho (1998).

Another interesting quantum phenomenon is encountered in tightly confined nanostructures, in particular in quantum dots. Here, the Coulomb interaction is considerable, and the electronic states depend on the number of electric charges (electrons) in the dot, which takes discrete values. A manifestation of this dependence is illustrated by the Coulomb blockade, in which an energy gap opens at the Fermi level in the electron energy spectrum in small metallic clusters or semiconductor quantum dots coupled by tunneling barriers to metallic leads; these confined structures are generically called islands. The configuration of a Coulomb blockade device is schematically represented in Fig. 1.16a (Ferry and Goodnick 2009). This gap in the energy spectrum, analogous to the energy gap in semiconductors, is modeled as an additional energy needed by an electron to tunnel in or out of the island, which originates in the Coulomb interaction between electrons in the island.

This additional energy is equal to $e^2/2C$ in metallic islands, where C is the capacitance between the island and the surrounding medium, and leads to a gap of e^2/C in the electron energy spectrum at the Fermi level since, in addition to electrons, holes need the same additional energy $e^2/2C$ to tunnel in or out of the island. The Coulomb blockade phenomenon is observed when $e^2/C \gg k_B T$, which implies low temperatures, and when the electron number on the island is constant. The last condition is fulfilled if e^2/C is much larger than the lifetime broadening $\hbar/\tau$, where τ denotes the electronic lifetime. Alternatively, in terms of an effective RC time, this requirement can be written as $R \gg h/e^2$, which means that the island is decoupled from the reservoirs through tunneling barriers with much larger resistances in comparison to the quantum resistance.

As displayed in Fig. 1.16b, c, in the presence of Coulomb blockade, an electron can tunnel into a metallic island only if a large enough bias is applied, $V > e/C$, which overcomes the charging energy. As a result, in the $I - V$ characteristic of the island, the current has very low values around zero bias. On the contrary, when $V = e/C$ a single electron from one lead can tunnel in the island, leading to an increase of the Fermi energy in the island by e^2/C, and a subsequent tunneling event is forbidden by the opening of another energy gap, unless the applied bias raises to $V > 3e/C$ or the additional electron in the island tunnels out of it. Summarizing, the average electron number in the island augments with one whenever the voltage increases with $2e/C$. As a consequence of these correlated tunneling events into and out of the island, the net current increases, and the $I - V$ characteristic acquires a staircase shape when the two junctions differ significantly (Ferry and Goodnick 2009).

The Coulomb blockade phenomenon in ballistic semiconductor islands can be treated in an analogous manner except that the quantization of energy must be included in the model (size quantization effects are not relevant in metallic clusters because the conditions for ballistic transport are not satisfied, in general). In this case, the additional energy needed to add an electron to the island is $e^2/C + \Delta E$, where ΔE is the difference between the energies of adjacent discrete quantum states.

The single-electron Coulomb blockade phenomenon is caused by the discreteness of electric charge that can be transferred to and from a conducting island connected through thin barriers to electron reservoirs. On the contrary, resonant tunneling devices are based on the discreteness of resonant energy levels in a quantum well coupled to electron reservoirs through thin barriers. The Coulomb blockade effect controls precisely the (small) number of electrons in the island and is employed in low-power switching devices that are essential for an increased level of circuit integration. In general, single-electron devices based on Coulomb blockade have an additional control of the electric charge in the island via a gate electrode, which leads to periodic current oscillations through the leads as the gate voltage is modified.

1.2 Nanotechnologies for Bionanoelectronic Devices

The nanotechnologies for bionanoelectronic devices rely on the technologies used in electronics to produce complex circuits based on AIII-BV semiconductors and micrometer and submicrometer-scale Si integrated circuits. Silicon is still the main material for micro- and nanoscale devices, but carbon allotropes, AIII-BV semiconductors, and biomolecular assemblies could end its dominance. There are two main approaches for nanotechnologies: top-down and bottom-up.

In the top-down approach, the major steps are (1) deposition of a single or multiple layers on a substrate, followed by (2) the transfer onto the layers grown in the first step of desired patterns, process that involves in turn several steps to remove unnecessary materials that are not included in the desired pattern. The first major step is called deposition technique, whereas the second is called lithography or nanolithography when the desired patterns have nanosized features.

The bottom-up nanotechnology techniques have no counterpart in the electronic technologies. For example, in the self-assembly technologies, which are mask-less techniques, the mask manufacturing process characteristic for the top-down approach is replaced by less expensive specific chemical reactions able to form desired three-dimensional (3D) patterns consisting of metallic or semiconducting nanosized materials. However, it is not uncommon to fabricate nanoelectronic devices, especially biosensors, by a combination of bottom-up and top-down approaches.

1.2.1 Deposition Techniques for Bionanoelectronic Devices

In this section, we examine briefly the most important deposition techniques for micro- and nanoscale devices. An extended review with comprehensive references can be found in Ziaie et al. (2004).

To start with, the fabrication of SiO_2 by oxidation of silicon is an essential process at both micro- and nanoscale, although other oxides, such as HfO_2, are increasingly studied. SiO_2 is a dielectric material used to isolate a variety of metallic electrodes from other conductive substrates and can be deposited on silicon substrates with thicknesses ranging from few nm up to 2 μm. In particular, SiO_2 is encountered in many transistors and micro-electro-mechanical systems (MEMS) or nano-electro-mechanical systems (NEMS), for example, switches. It is grown in the presence of oxygen or water at temperatures of about $1,000$–$1,200$°C into furnaces containing a quartz tube, an electrical resistance heater, and a wafer holder. The thickness of the SiO_2 depends on the temperature and the gas flow.

Another basic technological process in microelectronics, which is associated to deposition, is semiconductor chemical doping, performed with the aim of changing dramatically its electrical, optical, or mechanical characteristics. The process of changing these characteristics is generically referred to as functionalization or

material engineering. At the microscale, chemical doping is used to fabricate active electronic devices, such as diodes and transistors, at very large scale, and MEMS devices with specific mechanical characteristics (Dragoman and Dragoman 2001). Chemical doping of intrinsic silicon relies on a controllable introduction of p- or n-type impurities via ion implantation techniques or high-temperature diffusion from liquid or solid sources. At nanoscale, although chemical doping is still used in CNT transistors, for example, the diversity of functionalization techniques is greatly enhanced. For instance, nanowires or quantum dots can be functionalized also via electric fields, hydrogenation, oxygenation, adsorption of molecules or biomolecules, or other methods discussed throughout the book.

Deposition techniques for nanosized devices are based on chemical vapor deposition (CVD), in which thin films, with thicknesses up to fractions of nm, are deposited on a substrate using chemical reactions of specific gaseous components. In general, these chemical reactions need high amounts of energy generated by (1) plasma excitation, (2) optical excitation, or (3) heating the substrate at very high temperatures. The first two processes work at lower temperatures than the last one.

CVD techniques comprise two basic methods: low-pressure CVD (LPCVD) and plasma-enhanced CVD (PECVD). LPCVD requires electrical heated furnaces where a very low pressure (at 0.1–0.7 torr) is maintained by a pumping system. Thin films of a certain material are deposited on both faces of a wafer positioned on a holder inside the furnace. For instance, thin films of SiO_2 (using as gaseous components N_2O and $SiCl_2H_2$ at 900°C), polysilicon (from gaseous SiH_4 at 600°C), or Si_3N_4 (from gaseous components NH_3 and SiH_4 at 800°C), as well as Ti, Mo, Cu, and Ta metallic thin films can be deposited using LPCVD.

PECVD is based on the plasma produced by a high-power radio frequency (RF) source. The main advantage of PECVD is the lower temperature (100–300°C) needed to heat the substrate. PECVD consists of a plasma reactor with a RF source; a pumping system, which injects gases inside the reactor and produces a high vacuum inside the reactor chamber; and two parallel plates. On one plate, the RF signal is applied, and on the other, which contains also the wafer placed above an electrical heater, the RF generator is grounded. PECVD is used to deposit Si_xN_y (the nonstoichiometric form of silicon nitride), amorphous silicon, SiO_2, and CNTs produced from a mixture of NH_3 and C_2H_2 gases.

The CVD-based epitaxial techniques boosted the semiconductor industry because they made possible the growth of monolayers of several AIII-BV semiconductor heterostructures, such as InP/GaInAs or GaAs/AlAs, which are widely used in the advanced nanoelectronic devices based on quantum wells, wires, and dots. In the epitaxial growth based on CVD, a crystalline material is first grown on the substrate, and an additional crystalline material can be subsequently grown on top of it if their lattices match or are slightly different, case in which a strain between the crystals is induced. This strain can be used, for instance, to tune the bandgap of the resulting heterostructure to a desired value corresponding to a specific emission wavelength in quantum well or quantum wire lasers.

Two CVD-based techniques are mainly used to grow heterostructures: metal organic chemical vapor deposition (MOCVD) and molecular beam epitaxy (MBE).

The MOCVD technique relies on vapors of organic compounds with group-V hydrides and group-III atoms, which are placed inside the CVD chamber that can switch gases very fast. The MBE needs a CVD chamber containing a certain substrate in a high-vacuum environment. The substrate is exposed to molecular beams produced by several thermally evaporated basic sources, the beam exposure/growth order depending on the desired heterostructure.

1.2.2 Nanolithography

Nanoscale lithography is a key concept in nanotechnologies, the patterning of different surfaces with current nanolithographical methods attaining a resolution (the smallest achievable feature) of few nm; for an excellent review on nanolithography, see Harriot and Hull (2004).

In principle, lithography refers to the transfer of a desired geometrical pattern (a mask containing the desired features to be imprinted) onto a substrate. Lithography is typically performed by covering the substrate with a resist material, having a thickness of 0.2–2 μm, with the help of a spinner and then illuminating the deposited resist with light or energetic particles, such as electrons or ions. Because the resist is sensitive to illumination, which is made through a mask that transmits the excitation (optical, or particle beams) in wanted spatial regions and blocks it in undesired regions, the required pattern is imprinted in the resist after illumination and the following developing process. More recently, especially in the case of particle lithography, the pattern is directly written in the resist, the procedure being termed maskless lithography. In general, the fabrication of nanoelectronic devices is based on a combination of lithographical techniques at microscale, among which optical lithography is mostly used for the device contacts, for example, and nanolithographical techniques, where the nanosized features of devices are implemented.

In optical lithography, for instance, the mask is fabricated from glass or fused silica plates, which are selectively coated with an absorbing Cr layer with a thickness of 80–100 nm. Mask fabrication is a challenging step when nanoscale resolution and small number of defects are needed. In these situations, the desired pattern is written with a pattern generator into the resist deposited on Cr by a laser beam, process followed by resist developing and etching of the Cr layer.

Typically, the mask is a scaled down version of a pattern designed at a much larger scale, the reduction process having the net advantage that all cut defects and errors are reduced with the same scale as the pattern. Nevertheless, masks directly made at the 1:1 scale are used in X-ray lithography.

In the resist material, selective chemical reactions occur at illumination. Depending on the resist type, positive or negative, the illuminated areas are dissolved and removed in the developing stage of the resist or, respectively, are kept intact, as shown in Fig. 1.17. In both cases, the required pattern is imprinted in the resist, which is deposited on the substrate.

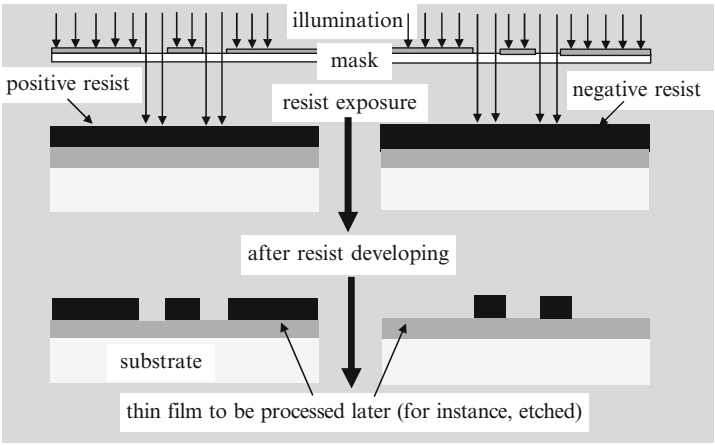


Fig. 1.17 The optical lithography process

If necessary, the lithographical process is repeated several times using various resists and masks, which need to be aligned at each step, the procedure being accompanied by etching and deposition processes in a prescribed order to obtain the final nanodevice or integrated circuit. The principal parameters of the lithographical system are resolution and mask alignment (interlevel alignment values).

The resist materials are generally polymers, such as diazonaphthoquinone (DNQ) or (poly)methylmethacrylate (PMMA) in the optical lithography case, the resist type being a function of the illumination source (particle beams, optical, X-ray). The exposure of the resist is performed in a direct write mode, when it is illuminated point-by-point, through a mask, which is fully exposed, and the pattern is imprinted in the resist by contact or proximity alignment of the resist surface to the mask (see Fig. 1.17), or by projecting the image of the mask. The mask is in direct contact with the resist in the contact exposure method, which assures a good resolution but damages the mask in the process, while in the proximity method, the mask is situated at a distance d_{m-r} from the resist, which determines the minimum feature size:

$$w_{\min} \approx (\lambda d_{m-r})^{1/2}. \tag{1.22}$$

In (1.22), λ denotes the wavelength of the illumination source. In the proximity method, d_{m-r} must be very small to obtain printed feature sizes the same as that on the mask, which is a challenge in nanotechnologies. Therefore, the projection exposure method is the most widespread high-resolution procedure in nanolithography. In this method, an optical system projects the image of the mask on the resist, improving the best resolution, which is given by

$$R = \beta\lambda/\mathrm{NA}, \tag{1.28}$$

where NA is the numerical aperture of the imaging lens and $\beta \cong 0.8$. The ratio β/NA is nearly 1 in the optical projection exposure method, the resolution being directly proportional with the excitation wavelength.

In a typical optical lithographic system, which uses mercury arc lamps with pronounced emission peaks at 365 and 435 nm, only geometrical line widths exceeding 0.25 μm can be obtained. If thinner line widths are needed, excimer lasers such as ArF ($\lambda = 193$ nm), F_2 ($\lambda = 157$ nm), and KrF ($\lambda = 248$ nm) must be used to lower the geometrical line widths up to 0.13 μm.

Higher resolutions, of 30 nm, and interlevel alignments as small as 10 nm can be achieved with extreme ultraviolet lithography (EUV), which uses wavelengths of 10–14 nm. However, there are many difficulties in mask fabrications at these wavelengths, the reflective elements of EUV masks being multilayered Bragg mirrors deposited on Si. A synchrotron or plasma source illuminates the mask, the tolerances of the imaging system in the projection exposure method being of only few angstroms. Despite these difficulties, the EUV displays very high yields, including a high speed of feature patterning, of about 10^{11} features/s. At the extreme end of the electromagnetic spectrum, lithography with X-rays uses high-energy (few keV). X-ray sources such as Cu target systems emitting X-rays or electron synchrotrons. X-ray lithography works in the proximity-printing mode and has a resolution of 50 nm. The main drawback of this method is the 1:1 scale mask fabrication.

The particle nanolithography is based on electrons or ions, the lithographical process displaying very good performances, with resolutions of 10 nm for ions and 50 nm for electron lithography. Even smaller features, of 5 nm, can be obtained using this method.

The electron beam lithography (EBL) technique relies on two configurations. In the first, direct writing EBL, a focalized electron source, is directed to a substrate or a substrate covered with a resist such as PMMA. The electron beam is then scanned with the help of magnetic or electrical deflection systems and writes the desired pattern in the resist. The disadvantage of this method is the very slow writing process, which needs hours to write a high-resolution pattern. The second EBL method is the electron beam projection lithography, similar to the optical lithographical technique with the same name described above. In this case, the electron beam passing through the mask, which is a membrane with holes, focuses an image of the pattern on a resist with the help of an imaging system. The interested reader can consult (Tseng et al. 2003) for an excellent review on this subject.

The focused ion beam (FIB) lithography consists of scanning directly a substrate with a focused high-energy ion beam. This technique does not require masks or resists, and the point-by-point lithographical process relies on two different principles: (1) subtraction of surface atoms or (2) decomposition over the substrate of an organic vapor. In the first case, the desired pattern is imprinted directly on the substrate by scanned sputtering of atoms from the surface, while in the second case, the desired pattern is formed by the material deposited on the substrate. The FIB lithographical process can be monitored/imaged in real time by the ions and electrons emitted as a result of the interaction between the substrate and the

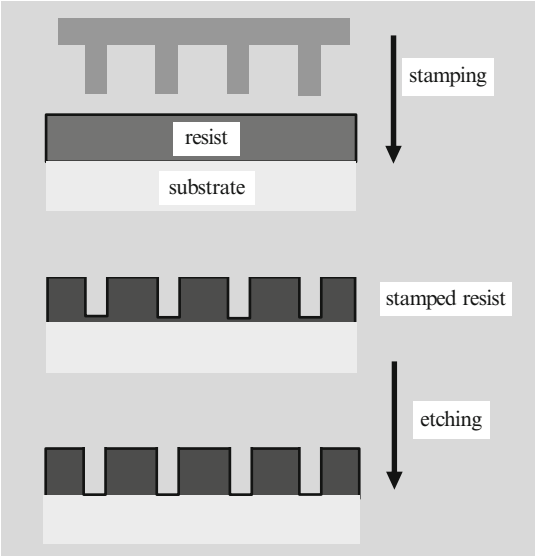


Fig. 1.18 The nanoimprint process

ion beam. The throughput, which represents the number of patterned features in one second, is quite low, of 10–10^2 features/s for the direct writing EB and FIB techniques, but increases dramatically, up to 10^{10} features/s, for the EB projection method.

The nanoimprint lithography is a high-resolution nanolithography, displaying high reproducibility in a short time. Its throughput is as large as 10^{12} features/s, while the resolution is as small as 10 nm. Typically, the nanoimprint technique utilizes a master or stamp fabricated from SiO_2 or Si, which contains the desired pattern fabricated via EB or FIB. This stamp is then used to imprint a resist with the desired pattern. The resist is generally a polymer, which can be UV-curable or thermoplastic (heated above its glass temperature), such as PMMA. After the polymer is stamped, the resist residues are removed by etching. The nanoimprint lithography, schematically displayed in Fig. 1.18, is excellently reviewed in Guo (2004).

The extreme top-down nanolithography technique is based on atomic beams, collimated and focused, similar to optical systems, by atomic lenses, atomic apertures, etc. In general, the atom beam interacts with laser beams and the resulted optical-atomic forces are utilized for (1) depositing atoms in order to built nanoscale structures or (2) patterning various shapes in a lithographic method (Meschede and Metcalf 2003). The atomic lithographic technique is based on an incoming atomic beam that is incident on a standing wave created by a laser and a mirror system, as shown in Fig. 1.19. The standing wave plays the role of a light mask for the incident atoms and is analogue to an array of lenses diffracting the atoms in the optical antinodes or nodes as a function of the sign of the difference between the laser frequency and the resonant frequency of atoms. In consequence, the desired pattern

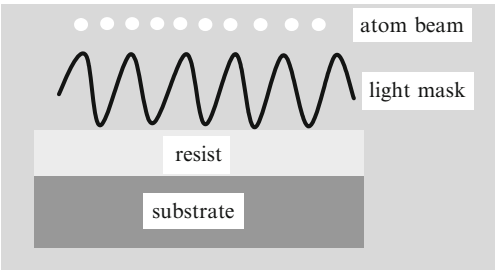


Fig. 1.19 The principle of atomic beam lithography

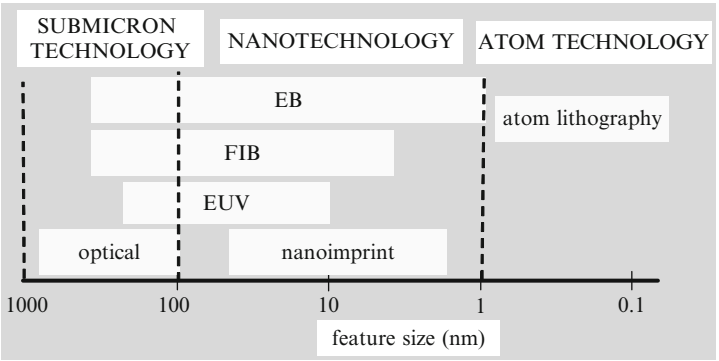


Fig. 1.20 The performances of the main lithographical techniques

is generated via the antinodes and nodes of the light beam. The incoming atoms pass through the mask and are deposited on a resist, which is subsequently developed.

The performances of various nanolithography techniques are illustrated in Fig. 1.20. There are several recent reviews dealing with nanolithography and its ultimate limits, which indicate also the strategies to be adopted when a certain device with certain features must be fabricated (Pease and Chou 2008; Saavedra et al. 2010).

The bottom-up approach to nanotechnologies is best represented by the self-assembly techniques. The self-assembly process stands for spontaneous organization of several components, such as molecules or nanoparticles, into a desired pattern. The organization process involves chemical, physical, or biochemical interactions, such as electrostatic and surface forces, or chemical hydrophilic and hydrophobic interactions. All these processes are extremely selective and have low defect rates, the resulted structures possessing a high quality (Huie 2003). Some basic self-assembly techniques are explained below.

The Langmuir–Blodgett (LB) technique is a well-known self-assembly method for thin film fabrication in which the pattern results by growing one monolayer at a time. A monolayer of a required material, originally adsorbed at a gas–liquid interface, is placed on the substrate on which the self-assembly structure will form. For instance, a monolayer of a fatty acid or, in general, a surfactant, is first spread

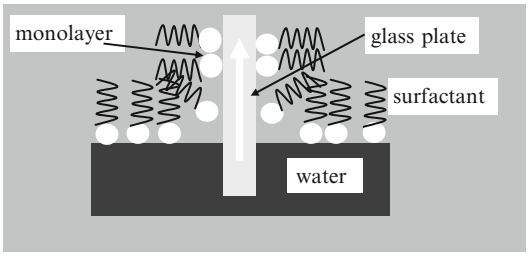


Fig. 1.21 The principle of the Langmuir–Blodgett technique

over a water surface. If a glass microscope slide acting as the self-assembly substrate is introduced in the water, a monolayer attaches to it when pulling the glass slide out of water, as shown in Fig. 1.21. The process can be repeated, and a monolayer is deposited each time the glass slide passes through the water surface. In practice, the LB technique requires specialized instruments such as a Langmuir trough, a dipping device for the lowering or raising substrate, and a movable barrier, which is controlled by a pressure sensor and maintains a constant surface pressure by sliding on the gas–liquid interface. The LB technique was used to fabricate polymeric films, 2D gold nanoparticle arrays, and semiconducting quantum dots. A good review about LB films is Rietman (2001).

Electrostatic self-assembly is a widespread technique based on the electrostatic interaction between nanoparticles or molecules. The origin of this self-assembly technique resides in the fabrication of multilayer films, in which each layer contains negatively and positively charged colloid particles, for example, Al and Si. The method is used in the fabrication of nanostructured films containing semiconductors, metals, polymers, magnetic materials, or organic molecules.

The thin films fabricated by electrostatic self-assembly are stable and uniform because of the strong ionic bonds between positive- and negative-charged particles, while the combination between polymers and the layers of electrically charged nanoparticles minimizes the eventual defects. In the electrostatic self-assembly process, a clean substrate is first immersed into a cationic solution, then rinsed, and dried, followed by another dip of the substrate already coated with cations into an anionic solution and subsequent rinsing and drying processes. This procedure can be repeated several times, the resulting nanostructured film growing with one layer/cycle (see Fig. 1.22). As an example, a heterostructure containing consecutive layers of gold and silver nanoparticles can be fabricated by first self-assembling positively charged Au nanoparticles on a negative glass substrate by dipping the substrate into 4-ATP (aminotiophenol)-capped Au solution having pH $=$ 4, then depositing a monolayer of negatively charged Ag nanoparticles by dipping the previous Au-coated glass into a 4-CTP (carboxythiophenol)-capped Ag solution having pH $=$ 8.5, and repeating the cycle as many times as necessary (Huie 2003). In a similar manner, it is possible to self-assemble a multilayer film consisting of conducting polymer chains and nanosized iron oxide, which combines the strength and flexibility of polymers with the magnetic properties of iron. In both cases, the

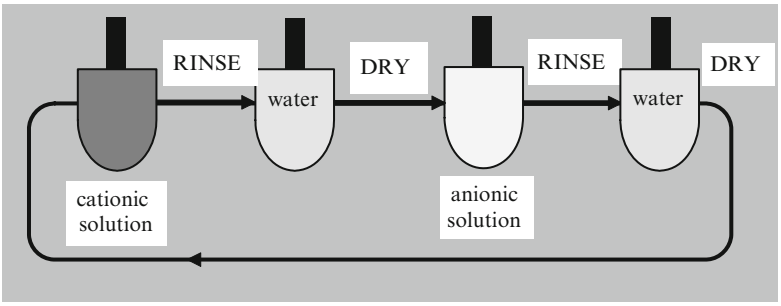


Fig. 1.22 The principle of the electrostatic self-assembly method

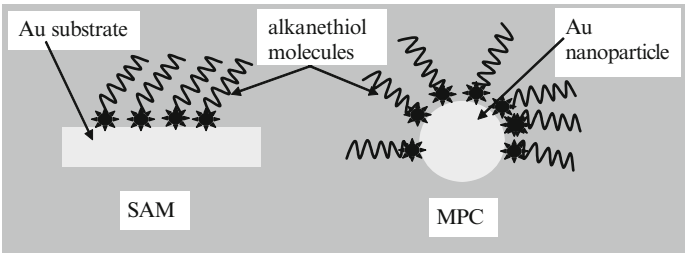


Fig. 1.23 Schematic representations of the SAM and MPC methods

properties of the final nanostructure are a mixture of the properties of the constituent layers, which is a typical characteristic for self-assembled nanostructures.

Chemical self-assembly methods are very frequently used to fabricate nanostructures, among them the SAM (self-assembled monolayers) and MPC (monolayer-protected clusters) techniques being prevalent. In the SAM method, monolayers form spontaneously by immersion of an appropriately chosen substrate into a solution, whereas MPC uses chemisorption to obtain nanoclusters with surfaces derivatized by ligand molecules, as displayed in Fig. 1.23.

Finally, proteins and DNA are the basic constituents of biomolecular self-assembly, which is used to fabricate (1) self-assembly of semiconducting nanoparticles, for example, ensembles of CdSe quantum dots; (2) functionalized metallic nanoparticles; or (3) biomolecular–metal complexes, for instance, DNA–Au complexes.

Many other self-assembly methods for nanostructures exist, the self-assembly strategy being one of the major techniques for fabricating new devices or materials (Shenhar et al. 2004). As an example, metal nanowires with lengths in the 20–50 – μm range and widths between 2 and 50 nm can be fabricated by self-assembly techniques in a defined 3D geometry, without lithography (Saif et al. 2003). These long and ultrathin metal nanowires are obtained by PECVD coating a Si substrate with SiO_2 incorporating OH impurities, which produce compressive stress in the Si/SiO_2 film. This film is subsequently annealed at 600°C for a long time to

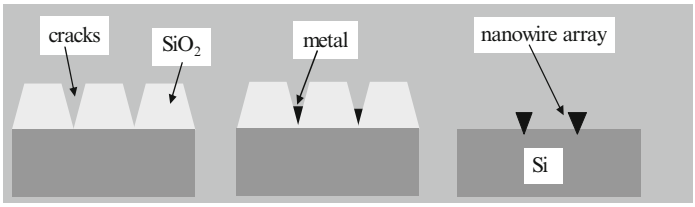


Fig. 1.24 Fabrication of a metal nanowire array by self-assembly

eliminate the OH impurities, process in which tensile stress develops in the film. As a result, the film cracks, the nanoscale cracks being able to extend up to the Si substrate. Finally, a metal is deposited in the cracks, and wet etching is used to remove the oxide. The geometry of the cracks depends on the type of the generated stress. For instance, a parallel array of cracks, and thus a parallel array of metal nanowires, as displayed in Fig. 1.24, is obtained due to an uniaxial stress.

1.2.3 Nanomaterials

The strength of bionanoelectronics resides in the fact that one or more dimensions of the nanoscale materials, termed as nanomaterials, are of the same size as bacteria, viruses, or biomolecules, such as DNA, from which the entire living matter is made (Gruner 2006). In what follows, we present briefly the main nanomaterials. The interaction of these nanomaterials with living matter is the main subject of the book.

Nanoparticles are frequently used in bionanoelectronics, especially for targeted drug delivery and sensing of biomolecules. Nanoparticles having few nm in diameter can be metallic, semiconducting, or even isolators, but their physical properties depend strongly on their size and differ dramatically from those of the bulk material from which they are made. A typical example is the transformation of the conduction band of bulk metals in discrete energy levels in a metallic nanoparticle, which resembles a quantum dot described in Sect. 1.1. The discrete nature of electronic states in a metallic nanosized particle is described by the average spacing between adjacent quantum levels, called Kubo gap:

$$\delta = 4E_F/3n, \tag{1.29}$$

where n is the number of electrons in the nanoparticle and E_F denotes the Fermi energy level. The nanoparticle behaves as a metal, i.e., the Kubo gap is not apparent, if $k_B T > \delta$. For instance, an Ag particle with a 3-nm diameter is metallic at ambient temperature because $k_B T \cong 25$ meV for $T = 300$ K, while $\delta = 10$ meV. Almost all metal nanoparticles (Au, Ag, Pd, Ni, Cu) are fabricated via evaporation of the bulk metal counterpart in vacuum.

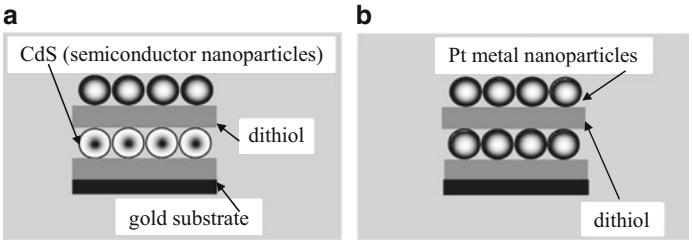


Fig. 1.25 Tridimensional assemblies of (**a**) metal nanoparticles–semiconductor nanoparticles and (**b**) metal nanoparticles–metal nanoparticles

Metal particles covered with organic molecules such as thiols are able to self-organize in 1D, 2D, and 3D arrays (Rao et al. 2000). In particular, 2D nanoparticle arrays are fabricated by mixing a hydrosol incorporating metal nanoparticles (Pd, Au) with a toluene solution of alkali thiol. Subsequently, the mixture of 3–4-nm metal nanoparticles and thiol is deposited on a surface and form a large and regular 2D array in which the nanoparticles are equally spaced in the x and y directions by about 5–6 nm (Rao et al. 2000). This 2D array is in fact an array of quantum dots in which the Coulomb blockade phenomenon manifests in electrical measurements.

Three-dimensional assemblies are fabricated via producing solutions of dithiol molecules and metal particles and then dipping a substrate into these solutions, followed by washing with toluene solution, and drying. Metal nanoparticle–semiconductor nanoparticle or metal nanoparticle–metal nanoparticle assemblies can be fabricated using this method (see Fig. 1.25a, b, respectively).

Nanoparticles are not only metallic but also semiconducting, such as silicon nanoparticles, elemental II–VI semiconductor quantum dots, and III–V nanoclusters. A review of fabrication of various types of nanoparticles is Adair et al. (1998). Nanoparticles play a central role in many areas of bionanoelectronics, especially in the area of controlled drug delivery and cancer therapies based on nanomaterials (Biswas et al. 2010). Also, core/shell nanoparticles are extremely used in biosensing, drug delivery, or cell labeling. The core/shell nanoparticles are of several types: (1) inorganic core/shell nanoparticles such as silver/silica or CdSe/CdTe nanoparticles, (2) organic/inorganic core/shell nanoparticles, in particular polymer/metal nanoparticles, such as CuS/PVA or TiO_2/celuloze, and (3) polymeric nanoparticles such as PMMA/PVC (Sounderya and Zhang 2008).

Also, many applications involve nanoparticles from group IV such as carbon nanoparticles. Because carbon is the key element of any known form of life, and in the human body carbon is the most encountered element after oxygen (Fan and Chu 2010), carbon nanoparticles such as diamond and graphite nanoparticles display important applications in the area of bioimaging and drug delivery nanosystems.

Similar to nanoparticles, nanowires are metallic, semiconducting, superconducting, or magnetic and are widespread used in many bionanoelectronic applications. The physical properties of nanowires, such as the quantized conductance behavior

showing a staircase shape, are not retrieved in the bulk materials from which they are fabricated.

A nanowire confines the carriers in two dimensions. The NW has a diameter of a few nm and lengths of 100–300 nm, but longer nanowires were also fabricated. Nanowires can be obtained through many methods, for example, bottom-up approaches (Bowler 2004). Another method to fabricate nanowires is based on Si(001), which is more reactive if not hydrogenated. Therefore, hydrogen can be used as a nanoscale mask, and other materials adsorbed on the nonhydrogenated regions can generate a nanowire. For example, (low-quality) Fe, Al, Co, Ga, or Ag nanowires can be obtained in the depassivated regions. There are several methods to fabricate self-assembled nanowires, such as methods based on specific chemical reactions and surface forces. For instance, rare-earth nanowires are fabricated by depositing the rare-earths on a Si(001) substrate and subsequent annealing. Similarly, Bi nanowires with a length of 200–600 nm and a width of 1.5 nm appear spontaneously when a Si(001) substrate covered with Bi is annealed around 600°C.

Template synthesis is, however, the most common method to fabricate nanowires. The template is, in fact, a prescribed pattern, for example, hexagonal, of nanosized pores or voids in a host material. Nanowires form when these pores are filled with the constituent material, termed source material. There are few reviews on nanowire growth such as Wang and Zhang (2008) or Bandaru and Pichanusakorn (2010).

Figure 1.26 illustrates the most common template for nanowire growth, which is anodic alumina. The alumina template results from anodizing Al thin films in specific acids. If an electrical current flows between the Al film and the cathode during anodization, an alumina film (membrane) displaying a regular hexagonal array of parallel and nearly identical cylindrical pores is obtained as the result of the etching process, as illustrated in Fig. 1.26. These nanopores have diameters within the 20–200-nm range, pitches of 50–400 nm, and a density that varies between 10^9 and 10^{11} cm^{-3} as a function of the etching conditions. The anodization process produces regular pore geometries. Nanochannels in glasses and etched polymers are also widely used to fabricate nanowires.

Precise templates can be created also by chemical etching of particle tracks resulted from ion bombardment. Nearly all types of nanowire can be fabricated using template techniques followed by a deposition method, and a subsequent extraction procedure of the source material.

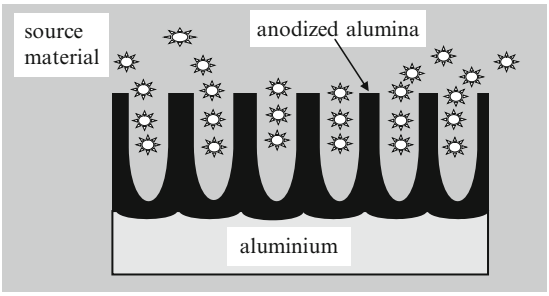


Fig. 1.26 Nanowire growth in an alumina template

Electrochemical deposition is frequently employed for growing not only thin films but also nanowires. The deposition procedure is restricted to a template consisting of nanopores, which is coated with a metallic thin film acting as a cathode on one side. Semiconductor nanowires of CdS, metallic nanowires of Co, Fe, Cu, Ag, Au, Pb, Ni, or superlattices A/B containing two constituents A and B, for instance, Co and Cu, were fabricated using the electrochemical deposition.

Extremely thin nanowires can be grown using MOCVD or CVD. In these methods, the precursor material of the nanowire is heated to generate vapors that fill the nanopores of the template, which is subsequently cooled to obtain the solid nanowires. Almost single-crystal nanowires are fabricated with the CVD method, polycrystalline nanowires being obtained otherwise. Examples of single-crystal nanowires grown with CVD techniques include GaAs, GaN, Bi, and InAs, with diameters smaller than 10 nm, as well as CNTs. A recent review of carbon growth techniques and applications in the area of electronics is Javey and Kong (2009).

The vapor–liquid–solid (VLS) growth method of nanowire relies on the fact that a liquid (L) droplet of a catalyst absorbs the vapors (V) of the source material. The nanowire forms as a result of the solidification (S) of the source material due to the liquid saturation and a subsequent nucleation process, which produces a preferential site for future deposition at the boundary of the liquid. Thus, other nucleation processes are avoided, and growth is allowed only in a single direction, as displayed in Fig. 1.27. The VLS method has been used to grow nanowires of Ge, Si, and ZnO. In situ electron microscopy can visualize and control the growth of Si and Ge NWs through the VLS method. Images and movies are produced in a transmission electron microscope, which is able to deposit catalysts and introduce CVD precursor gases in the sample under observation. In this way, the nucleation, the surface structure, and the growth kinetics are measured (Ross 2010).

Nanowires could be grown to form superlattices of the form XYXYXYXY… consisting of two types of nanowires, X and Y (Chik and Xu 2004). Figure 1.28 represents an InP/InAs superlattice (Björk et al. 2002), while Fig. 1.29 illustrates the energy band of a superlattice made of an alternation of intrinsic CNTs and

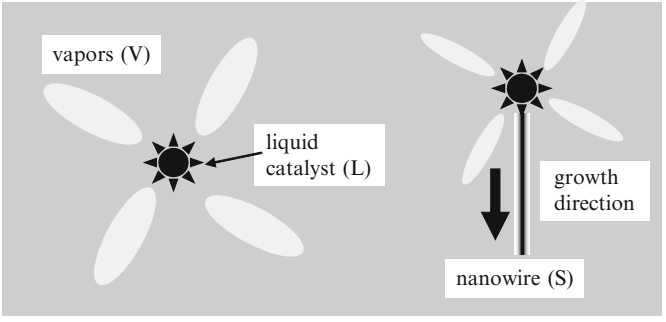


Fig. 1.27 The VLS method of growing nanowires

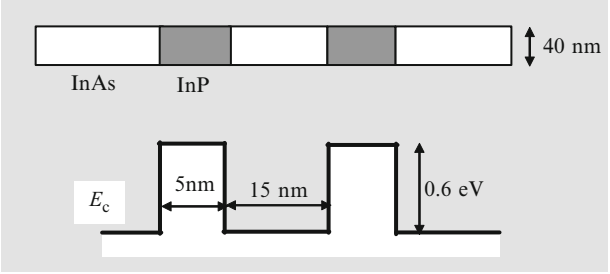


Fig. 1.28 Circular nanowire InP/InAs superlattice (*top*) and its conduction band diagram (*bottom*)

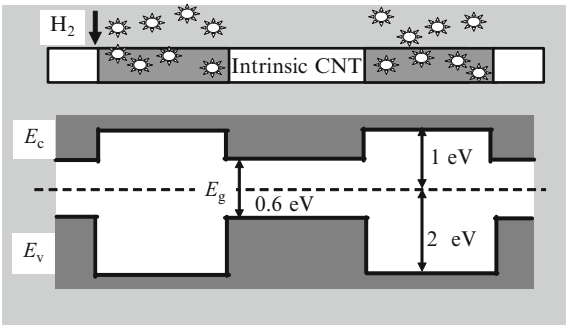


Fig. 1.29 Carbon nanotube superlattice (*top*) and its energy band diagram (*bottom*)

hydrogen functionalized semiconducting CNT; the energy bandgaps E_g of these material differ (Gülseren et al. 2003).

Both superlattices display a marked negative differential resistance and, as resonant tunneling diodes, can generate electromagnetic radiation with very high frequencies or can be employed as logic elements. The nanowire superlattice can be seen also as a series of three-dimensionally confined quantum dots, the transport through the structure taking place via coherent tunneling between adjacent quantum dots. If the width of the quantum barrier region increases, the tunneling process becomes sequential.

As in the case of core/shell nanoparticles, nanowire superlattices with radial periodicity can be fabricated by growing several nanowire shells with dissimilar properties. Typical examples are the coaxial Ge/Si and Si/Ge nanowires, obtained by initially growing one nanowire using the VLS method and then applying CVD deposition methods to grow the second nanowire over the first one (Lauhon et al. 2002).

Besides growing nanowires with prescribed properties, it is also important to align and position them using self-assembly techniques. Millions of CNTs, for example, can be aligned via a large-scale assembly method similar to biomolecular self-assembly processes (Rao et al. 2003). In this technique, millions of CNTs spread in solution are aligned on top of chemically functionalized patterns fabricated on a surface. The functionalization is achieved with two distinct regions, which are

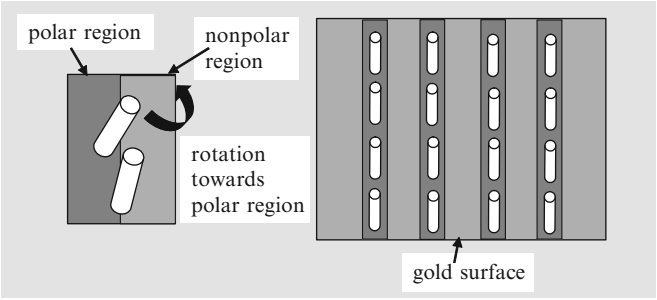


Fig. 1.30 The principle for aligning millions of carbon nanotubes by self-assembly

coated with nonpolar or polar groups. Because the CNTs in liquid suspension are attracted by the polar regions only, millions of CNTs align in less than 10 s, as illustrated in Fig. 1.30. The CNTs are rotated by the electrostatic attraction force in the direction of the polar region and are immobilized only in this region (Fig. 1.30).

Graphene is the most recent nanomaterial and has already attracted attention due to its unusual physics and potential applications, as demonstrated by awarding the Nobel Prize for physics in 2010 to A. Geim and K. Novoselov (Geim and Novoselov 2007) for groundbreaking research in this domain. Graphene is a monolayer sheet of graphite, with a thickness as small as 0.34 nm, consisting of carbon atoms in a sp^2 hybridization state, in which each atom is covalently bonded to three other atoms forming a honeycomb lattice. This lattice can be understood as consisting from two interpenetrating triangular sublattices. Graphene is at the origin of many carbon-based materials. For instance, graphite is formed by stacking millions of graphene layers, and single-walled carbon nanotube forms when graphene rolls up along a certain direction.

Graphene is a planar crystal and a natural 2D gas of charged particles. In its basic configuration, graphene is deposited on a 300-nm-thick SiO_2 layer grown on top of an n^+ silicon substrate. Only in this case it can be seen at an optical microscope. The wavelength at which graphene can be seen and its type, monolayer or bilayer graphene flakes, is determined by filtering a white light source and depends strongly on the SiO_2 thickness (Blake et al. 2007): for a 300-nm-thick SiO_2 layer graphene is optimally discriminate in green light, while blue light is most favorable for a SiO_2 layer with a thickness of 200 nm.

In the basic configuration mentioned above, containing a doped Si substrate, silicon is acting as gate which, upon applying a gate voltage V_g, controls the surface charge density n according to the expression (Novoselov et al. 2004)

$$n = \varepsilon_0 \varepsilon_d V_g / te. \tag{1.30}$$

In (1.30), ε_0 and ε_d are the dielectric permittivities of air and SiO_2, respectively, and t denotes the thickness of the SiO_2 layer. The carriers induced by the gate voltage exemplify the electrical doping effect, which is analogous to the chemical

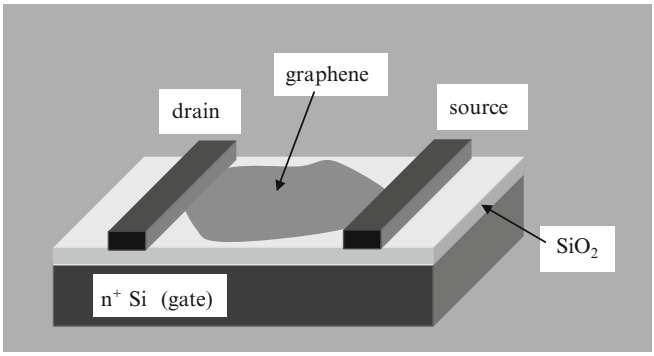


Fig. 1.31 Schematic representation of the graphene FET

Table 1.1 Characteristics of the main methods to produce graphene

Starting material	Brief description of method	Yield	Quality	Area
HOPG	Repetitive peeling	Low	Very high	Small
SiC	Reduction of Si at the surface of SiC at very high temperatures	Medium	Medium	Wafer size
GO	GO dispersion in hydrazine	High	Medium	Large
Gas mixture (CH_4 and H_2)	CVD	Very high	High	Very large

doping in semiconductor devices because both effects shift the Fermi energy level. In particular, the analogues of chemical p or n doping in graphene are obtained by applying negative or positive electrostatic gate voltages, respectively.

To implement specific devices, electrodes need to be patterned on the graphene sheet after its deposition on the Si/SiO_2 structure. An example of such a device is the graphene-based field-effect transistor (FET), displayed in Fig. 1.31.

The customary method to deposit graphene on Si/SiO_2 relies on mechanical exfoliation of highly ordered pyrolytic graphite (HOPG) using an adhesive tape, and a subsequent release of the graphene flake on Si/SiO_2 following tape removal. HOPG is a 3D structure consisting of vertically stacked graphene sheets. The fragments of HOPG obtained by mechanical exfoliation, including graphene, stick to the Si/SiO_2 surface due to van der Waals forces. This quite rudimentary method allows the production of graphene flakes with dimensions up to 1 mm (Geim 2009). More complicated techniques to obtain and handle graphene include the AFM microcleavage of HOPG pillars and subsequent deposition on Si/SiO_2 surfaces, epitaxial growth of graphene, and the growth of graphene by CVD techniques. All these graphene growth techniques were recently reviewed in Soldano et al. (2010). Table 1.1 presents the characteristics of the main methods to produce graphene.

Graphene has amazing physical properties. In particular, it has ambipolar transport characteristics, which can be controlled by the gate voltage in configurations similar to that displayed in Fig. 1.31. Moreover, the room-temperature mean-free

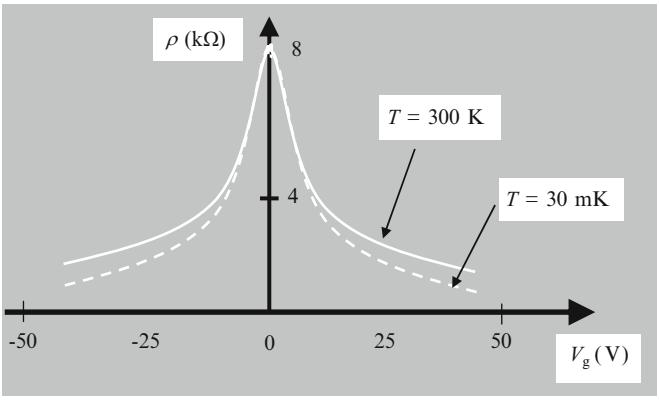


Fig. 1.32 Typical temperature and gate voltage dependence of graphene resistivity

path for ballistic transport is around 300 nm, and the carrier mobility takes a typical value of $15,000\ \mathrm{cm^2V^{-1}\ s^{-1}}$ at room temperature (Geim and Novoselov 2007) but can increase at $200,000\ \mathrm{cm^2V^{-1}\ s^{-1}}$ in suspended graphene sheets with a carrier concentration of $n = 2 \times 10^{11}\ \mathrm{cm^{-2}}$ (Bolotin et al. 2008). The graphene resistivity has only a slow variation with temperature, especially in samples with low mobility, but depends strongly on the gate voltage, as illustrated in Fig. 1.32. This behavior contrasts the strong resistivity dependence on temperature in other semiconductors or metals. The weak dependence of resistivity on temperature is determined by the dominant scattering mechanism, caused in low-mobility samples by static impurities and in high-density samples by residual electron-phonon scattering (Tan et al 2007).

In common semiconductors, the electron and hole transport are described by two independent Schrödinger equations, whereas in graphene, the electron and hole states are correlated through the so-called charge-conjugation symmetry (or chirality), and a Dirac-like equation for massless particles is needed to model their behavior. In particular, both electrons and holes in graphene are characterized by a linear dispersion relation, given by

$$E = \pm |\hbar \mathbf{k}|\, v_{\mathrm{F}}, \tag{1.31}$$

where E is the energy of charge carriers with wavevector $\boldsymbol{k} = \boldsymbol{i}\, k_x + \boldsymbol{j}\, k_y$, and v_{F} is the Fermi velocity. In Fig. 1.33, which illustrates this linear dispersion, the positive sign in (1.31) is assigned to electrons, while the negative sign corresponds to holes. The distinctive dispersion relation (1.31) shows that the valence and conduction bands touch each other in a point termed Dirac point, i.e., graphene is a gapless semiconductor.

A linear dispersion relation such as that in graphene is met only in photons propagating in vacuum, for which $E = \hbar\omega = hc/\lambda$, with c the speed of light. Note that the two situations correspond to completely different physical systems: graphene is a 2D gas of massless fermions, whereas photons are bosons.

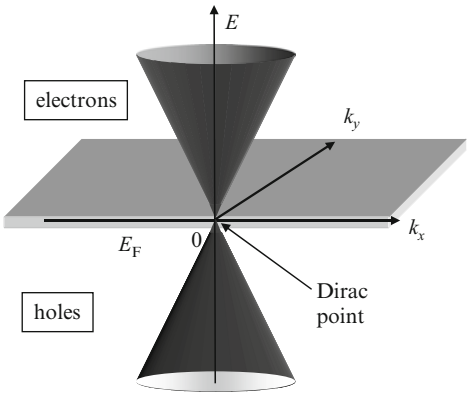


Fig. 1.33 The dispersion relation in graphene

In graphene, the linearity of the dispersion relation means that the effective mass of electrons and holes is zero, and the charge carriers propagate balistically with the velocity $v_F = 10^6$ m s$^{-1} \cong c/300$.

The gapless nature of graphene is in fact a disadvantage in many devices, so that different methods have been sought to open bandgaps in this material. A possibility is to confine the charge carriers by etching the graphene sheet contacted with electrodes. The resulting narrow strips, with typical length of 1–2 μm, are called graphene nanoribbons (GNR). The width W of such a strip controls the width of the energy gap, which opens at the Dirac point, according to the relation (Han et al 2007)

$$E_g = \alpha/(W - W^*), \tag{1.32}$$

where $\alpha = 0.2$ eV × nm and $W^* = 16$ nm are experimentally determined fitting parameters. Thus, band engineering becomes possible by tailoring the dimensions of the graphene nanoribbons. For example, the bandgap ranges from 100 to 3 meV if the width of the nanoribbon varies in the 20–90-nm interval.

The physics and applications of graphene are reviewed in many recent papers such as (Dragoman and Dragoman 2009b) and Wei and Liu (2010). Graphene has also important applications in the area of bionanoelectronics, biomolecule sensing, DNA sequencing, and drug delivery. More details are found throughout this book, especially in the chapter dedicated to sensing of biomolecules, such as DNA.

1.3 Conduction Properties of Biological Materials

The conduction of biological materials relevant for nanoscale electronics is related to the charged ions, which can be small molecules (≅ 0.2 nm), protein composites (≅ 10 nm), or giant polymers (DNA is often few centimeters long and contains millions of negative-charged groups in its backbone) (Waigh 2007). The charged ions are surrounded by water molecules. Examples of charged ions are the COO

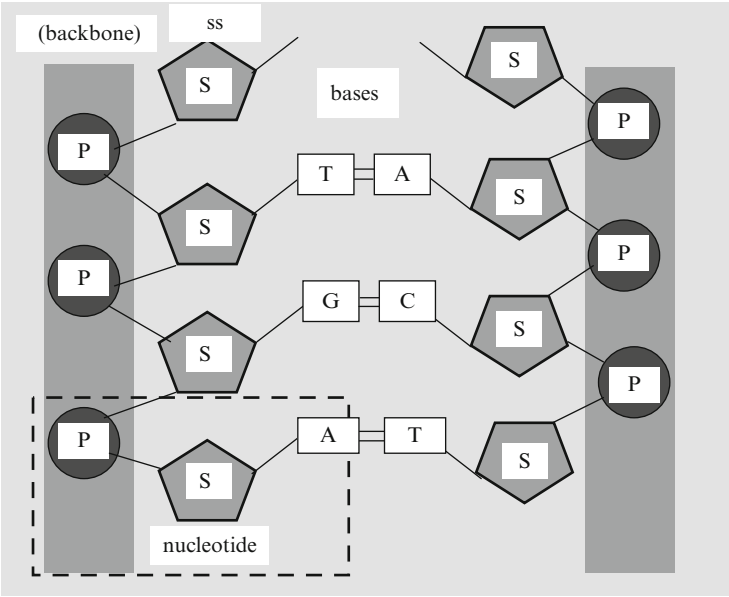


Fig. 1.34 The schematic DNA structure

carboxylic acid encountered in aspartic and glutamic acid in proteins and fatty acids, or the phosphate group in DNA. Their role is essential in the main functions of live: Ca^{2+} ions regulate the molecular motors in muscles, while a combination of Na^{+}, K^{+}, and Ca^{2+} ions control the propagation of electrical pulses in nerves. The role of charged ions in living organisms is excellently reviewed in Waigh (2007).

In what follows, we will focus on the conduction properties relevant for nanoscale electronics. DNA, the blueprint of life, is considered the most complex molecule, which not only contains the code of life but is considered as a good candidate for a series of devices associated with molecular electronics and thus is seen as a "material" due to the hybridization of its single strands, which is associated with self-assembly processes, and to its ability to be synthesized in an huge number of sequences (Kwon et al. 2009; Yang et al 2010). The single-stranded DNA (ssDNA) is formed by four bases and the phosphate backbone and is schematically represented in Fig. 1.34. The four bases of DNA are thymine (T), adenine (A), cytosine (C), and guanine (G). In the double-stranded DNA (dsDNA), the bases are bonded by hydrogen bonds in complementary manner: A–T and G–C. The dsDNA has a double-helix structure, with a width of 2 nm and a distance between stacked base pairs of about 3.4 Å. Billions of such base pairs in the folded dsDNA form the genome, which is found in any cell of eukaryote organisms.

On the other hand, in the last years, many nanoscale devices were tested for DNA sequencing, all based on the negative charging of the DNA backbone in solution. DNA is the best example that conduction of biological materials is quite

far from the conduction in crystals, which is quite well understood. The DNA is conducting due to the interbase hybridization of π_z orbitals, placed perpendicular to the planes of staked bases (Enders et al. 2004). However, DNA is not a crystal; it is not periodic. The potential barrier between two bases is very large, reaching 0.6 eV, which indicates an Anderson localization of electronic states of base pairs. In the double-helix DNA conduction experiments, the hydrophobic bases avoid contact with water, and the immediate environment consists of counterions formed by positive charges, which neutralize the negatively charged backbone. These considerations suggest that the conduction of biomolecules is strongly dependent on environment, in contrast to many electrical experiments at room temperature on inorganic nanodevices, where the environment plays no role. Moreover, at room temperature, the root-mean-square vibration amplitudes of DNA bases is ten times smaller than the distance between stacked base pairs, and one order of magnitude greater than in any crystal, DNA being nearly in a "melted" state, essential, however, for an easy replication or repair of its sequence. So, when dealing with biomolecules in general, and with DNA in particular, the precautions that must be taken in conduction measurements and the interpretation of results are far from the common experience with usual crystals in electronic laboratories.

The studies of DNA conductance have indicated a rich variety of DNA behavior, encompassing insulator, semiconductor, or conductor characteristics. In principle, ssDNA molecules are insulating at room temperature, but dsDNA in the form of short periodic structures such as poly(G)–poly(C), or bundles of λ-DNA, can conduct electricity due to hybridization.

As we have mentioned, the conduction in dsDNA is due to $\pi - \pi$ interactions between base pairs, which originate in the delocalized π bonding and π^* antibonding orbitals separated by an energy bandgap E_g of almost 4 eV, and which are produced by the p_z atomic orbitals perpendicular to the base plane. The DNA can be doped, as can common semiconductors, the doping involving chemical oxidation or reduction reactions. So a band engineering of DNA is possible, its transformation in a conducting wire occurring by further functionalization with metal particles, such as Ag.

In experiments, DNA acts as insulator, semiconductor, or conductor, depending on the different types of electrodes used to contact DNA; the various types of DNA strands, ssDNA or dsDNA; the length of the sample; and the specific base sequence. Many experiments are summarized in Table 1.1 of the key reference Enders et al. (2004). In Fig. 1.35, we have represented the work functions of DNA bases and metals, to get an insight into how DNA must be contacted to behave resistively or in a Schottky-like manner. In this figure, LUMO and HOMO stand for lowest unoccupied molecular orbital and highest occupied molecular orbital, respectively.

The conduction of DNA could be measured due to advancements in self-assembly and nanoscale lithographic technique, as well as in STM (scanning tunneling microscope), which is able to measure the tunneling current with high accuracy by approaching a sharp tip to the conductive sample (the distance between them is 1–2 nm). The STM will be extensively described in Chap. 3.

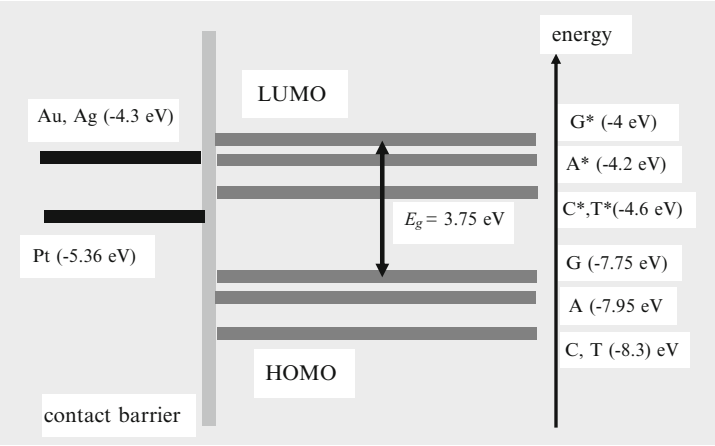


Fig. 1.35 Work functions of metals and DNA bases

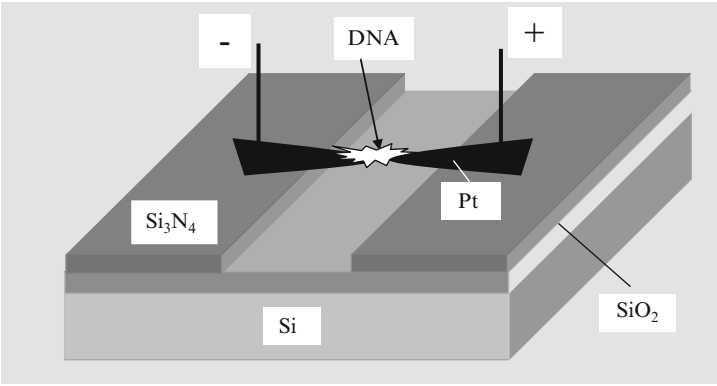


Fig. 1.36 DNA molecule trapped between gap electrodes for conduction measurements

The wide bandgap semiconducting behavior of dsDNA was demonstrated in Porath et al. (2000), in which it was considered that there is an overlap of π orbitals of neighboring base pairs in a 10.4-nm-long dsDNA structure, which covers the gap between two Pt electrodes and contains 30 base pairs of the same type (poly(G)–poly(C)). The DNA was trapped by electrostatic trapping when a bias was applied between the two electrodes (see Fig. 1.36).

The 1 μl of DNA from a dilute aqueous solution, which contained one DNA molecule per $(100\ \text{nm})^3$, was trapped between the electrodes. The gap between electrodes is 8 nm and was obtained by EBL. The $I - V$ characteristic of the device, similar to that of a Zenner diode, is displayed in Fig. 1.37. The authors in Porath et al. (2000) have attributed this behavior to the offset between the Fermi level of the nanogap electrodes and the molecular energy bands of DNA (see Fig. 1.35).

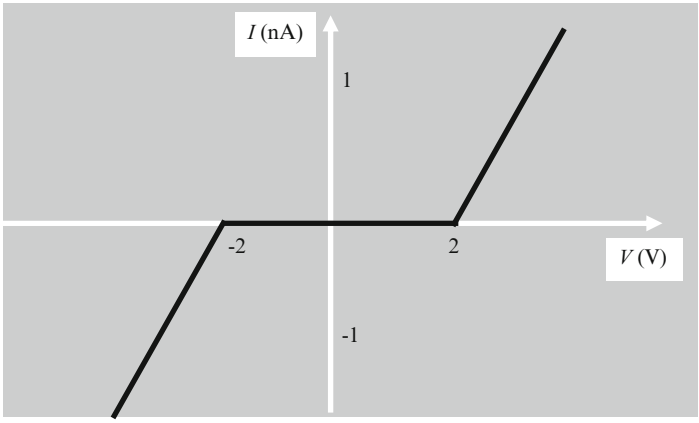


Fig. 1.37 The Zenner behavior of DNA between two electrodes

This is the typical semiconductor behavior of DNA, with a bandgap of about 1 eV at room temperature. It was shown that the overlapping of the π orbitals in the base-pair stack coupled to the backbone is enough to explain the opening of the bandgap in short DNA base sequences at room temperatures (Cuniberti et al. 2002). Using a Hamiltonian approach, it was evidenced that backbone coupling of overlapping π orbitals controls the energy gap opening in electron transmission along a DNA oligomer consisting of 30 G–C base pairs.

In this case, the transmission is written as

$$T = \frac{4\delta^2 \sin^2\theta}{[\sin(N+1)\theta - \delta^2 \sin(N-1)\theta]^2 + 4\delta^2 \sin^2 N\theta}, \tag{1.33}$$

where $\Phi = \cos\theta$ is the backbone coupling parameter and N is the number of resonances at the molecular orbitals, which are broadened by the spectral density parameter δ. The gap opening is easily understandable by the dispersion equation of the infinite G–C lattice: $\Phi(E) = -\cos q$, where q is the adimensional longitudinal momentum of the lattice. The lack of electronic states between LUMO and HOMO determines the bandgap opening.

The charge propagates along the π orbitals by nearest-neighbor hopping with probabilities t_a and/or is hybridized at the two edges of the structure with probabilities $t_{b\pm}$. Based on this conduction mechanism model, a model of charge propagation in C–G periodic base pairs forming a short DNA molecule is depicted in Fig. 1.38. The localized charges at the central and edge states are denoted by b and $c_\pm$, respectively, with an additional label indicating the specific base pair.

Using the above model and assuming that $t_{b+} = t_{b-} = t_b$, the gap in transmission, which corresponds to a gap in the current according to the Landauer formula, is given by

$$T_g = 2(t_a^2 + t_b^2)^{1/2} - 2t_a \tag{1.34}$$

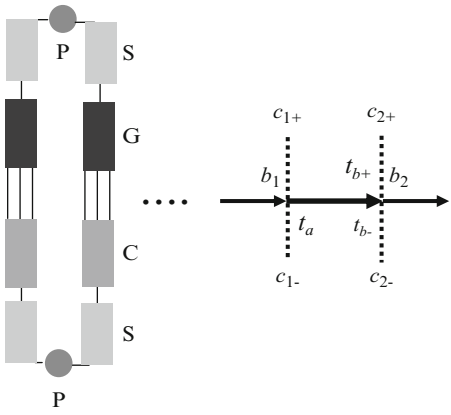


Fig. 1.38 The charge propagation in a GC periodic structure

and fits the $I - V$ curve at 18 K for $t_a = 0.37$ eV and $t_b = 0.74$ eV. Either the charge propagation or the coupling backbone model described above indicates that the bandgap is opened due to high reflections near zero energies, which correspond to a gap in electronic states between LUMO and HOMO.

DNA is not behaving only as a semiconductor. For example, DNA behaves as an electrical conductor with conductivity similar to that of conducting polymers, if several DNA molecules associate in ropes with lengths of 600 nm (Fink and Schönenberger 1999). In this case, the DNA is much longer than in the previous example where only 30 base pairs are used. The length and shape of DNA molecules and the number of bases dictate the conduction properties, as can be seen from the examples collected in Di Ventra and Zwolak (2004).

Thus, if DNA is conducting, there are three possible responsible mechanisms: thermal hopping, sequential tunneling from site to site, or coherent tunneling along the entire length of DNA. The tunneling rate, expressed as

$$R \cong \exp(-\gamma L), \tag{1.35}$$

is exponentially decreasing with the length of DNA, L. The γ parameter in (1.35) is 0.01 nm^{-1} for sequential tunneling and 0.1 nm^{-1} for coherent tunneling.

A step forward is to identify the conduction characteristics of the bases themselves. This was done by STM measurements (Xu et al. 2007). The electrical signals, and in particular the work functions of the four DNA bases, are different and, hence, unique, as resulted from STM measurements. A simple fitting formula for the $I - V$ signature of the four bases can be written as Dragoman and Dragoman (2009c)

$$I(V) = V/R + aV^2 \exp(-b/V), \tag{1.36}$$

where the first term in the rhs is due to a series resistance R and the last term originates from the Fowler–Nordheim behavior, which expresses the one-dimensional field emission from a triangular barrier of height ϕ in the presence of an applied electric field. Thus, the DNA bases are understood as potential barriers for

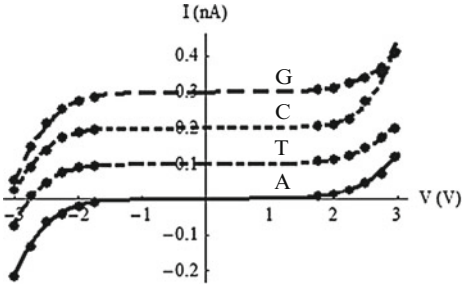


Fig. 1.39 The electric signature of the four bases of the DNA

field-emitted electrons, with heights ϕ. The work functions of the four DNA bases are then determined from

$$b = \frac{4L}{3}\frac{\sqrt{2m_0}}{e\hbar}\phi^{3/2}, \tag{1.37}$$

where m_0 is the free electron mass and L is the distance between the STM tip and the sample. The fitting $I - V$ curves are illustrated in Fig. 1.39 with solid line for A, dashed–dotted line for T, dotted line for C, and dashed line for G, the points on all the curves representing experimental data taken from Fig. 1.3a of (Xu et al. 2007). The characteristics for G, C, and T have been raised with 0.3, 0.2, and 0.1 nA, respectively, to render them easily readable. As can be seen from Fig. 1.39, the $I - V$ characteristics are not symmetric with respect to the origin. The forward polarization data were fitted with a series resistance value of 500 Ω for A, 1,000 Ω for T, 1,000 Ω for C, and 454 Ω for G, and the corresponding a parameters in (1.36) have values of 0.34, 0.27, 5.51, and 0.35 nS V^{-1}. On the contrary, the fitting series resistances for the backward polarization data are 555 Ω for A, 100 Ω for T, 500 Ω for C, and 666 Ω for G, the corresponding a parameters being 0.95, 1.35, 1.15, and 0.75 nS/V. Using (1.37), the work functions of the bases are found from the b parameters of the fit. These have, again, different values for positive and negative polarizations. For positive voltages, they are 9.7 V for A, 9.3 V for T, 16 V for C, and 10.1 V for G, while for negative voltages, the corresponding values are 11.2 V for A, 12.9 V for T, 12.2 V for C, and 10 V for G.

The different fitting parameters for the two polarizations are caused by specific base interactions with the substrate. The average work functions ϕ obtained from data at positive and negative polarizations are 1.74 eV for A, 1.81 eV for T, 2.12 eV for C, and 1.7 eV for G, for an estimated L value of 0.66 nm. It is interesting to note from Fig. 1.39 that all bases behave as semiconductors, displaying a Zenner diode-like behavior.

The conduction of DNA is an active research issue. For example, short dsDNA molecules formed from 13 base pairs of poly(dA)–poly(dT) were measured on a gold substrate, with gold nanoparticles with diameters of 10 nm as electrodes (Qian et al. 2009). The conduction type detected with the STM for a single dsDNA molecule was discovered to be semiconducting. Double-stranded DNA molecules were bonded with thiol to the gold film and are terminated with a single gold nanoparticle as electrode for STM tip measurements (see Fig. 1.40).

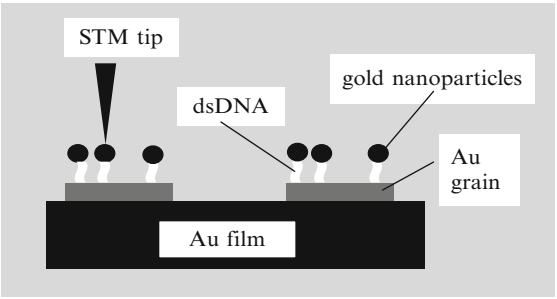


Fig. 1.40 The dsDNA terminated with nanoparticles as electrodes

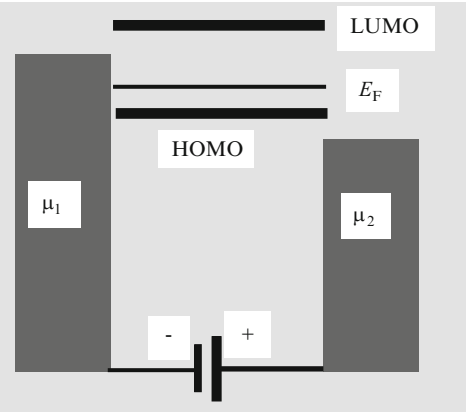


Fig. 1.41 The energy diagram of the single DNA molecule connected to a nanoparticle STM electrodes

The Landauer model is again used to explain the conduction mechanisms encountered in the above nanostructure. Considering that the dc voltage V is applied between the STM tip and the gold substrate, the electrochemical potential of the tip, μ_2, differs by the quantity eV from the electrochemical potential of the substrate μ_1: it is higher or smaller than μ_1 if a negative or a positive voltage V, respectively, is applied. The energy level diagram is presented in Fig. 1.41 when a positive voltage is applied on the tip. Thus, the current is given by

$$I = \frac{2e}{\hbar}\frac{\Gamma_1\Gamma_2}{\Gamma_1+\Gamma_2}[f(E,\mu_1) - f(E,\mu_2)], \tag{1.38}$$

where $\Gamma_{1,2}$ are the widths of the broadened molecular energy levels in the presence of metallic contacts, E is the energy level of the molecule, and $f(E,\mu_{1,2})$ are the Fermi functions of the contacts. Charging effects are expressed by a built-in electric potential $U_{\text{bi}} = U(N - 2f_0)$, where U is the potential per electron in the nonequilibrium state,

$$N = 2\frac{\Gamma_1 f(E,\mu_1) + \Gamma_2 f(E,\mu_2)}{\Gamma_1+\Gamma_2} \tag{1.39}$$

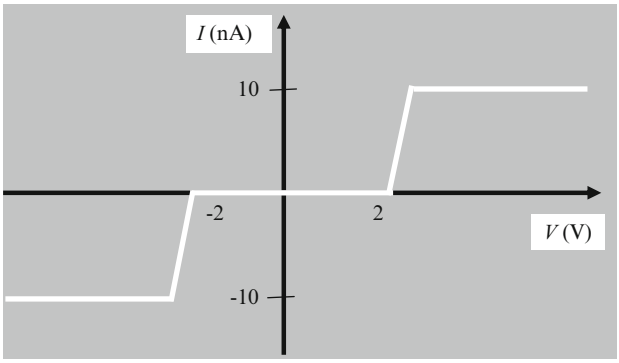


Fig. 1.42 The $I - V$ characteristic for a single DNA molecule having as contacts nanoparticles and a STM tip

is the number of electrons on the molecular level with energy E, and $f_0 = f(E_0, E_F)$ is the Fermi–Dirac distribution function at equilibrium, when the molecular level has energy E_0. In the presence of an applied potential, $E = E_0 + U_{bi}$.

Equations (1.38) and (1.39) were solved, the theoretical $I - V$ characteristics were found to be in good agreement with experimental data when $\Gamma_1 = 0.1\,\text{eV}$, $\Gamma_2 = 0.4\,\text{eV}$, $U = 0.1\,\text{V}$, and the LUMO and HOMO levels of the DNA are of about -7.99 and $-5.29\,\text{eV}$, respectively. The electrochemical potentials were written as

$$\mu_1 = E_F - \gamma eV \tag{1.40a}$$

$$\mu_2 = E_F + (1 - \gamma)eV, \tag{1.40b}$$

where $\gamma \cong 0.43$ describes the electrostatic potential distribution between the (DNA–nanoparticle)–substrate junction and the tip–(DNA–nanoparticle) junction. Under these conditions, a typical $I - V$ characteristic is displayed in Fig. 1.42.

Strong nonlinear current–voltage behavior was found in circular DNA bundle samples immersed in an aqueous solution, due to charge hopping through the DNA network (Romano et al. 2007). Inside DNA molecules, the tunneling mechanism describes the transport of carriers through the bases. The water solution containing the circular DNA bundle was deposited over the 1-mm gap between vacuum-evaporated gold electrodes deposited on a nylon substrate containing pores of 450 nm diameter (see Fig. 1.43). In a hopping model for charge transport, the charges hop from one localized site to the adjacent one due to the energy provided by the applied bias and thermal vibrations. The thermally activated current is in this case

$$I \cong \sinh(erV/2k_BT)\exp(-E_a/k_BT), \tag{1.41}$$

where E_a is the thermal activation energy and $r = a/d$, with a the hopping distance and d the gap between electrodes. By fitting the current–voltage data with the above formula, the experimental and theoretical curves are found to be similar for $E_a =$

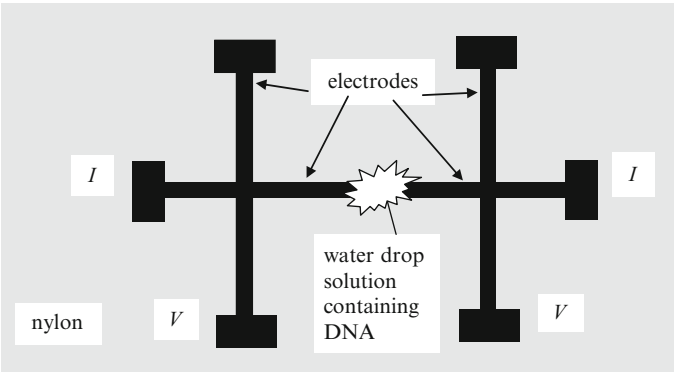


Fig. 1.43 The gap electrode configuration for circular DNA in water

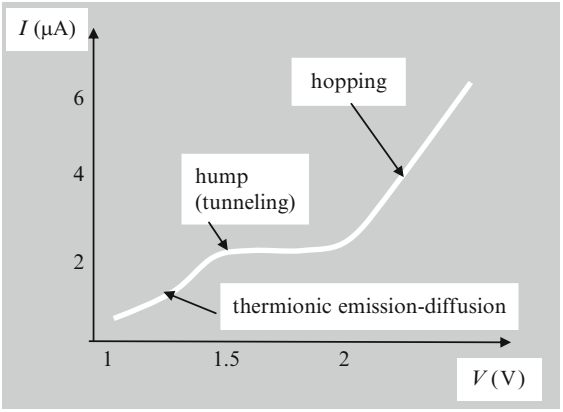


Fig. 1.44 Conduction mechanisms in circular DNA in water

300 meV. However, a hump located at $V_h = 1.5$ V in the curve, and present in nearly all samples, cannot be explained by the above equation (Fig. 1.44).

As pointed out above, the transport in DNA and in general in any organic molecule is the result of several different charging mechanisms. The hump is explained by the potential barrier that the charges must tunnel through inside the DNA molecules. For a rectangular barrier, the tunneling current is given by

$$I \cong \exp(2\varphi_T - eV)^{1/2}, \tag{1.42}$$

where ϕ_T is the barrier energy, which has values between 2.5 and 4.5 eV in different samples, consistent with the gap between HOMO and LUMO.

At low voltages, the effects of the contact between DNA and gold become visible in the $I - V$ characteristic, and the dominant transport mechanism is the thermionic emission–diffusion process expressed by the formula

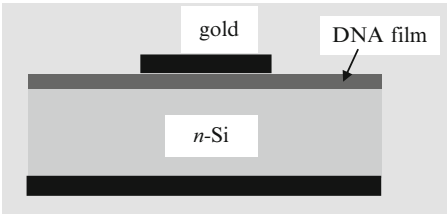


Fig. 1.45 The DNA thin film Schottky contact diode

$$I \cong \exp\left(-\frac{n\varphi_b - eV}{nk_BT}\right), \tag{1.43}$$

where ϕ_b is the potential barrier between electrodes and sample and $n \cong 18$ is a system-dependent constant.

So there are three conduction mechanisms which can be retrieved in a single experiment on circular DNA dispersed in water. In Fig. 1.44, we have displayed the regions of the current–voltage dependence in which each of these conduction mechanisms dominate. From fitting data analysis, the ϕ_b parameter is determined to be in the 400–500-meV range.

DNA thin films sandwiched between a *n*-doped Si substrate and a gold top contact display rectifying characteristics. The thickness of the thin film containing DNA extracted from wheat leaf tissues is about 19 nm, and the Au/DNA/*n*-Si structure behaves like a Schottky diode with an ideality factor of 1.22 and a barrier height of 0.78 eV (Sönmezolu et al. 2010). So, in the case of DNA thin films, the thermionic transport mechanism is dominant, and a Schottky-like diode with good performances can be obtained.

We can see from the above examples that there are two basic methods to connect biomolecules to metallic electrodes. One method is the top contact, as in the case of STM tip–(nanoparticle–DNA) electrodes in Fig. 1.40 or the DNA Schottky diode in Fig. 1.45. The other method consists of the nanogap electrodes, as in the case of the DNA molecule in Fig. 1.36 formed from 30 base pairs of poly(G)–poly(C), which displays a semiconducting behavior.

Nanogap electrodes are mostly used when one or few molecules must be connected to metallic contacts. Because the nanogap electrode has a planar configuration, this can be used for further integration with other electronic devices. However, it is quite difficult to fabricate gap widths of 1–8 nm, but many methods recently reviewed in Li et al. (2010) are presently able to produce a nanogap electrode or an array of nanogap electrodes. For example, FIB is a maskless method able to create nanogaps with dimensions between 10 and 20 nm in combination with EBL and optical lithography. The two electrodes of the nanogap are the drain and source of a transistor, and between them a CNT, a nanowire, or graphene can be deposited. For the subject treated in this book, nanogap electrodes are bridged by biological molecules. Conversely, nanogap electrodes can be fabricated over CNTs or nanowires, which are used as channels for ballistic field-effect transistors (FETs) where the two arms of the nanogap act as drain and source. In these cases, nanogap

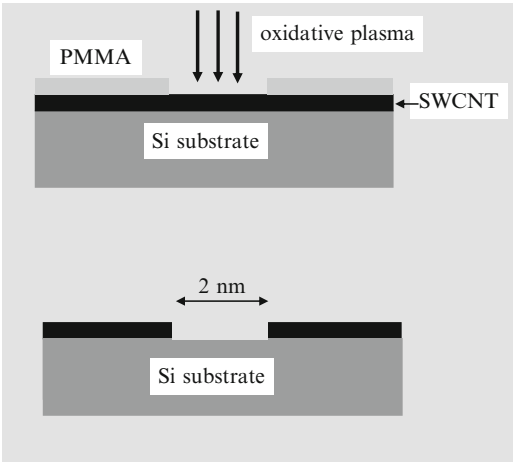


Fig. 1.46 Nanogap made in a single-walled CNTs

widths of 10–50 nm obtained by particle lithographic techniques are more than enough, but when we want to contact electrically only one molecule, the oxidative plasma is used to cut the electrodes in a precise manner, as in the example presented in Fig. 1.46. Nanogaps as small as 2 nm can be obtained in this way. Further, the CNT can be functionalized with the COOH group, and the device is ready to detect the hybridization of DNAs via conduction measurements (Guo et al. 2008).

Not only metallic contacts but also nanowires (Tian et al. 2010) and graphene (He et al. 2010) can be used as nanogap electrodes. The biosensing applications of electric nanogaps are reviewed by Chen et al. (2010).

1.4 Microfluidics and Nanofluidics

The last section of this chapter is addressed to micro- and nanofluidics. Micro- and nanofluidics is an area dedicated to the miniaturization of plumbing and the control of fluidic flows, which enable system integration of various biological sensors and the development of devices such as lab on a chip, which are analogous to the electrical integrated circuits. However, some microfluidics applications are encountered in the daily life: inkjet printing and liquid-crystals displays.

An interesting fact about the scale of electronic and fluidic processes is pointed out in Squires and Quake (2005). The success of integration of transistors on a chip reaching nowadays a density of 2–3 billions/chip is due to the fact that the downscaling of the geometrical dimensions of electronic devices does not affect their physical properties (up to gate lengths of 10–30 nm, where short channel effects strongly perturb the behavior of CMOS transistors). On the contrary, fluidic systems change strongly their properties depending on the length scales. Microfluidics deals with channels with dimensions of 100 μm, the corresponding volume of a cube with 100 μm sides being 1 nl. Some microchannel shapes are displayed in Fig. 1.47.

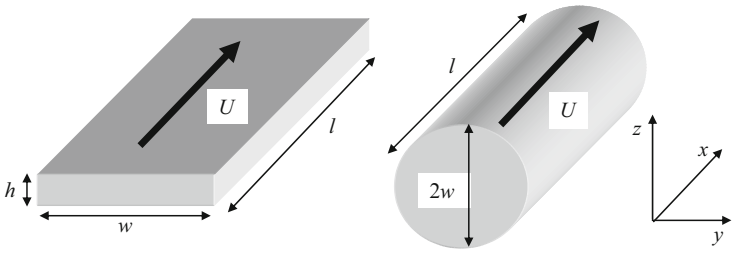


Fig. 1.47 Microchannels of various shapes

The fluid inside microchannels is considered a continuous media, and the velocity flow is described by the Navier–Stokes equation

$$\rho[(\partial \mathbf{v}/\partial t) + \mathbf{v}\nabla\mathbf{v}] = \nabla\hat{\boldsymbol{\sigma}} + \mathbf{f} = -\nabla p + \eta\Delta\mathbf{v} + \mathbf{f}, \tag{1.44}$$

where $\mathbf{v}$ is the fluid velocity, $\mathbf{f}$ denotes the external forces (gravitational, electric, or magnetic) per unit volume exerted on the fluid, ρ is the fluid density, $\hat{\boldsymbol{\sigma}}$ is the fluid stress tensor (viscosity contribution), η is the shear viscosity, and p is the pressure.

If the inertial force, which is nonlinear, is small (situation encountered in almost all microfluidic devices), the Navier–Stokes equation becomes the Stoke equation

$$\rho(\partial \mathbf{v}/\partial t) = \nabla\hat{\boldsymbol{\sigma}} + \mathbf{f} = -\nabla p + \eta\Delta\mathbf{v} + \mathbf{f}. \tag{1.45}$$

In both equations, the mass conservation requirement imposes that

$$\partial\rho/\partial t + \nabla(\rho\mathbf{v}) = 0 \tag{1.46}$$

holds. To decide if the inertial forces are important in the steady state condition

$$\partial\mathbf{v}/\partial t = 0, \tag{1.47}$$

we must calculate the Reynolds number, defined as the ratio between the inertial force and the viscous force

$$\mathrm{Re} = \left|\frac{\rho(v\nabla v)}{\eta\Delta v}\right| = \rho UL/\eta = UL/\nu, \tag{1.48}$$

where U is a characteristic flow velocity scale, L is a typical length scale, and $\nu = \eta/\rho$ is the viscous coefficient.

Relation (1.48) tells us that the Reynolds number decreases as the dimensions of the systems decrease. If the fluid velocity is in the range of cm s^{-1} and $L = 10$ μm, we obtain $\mathrm{Re} \cong 10^{-1}$, so that the convective (inertial) force is negligible compared to the viscous force. No convection means that there is no turbulence in microfluidic systems. In microfluidic devices, the flow is linear or laminar, is described by the

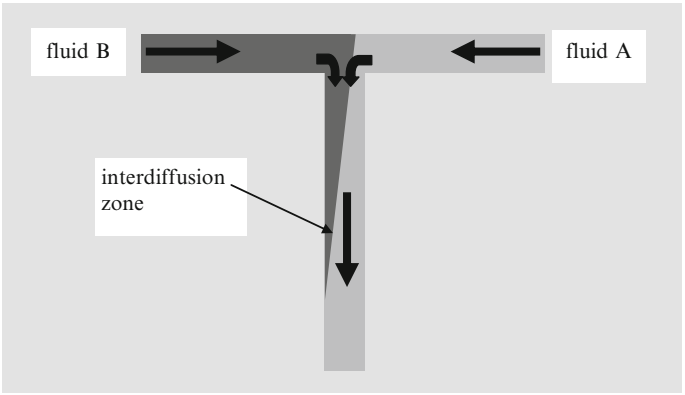


Fig. 1.48 The T-mixer

Stokes equation, and is deterministic and predictable. However, if the microchannels are curved, convection appears if the curvature is much greater than the channel radius, and the centrifugal and the convective (inertial) forces generate a secondary flow, the Dean flow. This effect is used to create microfluidic devices such as micromixers.

Another quantity of interest is the inertial time scale

$$\tau_{\mathrm{i}} = \rho L^2/\eta, \tag{1.49}$$

which indicates the duration after which the flow reaches the steady state. This time is small, of about 10 ms in a $100-\mu$m channel.

Although almost all microfluidic devices display regular, deterministic flows, there are effects, such as capillarity or electrokinetic effects, where the nonlinearity could increase if the dimensions are shrinking. However, in the real world, the fluids are described by high Re numbers.

The ratio between the regimes of convection and diffusion is named Péclet number

$$\mathrm{Pe} = Uw/D, \tag{1.50}$$

where D is the diffusion coefficient and w is the channel width. In the case of a biomolecule flow in a 100 μm microchannel with velocity of 100 $\mu\mathrm{m\ s}^{-1}$, we get $\mathrm{Pe} = 250$ if $D = 40\ \mu\mathrm{m}^2\ \mathrm{s}^{-1}$. The Pe parameter is crucial for the mixing of liquids.

An example of a microfluidic device, which mixes liquids, is the T-shape mixer presented in Fig. 1.48. The two fluids are flowing alongside each other in the channel, and an interdiffusion zone is produced since the molecules of the two fluids diffuse into each other. This simple system was used to measure the analyte concentrations, or antigen–antibody binding via the injection of antibody solution along a labeled solution of antigen.

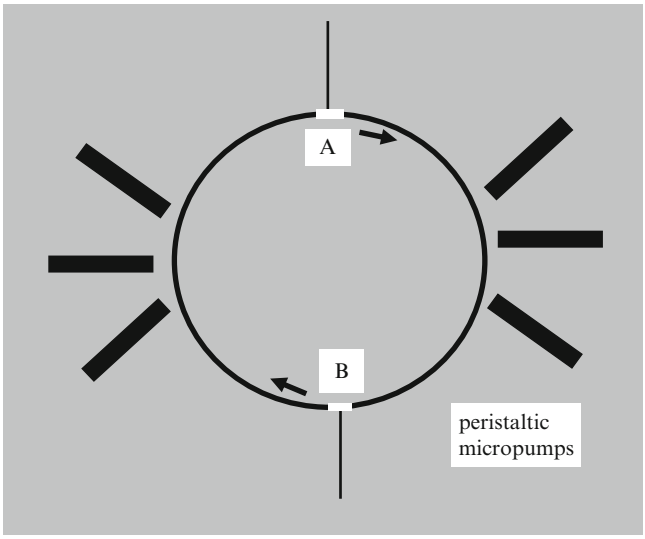


Fig. 1.49 The schematic view of a rotary mixer

Mixing of fluids is of paramount importance in micro- and nanofluidics. A basic physical mechanism for mixing is the Taylor dispersion, which is based on the dispersion of a tracer along the stationary flow in a small-dimension microchannel or a capillary. If we consider that the concentration C is homogenous along the transverse direction, the equation for the tracer spot spreading is

$$DC/Dt = \partial C/\partial t + U\,\partial C/\partial x = D_{\text{eff}}\Delta C, \tag{1.51}$$

where the flow direction is denoted by x, and the effective diffusion coefficient is given by

$$D_{\text{eff}} = D[1 + \beta(Uh/D)^2], \tag{1.52}$$

where h is the channel height and β is a parameter dependent on the cross-sectional shape of the microchannel. Expression (1.52) tells us that when Péclet is high, the diffusion coefficient is increasing with the square of the velocity of the fluid. This effect is called Taylor diffusion and, based on it, a rotary mixer device as that in Fig. 1.49 can be designed.

The peristaltic pumps in Fig. 1.49 are driving the flow around the ring. The two fluids A and B are mixed after few rotations and, due to the Taylor diffusion, this occurs in tens of milliseconds. In this way, such a rotary mixer can extract DNA from a single cell and amplify it. Another way of rapid mixing is flow focusing, where a central fluid stream is narrowed using a constriction. The mixing in this case is performed in tens of μs. There are various ways of driving or manipulating fluids via electrokinetics effects, such as electroosmosis and electrophoresis. We will not detail further the theory of microfluidics since this is beyond the scope

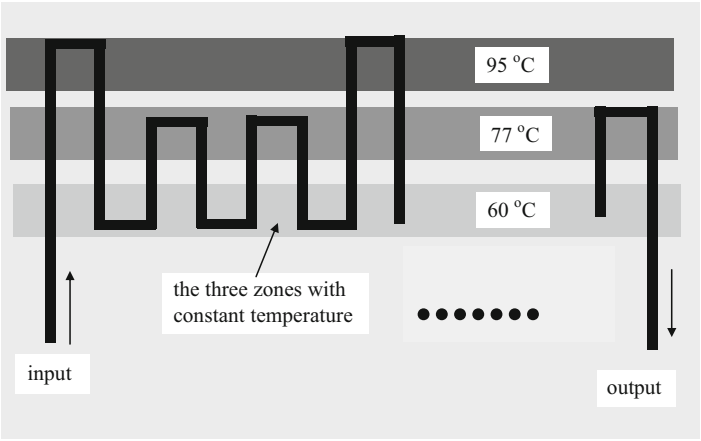


Fig. 1.50 PCR lab on chip

of our book. However, as we will see in the next chapter, micro- or nanofluidics devices are integrated with nanoelectronic devices for biomolecule sensing. For the interested readers, there are a lot of reviews and books about microfluidics. We mention here just a few of them. An early review on microfluidic integrated devices is found in Erickson and Li (2004), while a brief review on droplets processing is found in Tabeling (2009). Droplets microfluidics is reviewed by Teh et al. (2008), while the biological applications of microfluidics is found in the book of Tian and Finehout (2008). The ultimate device of microfluidic is a lab-on-chip, which is a device enabling the miniaturization, integration, and automation of biochemical assays. There are many types of lab-on-chips, such as polymerase chain reaction (PCR), which is used to replicate DNA and amplify it in a large amount.

This lab-on-chip, depicted in Fig. 1.50, acts like a chemical amplifier and so is the analogue of an electronic amplifier (Kopp et al. 1998). PCR is based on three steps: (1) heating the DNA solution at 95°C to separate dsDNA in single strands, (2) heating at 50–65°C for primer binding to targeted sites (annealing), and (3) heating at 72–77°C for primer extension with DNA polymerase, such as Taq. PCR is able to produce 10^3–10^6 copies of a DNA strand.

The fluid containing DNA is pumped into various temperature regions to accomplish the PCR steps indicated in different gray colors in Fig. 1.50 (Kopp et al. 1998). The three temperature zones are preserved at constant values with the help of thermostated copper blocks. Each cycle of melting doubles the number of DNA molecules by changing the temperatures as indicated above.

The PCR lab-on-chip is formed by a single microchannel having a meandered shape, which is hydrostatically pumped. The microchannel is etched in glass and has a depth of 40 μm and a width of 90 μm. The output of the PCR is collected by a capillary outlet.

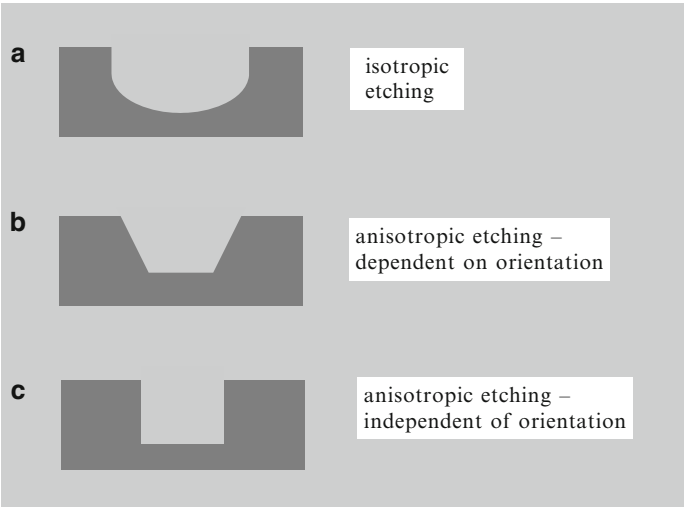


Fig. 1.51 Si etching techniques

There are many other types of lab-on-chips, all performing the following tasks:

- Fluid transport
- Fluid control (valves)
- Fluid mixing
- Fluid separation
- Detection
- Readout

The development of microfluidic technology is related to the development of microelectromechanical systems (MEMS). Although there are other materials in which microfluidic devices are made, in what follows, we will briefly describe some Si MEMS techniques to fabricate them because in this way, the microfluidic devices can be integrated with electronic devices, which are able to process the information. The microfluidic devices are fabricated by etching silicon.

Various etching shapes result from different techniques, as shown in Fig. 1.51. There are two main categories of etching techniques: wet or chemical etching, illustrated in Fig. 1.51a, b, and dry or plasma etching, shown in Fig. 1.51c. The final goal is to fabricate high-precision and low-roughness microchannels with controllable shapes or other microfluidic devices using the etching process. However, ideal vertical walls can be realized only with plasma-etching techniques.

There are two basic types of micromachining techniques: bulk micromachining and surface micromachining. In bulk micromachining, a specific mechanical structure is fabricated from a wafer by etching the back side of it. A typical example, of producing a micromechanical membrane, is illustrated in Fig. 1.52. Such a membrane, with a thickness of 1 μm and a surface as large as 2×2 cm^2, can be produced using chemical or plasma-etching methods. In the case of bulk micromachining,

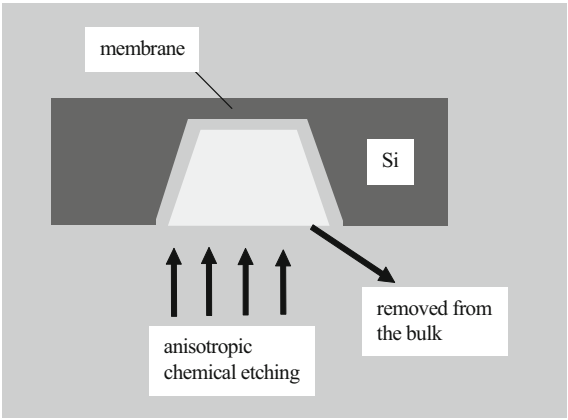


Fig. 1.52 The bulk Si micromachining technique

the desired shape and the precision of micro- or nanosized mechanical structures is controlled by etching stop techniques. On the contrary, in surface micromachining, MEMS or NEMS are fabricated by consecutive depositions of thin film layers on a common substrate, and a subsequent selective etching of one of these layers, referred to as the sacrificial layer.

Nanofluidics deals with fluid motion in narrow channels, with dimensions smaller than 100 nm. There are some important differences in the fluidic motion at such small dimensions in comparison to microfluidics. In principle, due to the very high aspect ratio between the surface and the volume of the devices, the fluid interaction with the walls becomes significant. It is important to note that according to Sparreboom et al (2010), the Stokes equation is still valid for channels with dimensions of at least 10 nm. If the dimensions of the channels are less than 10 nm, as in the case of single-walled CNTs, the discrete nature of molecules participating at the flow must appear explicitly in the flow equations. It is extremely interesting to note that liquids preserve their continuous character up to very small dimensions, while in a transistor channel with widths of 10 nm, the transport of electrons is fully quantized.

In nanofluidics, the motion of fluids is no longer induced by the pressure difference as in the case of microfluidic channels. The typical fluid motion in nanofluidics channels is an electroosmotic flow. The electroosmotic flow, i.e., the liquid transport driven by an axial applied electric field, is encountered also in microfluidic systems, but in nanofluidics, it becomes a basic tool for fluid transport.

Electrical charges are transported by a solid surface in contact with an aqueous solution, which is conductive due to the ionic groups at the surface, dissolved salts, etc. This excess of surface charge attracts the counterions and repeals similarly charged co-ions. The thickness of this ionic double layer, which depends on the ionic concentration, is named Debye length. Beyond the Debye length, the fluid has no charge; it is neutral. The electroosmotic flow is the result of an applied electric field along the fluid. The velocity profiles for the flow due to pressure difference (microfluidics) and electroosmotic flow (nanofluidics) are depicted in Fig. 1.53.

From Fig. 1.53, it can be seen that in nanofluidics channels, the velocity of the fluid is constant (outside the Debye length), in contrast with the microfluidic

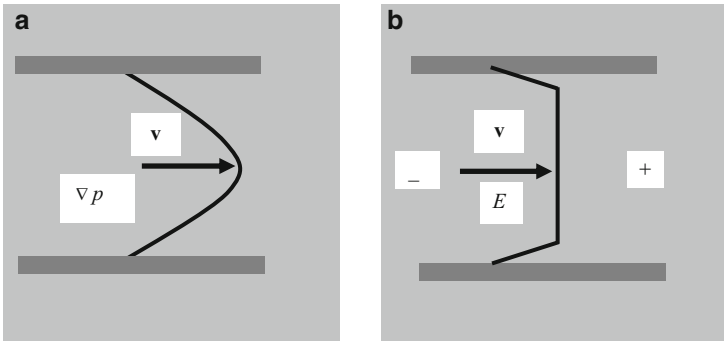


Fig. 1.53 Typical (**a**) microfluidics and (**b**) nanofluidics flow mechanisms

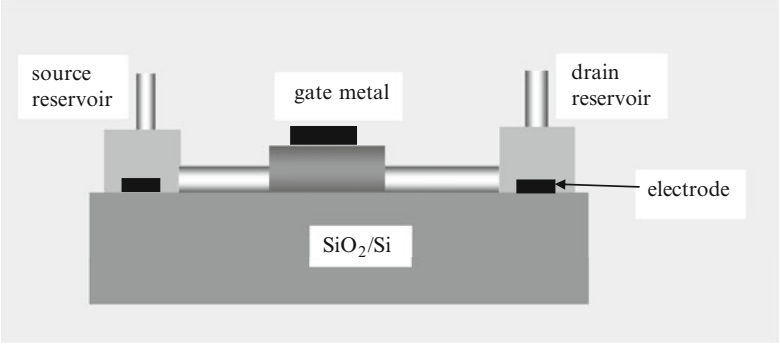


Fig. 1.54 The nanofluidics transistor

channels where, due to pressure difference, the velocity distribution has a parabolic shape (Poiseuille flow).

The velocity of the electroosmosis flow profile beyond the Debye length is given by

$$v = -\varepsilon_0 \varepsilon_r E \phi_s / \eta, \tag{1.53}$$

where ϕ_s is the potential at the surface and E is the applied electric field in the flow direction. Formula (1.53) reveals the constant velocity profile for electroosmotic flow.

The Debye length l_D decreases when the ion concentration n_i increases, $l_D \propto n_i^{-1/2}$, and ranges from 1 to 100 nm for aqueous solutions. In microchannels, the Debye length is much smaller than the channel dimensions, so that the solution is isolated from the surface charge and is neutral. On the contrary, in nanochannels with a dimension smaller than the Debye length, the electrostatic field penetrates through the channel and thus controls the flow via the field effect. This effect is analogue to the effect of the gate on a FET transistor, and the devices with 1D nanochannels analogous to FET channels and working on the above principle are termed nanofluidic transistors (Karnik et al. 2005) (see Fig. 1.54). In the case of nanofluidics channels, the ion concentration is modulated by an electrostatic gate.

More advanced nanofluidics devices based on CNTs are used in various applications such as "water wires," membranes, gas transport, filtration (Noy et al. 2007), or as biosensors (Lee et al. 2009). The biosensors, which are of paramount importance for bionanoelectronics, will be the subject of the next chapter. The technologies for the fabrication of nanochannels and nanofluidic devices are the top-down and bottom-up approaches described in this chapter (Mijatovic et al. 2005).

References

Adair JH, Li T, Havey KK, Moon J, Mecholsky J, Morrone A, Talham DR, Ludwig MH, Wang L (1998) Recent developments in the preparation and properties of nanometer-size spherical and platelet-shaped particles and composite particles. Mater Sci Eng R 23:139–242

Appenzeller J, Radosavijevic M, Knoch J, Avouris Ph (2004) Tunneling versus thermionic emission in one-dimensional semiconductors. Phys Rev Lett 92:048301

Bandaru PR, Pichanusakorn P (2010) An outline of synthesis and properties of silicon nanowires. Semicond Sci Technol 25:024003

Biswas A, Wang T, Biris AS (2010) Single metal nanoparticles spectroscopy: optical characterization of individual nanosystems for biomedical applications. Nanoscale 2:1560–1572

Björk MT, Ohlsson BJ, Thelander C, Persson AI, Deppert K, Wallenberg LR, Samuelson L (2002) Nanowire resonant tunneling diode. Appl Phys Lett 81:4458–4460

Blake P, Noviselov KS, Castro Neto AH, Jiang D, Yang R, Booth TJ, Geim AK, Hill EW (2007) Making graphene visible. Appl Phys Lett 91:063124

Bowler DR (2004) Atomic-scale nanowires: physical and electronic structure. J Phys: Condens Mater 16:R721–R754

Bolotin KI, Sikes KJ, Jiang Z, Klima M, Fudenberg G, Hone J, Kim P, Stormer HL (2008) Ultrahigh electron mobility in suspended graphene. Solid State Commun 146:351–355

Chen X, Guo Z, Yang G-M, Li J, Li M-Q, Liu J-H, Huang X-J (2010) Electrical nanogap devices for biosensing. Mater Today 13:28–41

Chik H, Xu JM (2004) Nanometric superlattices: non-lithographic fabrication, materials, and properties. Mater Sci Eng R43:103–138

Cuniberti G, Craco L, Porath D, Dekker C (2002) Backbone-induced semiconducting behaviour in short DNA wires. Phys Rev B65:241314

Datta S (1997) Electronic transport in mesoscopic systems. Cambridge University Press, Cambridge

del Alamo JA, Eugster CC, Hu Q, Melloch MR, Rooks MJ (1998) Electron waveguide devices. Superlattice Microstruct 23:121–137

Di Ventra M, Zwolak M (2004) DNA electronics. In: Encyclopedia of nanoscience and nanotechnology. Nalwa HS (ed), American Scientific Publishers, California 1–19

Dragoman D, Dragoman M (1999) Advanced optoelectronic devices. Springer, Berlin

Dragoman D, Dragoman M (2001) Micro/nano-optoeletromechanical systems. Prog Quantum Electron 25:229–290

Dragoman D, Dragoman M (2004) Quantum-classical analogies. Springer, Berlin

Dragoman M, Dragoman D (2009a) Nanoelectronics. Principles and devices. Artech House, London

Dragoman M, Dragoman D (2009b) Graphene-based quantum electronics. Progr Quantum Electronics 33:165–214

Dragoman D, Dragoman M (2009c) The real-time detection of deoxyribonucleic acid bases via their negative differential conductance signature. Phys. Rev E 80:022901

Enders RG, Cox DL, Singh RRP (2004) The quest for high conductance DNA. Rev Mod Phys 76:195–214

Erickson D, Li D (2004) Integrated microfluidic devices. Anal Chim Acta 507:11–26
Fan J, Chu PK (2010) Group IV nanoparticles: synthesis, properties and biological applications. Small 6:2080–2098
Ferry DK, Goodnick SM (2009) Transport in nanostructures. Cambridge University Press, Cambridge
Fink H-W, Schönenberger C (1999) Electrical conduction through DNA molecules. Nature 398:407–410
Geim AK, Novoselov KS (2007) The rise of graphene. Nat Mater 6:183–191
Geim AK (2009) Graphene: Status and prospects. Science 324:1530–1534
Gruner G (2006) Carbon nanotube transistors for biosensing applications. Anal Bioanal Chem 384:322–335
Gülseren O, Yildirim T, Ciraci S (2003) Formation of quantum structures on a single nanotube by modulating hydrogen adsorption. Phys Rev B 68:115419
Guo LJ (2004) Recent progress in nanoimprint technologies and its applications. J Phys D 37:R123–R141
Guo X, Gorodetsky AA, Hone J, Barton JK, Nuckolls C (2008) Conductivity of a single DNA duplex bridging a carbon nanotube gap. Nat Nanotechnol 3:163–167
Han MY et al (2007) Energy band-gap engineering of graphene nanoribbons. Appl Phys Lett 98:206805
Harriott LR, Hull R (2004) Nanolithography. In: Di Ventra M, Evoy S, Heflin JR Jr (eds) Introduction to nanoscale science and technology. Kluwer Academic Publishers, Dordrecht 7–40
He Y, Dong H, Li T, Wang C, Shao W, Zhang Y, Jiang L, Hu W (2010) Graphene and graphene oxide nanogap electrodes fabricated by atomic force microscope nanolithography. Appl Phys Lett 97:133301
Huie JC (2003) Guided molecular self-assembly: a review of recent efforts. Smart Mater Struct 12:264–271
Javey A, Kong J (eds) (2009) Carbon nanotube electronics. Springer, Heidelberg
Karnik R, Fan R, Yue M, Li D, Yang P, Majumdar A (2005) Electrostatic control of ions and molecules in nanofluidic transistors Nano Lett 5:943–948
Kopp MU, de Mello AJ, Manz A (1998) Chemical amplification: continuous-flow PCR on a chip. Science 280:1046–1047
Kwon Y-W et al (2009) Material science of DNA. J Mater Chem 19:1353–1380
Lauhon LJ et al (2002) Epitaxial core-shell and core-multishell nanowire heterostructures. Nature 420:57–61
Lee M, Baik KY, Noah M, Kwon Y-K, Lee J-O, Hong S (2009) Nanowire and nanotube transistors for lab-on-a-chip applications. Lab on a Chip **9,** 2267–2280
Li T, Hu W, Zhu D (2010) Nanogap electrodes. Adv Mater 22:286–300
Meschede D, Metcalf H (2003) Atomic nanofabrication: atomic deposition and lithography by laser and magnetic forces. J Phys D 36:R17–R38
Mijatovic D, Eijkel JCT, van den Berg A (2005) Technologies for nanofluidic systems: top-down vs. bottom-up—a review. Lab on a Chip 5:492–500
Novoselov KS, et al (2004) Electric field effect in atomically thin carbon films. Science 306:666–669
Noy A, Park HG, Farnasiero F, Holt JK, Grigoropoulos CP, Bakajin O (2007) Nanofluidics in carbon nanotubes. Nanotoday 2:22–29
Pease RF, Chou SY (2008) Lithography and other patterning techniques for future electronics. Proc IEEE 96:248–270
Porath D, Bezryadin A, de Vries S, Dekker C (2000) Direct measurement of electrical transport through DNA molecules. Nature 403:635–637
Qian J, Liao S, Xu S, Stroscio MA, Dutta M (2009) Direct measurement of electrical transport through single DNA molecules. J Appl Phys 106:033702
Rao CNR, et al (2000) Metal nanoparticles, nanowires and carbon nanotubes. Pure Appl Chem 72:21–33

Rao SG, Huang L, Setyawan W, Hong S (2003) Large-scale assembly of carbon nanotubes. Nature 425:36–37

Rietman EA (2001) Molecular engineering of nanosystems. Springer, New York 158–185

Romano P, Polcari A, Verruso B, Colantuoni V, Saldarriaga W, Baca E (2007) Nonlinear current-voltage characteristics measured across circular deoxyribonucleic acid (DNA) molecule bundles. J Appl Phys 102:103720

Ross FM (2010) Controlling nanowire structures through real time growth studies. Rep Prog Phys 73:114501

Saavedra HS, Mullen JT, Zhang P, Devey DC, Claridge SA, Weiss PS (2010) Hybrid strategies in nanolithography. Rep Prog Phys 73:036501

Saif T, Alaca E, Sehitoglu H, Nano wires by self assembly, 16th IEEE Annual Conference on Micro Electro Mechanical Systems MEMS 03, Kyoto, Japan, 19–23 January 2003, pp. 45–47

Shenhar R, Norsten TB, Rotello VM (2004) Self-assembly and self-organization. In: Di Ventra M, Evoy S, Heflin JR Jr (eds) Introduction to nanoscale science and technology. Kluwer Academic Publishers, Dordrecht pp 41–74

Soldano C, Mahmood A, Dujardin E (2010) Production, properties and potential of graphene. Carbon 48:2127–2150

Sönmezolu S, Sönmezolu ÖA, Çankaya G, Yildirim A, Serin N (2010) Electrical characteristics of DNA-based metal-insulator-semiconductor structures. J Appl Phys 107:124518

Sounderya N, Zhang Y (2008) Use core/shell structured nanoparticles for biomedical applications. Recent Patents Biomed Eng 1:34–42

Sparreboom W, van den Berg A, Eijkel JCT (2010) Transport in nanofluidic system: a review of theory and applications. New J Phys 12:015004

Squires TM, Quake SR (2005) Microfluidics: fluid physics at the nanoliter scale. Rev Mod Phys **77**:977–1026

Tabeling P (2009) A brief introduction to slippage, droplets and mixing in microfluidic systems. Lab on a Chip 9:2428–2439

Tan Y-W, et al (2007) Temperature dependent electron transport in graphene. Eur Phys J Special Topics 148:15–17

Teh S-Y, Lin R, Hung L-H, Lee AP (2008) Droplet microfluidics. Lab on a Chip 8:108–220

Tian W-C, Finehout E (eds) (2008) Microfluidics for biological applications. Springer, Heidelberg

Tian X, Li J, Xu D (2010) Nanowire-based nanogap electrodes by annealing of multisegmented Pt/Au/Pt nanowire stir. Electrochem Commun 12: 1081–1083

Tseng AA, Chen K, Ma KJ (2003) Electron beam lithography in nanoscale fabrication: recent developments. IEEE Trans Electronic Packag Manufact 26:141–1949

Wacker A, Jauho A-P (1998) Quantum transport: the link between standard approaches in superlattices. Phys Rev Lett 80:369–372

Waigh T (2007) Applied biophysics: a molecular approach for physical scientists. Wiley, New York, p 206

Wang N, Zhang RQ (2008) Growth of nanowires. Mater Sci Eng R 60:1–51

Wei D, Liu Y (2010) Controllable synthesis of graphene and its applications. Adv Mater 22:3225–3241

Xu M, Enders RG, Arakawa Y (2007) The electronic properties of DNA bases. Small 3:1539–1542

Yang D, et al (2010) Novel DNA materials and their applications. Wiley Interdisciplin Rev Nanomed Nanobiotechnol 2:648–669

Ziaie B, Baldi A, Atashbar MZ (2004) Introduction to Micro/Nanofabrication. In: Bhushan B (ed), Springer handbook of nanotechnology. Springer, Berlin 147–184

Zhu W (ed) (2001) Vacuum microelectronics. Wiley, New York

http://www.springer.com/978-3-642-25571-7

Bionanoelectronics
Bioinquiring and Bioinspired Devices
Dragoman, D.; Dragoman, M.
2012, X, 254 p., Hardcover
ISBN: 978-3-642-25571-7