TOPICAL REVIEW

Properties and applications of high-mobility semiconducting nanotubes

T Dürkop, B M Kim$^1$ and M S Fuhrer$^2$

Department of Physics and Center for Superconductivity Research, University of Maryland, College Park, MD 20742, USA

E-mail: mfuhrer@physics.umd.edu

Received 30 December 2003
Published 23 April 2004
Online at stacks.iop.org/JPhysCM/16/R553
DOI: 10.1088/0953-8984/16/18/R01

Abstract

Experiments to determine the resistivity and charge-carrier mobility in semiconducting carbon nanotubes are reviewed. Electron transport experiments on long chemical-vapour-deposition-grown semiconducting carbon nanotubes are interpreted in terms of diffusive transport in a field-effect transistor. This allows for extraction of the field-effect and saturation mobilities for hole carriers, as well as an estimate of the intrinsic hole mobility of the nanotubes. The intrinsic mobility can exceed 100 000 cm$^2$ V$^{-1}$ s$^{-1}$ at room temperature, which is greater than any other known semiconductor. Scanned-probe experiments show a low degree of disorder in chemical-vapour-deposition-grown semiconducting carbon nanotubes compared with laser-ablation produced nanotubes, and show conductivity and mean-free-path consistent with the high mobility values seen in transport experiments. The application of high-mobility semiconducting nanotubes to charge detection and memory is also reviewed; it is shown that single electronic charges may be detected with a semiconducting nanotube field-effect transistor at operating temperatures up to 200 K.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction to carbon nanotubes 554
   1.1. Electronic structure: metallic and semiconducting nanotubes 554
   1.2. Predictions for transport experiments: mobility and mean-free-path 558

$^1$Present address: Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA 19104, USA.
$^2$Author to whom any correspondence should be addressed.
1. Introduction to carbon nanotubes

Multiwalled carbon nanotubes (MWNTs), initially dubbed ‘helical microtubules of graphitic carbon’ (Iijima 1991), were discovered by Iijima while researching methods of producing fullerenes. Using transmission electron microscopy (TEM) he discovered unusual fibres in the soot produced by an arc discharge between carbon electrodes, and identified them as seamless, concentrically nested, tubular sheets of graphite. Soon methods were developed to produce single-walled carbon nanotubes (SWNTs), i.e. nanotubes that consist of only a single sheet of graphite, by Iijima and Ichihashi (1993) and Bethune et al (1993). The production of SWNTs in large quantities by laser ablation (Thess et al 1996) spurred research on this material, revealing outstanding electrical, thermal and mechanical properties.

1.1. Electronic structure: metallic and semiconducting nanotubes

At the time of the discovery of carbon nanotubes (CNTs) there were already theoretical discussions of their possible atomic and electronic structure (Hamada et al 1992, Mintmire et al 1992, Saito et al 1992b). A carbon nanotube can be seen as a single sheet of graphene that has been rolled up along a direction given by the circumference vector \( c = na_1 + ma_2 \), or simply \((n, m)\), where \(a_1\) and \(a_2\) are the lattice vectors in graphene (see figure 1(a)). Depending on the indices \( n \) and \( m \) the nanotube can be either metallic or semiconducting (Hamada et al 1992, Saito et al 1992a, 1992b): if the difference between \( n \) and \( m \) is divisible by three the nanotube is metallic, otherwise it is semiconducting. This behaviour occurs because by rolling up a sheet of graphene the electron wavefunctions are subjected to an additional quantization condition, essentially ‘cutting’ 1D slices out of the formerly 2D band structure of graphene.

Here we present a simple picture of a carbon nanotube band structure which contains the relevant features needed for understanding the rest of this article. A simple tight-binding model for the \( \pi \) and \( \pi^\ast \) graphene bands was formulated by Wallace (1947). We further simplify this description by looking at just the low-energy portion of the band structure, which we approximate as cones with apices at the \( \mathbf{K} \) point in the graphene Brillouin zone. We take the slope of the cones to be a constant \( v_F \approx 8 \times 10^7 \text{ cm s}^{-1} \) (see figure 1(b)). This approximation ignores the ‘trigonal warping’ of the bands, i.e. the different slope of the bands in the \( \mathbf{K}–\mathbf{K} \) and \( \mathbf{K}–\mathbf{\Gamma} \) directions in \( k \)-space. It also fails for higher energy (and hence smaller diameter nanotubes) due to the non-linearity of the bands. The nanotube band structure is then obtained from the quantization of the circumferential wavevector, i.e. \( \mathbf{c} \cdot \mathbf{k}_F = 2\pi i \), where \( i \) is an integer,
Figure 1. Carbon nanotubes: connection between real-space and electronic structure. (a) Lattice of a single-atomic layer of graphite (graphene). The dashed red arrows indicate a common choice of coordinates corresponding to indices \((n, m)\) for describing the rolling up of graphene into nanotubes. The solid blue arrow is the roll-up vector for a metallic \((5, 5)\)-armchair nanotube. (b) Schematic image of the graphene band structure consisting of cones whose apices touch at the Fermi-level. The red lines demonstrate how the additional quantization condition introduced by rolling up the graphene into nanotubes cuts the band structure into 1D slices. This particular cut corresponds to a semiconducting tube; the cut does not intersect the cones’ apices. (c) Simplified model of the nanotube band structure for small \(k\). The curves are calculated from \(E = ((n\Delta)^2 + h^2v_F^2k^2)^{1/2}\), where \(\Delta\) is the band gap and \(n = 0, 1, 2, 3, \ldots\) counts the first subbands. The blue curves at \(n = 0, 3, 6, \ldots\) occur in metallic nanotubes; the red bands at \(n = 1, 2, 4, 5, \ldots\) occur in semiconducting tubes. (d) Examples for nanotubes with high symmetry. So-called armchair tubes, which are always metallic, have indices \((n, n)\). Tubes with indices \((n, 0)\) are semiconducting if \(n\) is not divisible by three. These tubes are often called ‘zigzag’ tubes. Both zigzag and armchair tubes have been treated extensively in theoretical publications, because their small unit cells allow for much easier computer simulations than the large unit cells of general \((n, m)\) tubes.

Effectively slicing the 2D band structure along lines where the quantization condition is met. Depending on whether or not these cuts pass through the apices of the cones the 1D band structure is metallic or contains a bandgap. It can be shown that the condition for metallicity is \(n - m = 3i\), with \(i\) an integer, thus 1/3 of the nanotubes are predicted to be metallic, and 2/3 semiconducting.

Figure 1(c) illustrates the resulting bands. The red lines are the bands and the first subbands of a semiconducting nanotube located at \(\pm \Delta\) and \(\pm 2\Delta\) from the Fermi level where \(\Delta\) is one-half of the bandgap, i.e. \(\Delta = E_g/2\). The blue lines are the bands present in a metallic nanotube with the same diameter as the semiconducting nanotube. Here the first subbands are at \(\pm 3\Delta\). Finally figure 1(d) shows two special cases of nanotubes, which are especially suited for theoretical investigations because of their high degree of symmetry. The first one is a so-called armchair nanotube with the indices \((n, n)\), which is always metallic. The second is called a zigzag nanotube with \((n, 0)\). Zigzag nanotubes are semiconducting unless the index \(n\) is divisible by three.

Within this picture the bandgap can be seen to depend inversely on diameter; as the diameter becomes smaller, the spacing between the lines in \(k\)-space for which the quantization condition is met becomes larger, and hence the lines lie further from the metallic apex of the cone. For semiconducting nanotubes the result is a bandgap that to first order is given by
Figure 2. Examples of the gate-voltage dependence of the conductance through a SWNT at room temperature for a metallic nanotube (a) and a semiconducting nanotube (b).

\[ E_g = 2\Delta = 0.7/d \text{ eV nm}^{-1}, \]  
where \( d \) is the nanotube diameter independent of the particular chirality (Kane and Mele 1997). If one takes into account the curvature of the graphene sheet in a nanotube there is an additional contribution to the gap proportional to \( d^{-2} \), which also causes metallic nanotubes with indices other than \( (n,n) \) to have a small bandgap of around 0.05 eV or less (Hamada et al 1992, Kane and Mele 1997).

After these theoretical predictions had been made it took several years before it was possible to verify them experimentally by making electrical contacts to single nanotubes and test whether they showed metallic or semiconducting behaviour. (A detailed description of nanotube devices and their fabrication is given in section 2.2.) The first devices made with single metallic nanotubes or bundles of a few nanotubes were presented by Tans et al (1997) and Bockrath et al (1997). Their devices showed Coulomb-blockade effects at low temperatures and Ohmic behaviour without dependence on applied gate voltage at higher temperatures (for a good review of Coulomb-blockade and other low-temperature effects see Kouwenhoven et al (1997)). The first investigation of single semiconducting SWNTs was published by Tans et al (1998). Their devices show the oft-reproduced behaviour of a field-effect-transistor (FET) at room temperature that makes semiconducting SWNTs so interesting for possible microelectronics applications: the conductivity of these devices can be turned off by applying a gate voltage.

Figure 2(a) shows a plot of the gate voltage dependence of the conductance of a metallic nanotube device at room temperature. While applying a constant bias-voltage along the nanotube, the gate voltage is being varied. The current is largely independent of gate voltage. Figure 2(b) shows the same plot for a device made from a semiconducting nanotube. Clearly this device can be turned off by applying a positive gate voltage; conductance changes of over five orders of magnitude between on and off states have been seen (Martel et al 1998).

Although transport experiments can demonstrate the existence of an energy gap (Tans et al 1998) and can provide information about the low-lying excitations in the nanotube (Bockrath et al 1997, Tans et al 1997), they do not give the details of the band structure, for example the onset of various subbands. One method of verifying the theoretical predictions (Hamada et al 1992, Mintmire et al 1992, Saito et al 1992b) is to use tunnelling measurements probing the density of states (DOS). The first measurements of this kind on SWNTs were published by Odom et al (1998), Wildöer et al (1998). In a 1D electron system like the one
suggested for carbon nanotubes the onset of new subbands corresponds with the occurrence of singularities in the DOS (Saito et al 1992a). At these van Hove singularities the DOS as a function of energy $E$ diverges with an $E^{-1/2}$-dependence. Figure 3 shows examples of scanning tunnelling spectroscopy (STS) measurements presented in Wildöer et al (1998). The position of the singularities provides information about the size of the bandgap and the position of the subbands, and thus allows us to identify and distinguish various chiralities of nanotubes, i.e. together with real-space scanning tunnelling microscopy (STM) images it allows determination of the circumference vector $(n, m)$ for individual nanotubes. Unfortunately this type of experiment cannot be carried out on nanotubes that have been incorporated into devices because the STS and STM measurements require the nanotubes to reside on a conducting surface (e.g. gold) whereas transport measurements are only possible if the nanotubes are on an insulating surface, most commonly SiO$_2$.

Additional insight into the band structure can be gained from optical experiments. Because the Raman scattering from nanotubes is greatly enhanced when the incident photon energy is resonant with a transition between van Hove singularities in the density of states, Raman spectroscopy gives information about the electron energy spectrum as well as the phonon spectrum of nanotubes (Dresselhaus and Eklund 2000, Saito and Kataura 2001). Recent progress has allowed Raman spectroscopy investigations on individual nanotubes (Jorio et al 2001); combined knowledge of the electron and phonon spectra allow unique determination of the $(n, m)$ indices of individual nanotubes. Transitions between van Hove singularities are also observed in the excitation and emission spectra in fluorescence experiments on semiconducting nanotubes in solution. Recently the position of absorption and emission peaks in fluorescence spectroscopy have been used to determine the $(n, m)$ indices of semiconducting nanotubes (Bachilo et al 2002). Raman spectroscopy and fluorescence spectroscopy have since been performed on single nanotubes to verify the $(n, m)$ assignments (Hartschuh et al 2003).
1.2. Predictions for transport experiments: mobility and mean-free-path

Within the Drude model the conductivity $\sigma$ is defined by $\sigma = j / E = ne^2 \tau / m$, where $j$ is the current density, $E$ the electric field, $n$ the charge carrier density, $e$ the absolute value of the electronic charge, $\tau$ the momentum scattering rate and $m$ the effective mass of the carriers. In one dimension, we have $\sigma = GL$ and $n = N / L$, where $G$ is the conductance, $L$ the length of the sample and $N$ the number of carriers. The quantity $\mu = et / m = \sigma / ne$ is the mobility, and has the same units in any dimension.

Mobility measures the momentum scattering rate of carriers, and therefore is of fundamental interest in understanding the scattering processes in a given system. In addition, the mobility of charge carriers in the FET channel is one of the most important parameters determining the performance of an FET. Mobility ultimately determines the high-frequency performance of the FET, and is also important for the determination of the transconductance and drive current. Nanotube FETs (NT-FETs) have been proposed for a variety of applications including chemical sensors (Besteman et al 2003, Chen et al 2003, Star et al 2003) and single-electron memories (Fuhrer et al 2002, Kim et al 2002a); in these applications mobility determines the sensitivity of the NT-FET to charge or chemical species.

In this section we will discuss theoretical and experimental results on the conductivity of metallic and semiconducting nanotubes. In section 3.1.2 we will discuss the determination of the mobility directly from the transistor characteristics of a NT-FET, and when this analysis is applicable.

Mobility is difficult to define for a metallic nanotube, since it is not clear whether the relevant charge density should be measured from the subband bottom or the band bottom. In addition, the nearly flat bands at the Fermi level imply a near-infinite effective mass. However, the mean free path may be defined unambiguously, and we will use this quantity to compare theoretical and experimental results for the conductivity of metallic and semiconducting nanotubes.

In one dimension, the maximum conductance is quantized: in the absence of scattering the conductance is $G_{\text{max}} = e^2 / h$ (Datta 1995) for each conductance mode, where $h$ is Planck’s constant. For the nanotube, with two bands with two spins, $G_{\text{max}} = 4e^2 / h$. In an imperfect wire with scattering (neglecting quantum interference) the conductance is given by $G^{-1} = G_{\text{max}}^{-1} + G_{\text{wire}}^{-1}$, where $G_{\text{wire}} = G_{\text{max}} T / (1 - T)$ and $T$ is the transmission probability for the wire (Datta 1995). The mean free path $l$ is the length of wire over which $T = 1/2$ or $G_{\text{wire}} = G_{\text{max}}$. This allows us to relate the conductance to the mean free path by the relation $l / L = G_{\text{wire}} / G_{\text{max}}$, where $L$ is the length of the nanotube. The mobility is related to the mean free path: $\mu = G_{\text{wire}} L / ne = G_{\text{max}} l / ne$.

The observation of conductances which are an appreciable fraction of $G_{\text{max}}$ in metallic nanotubes with lengths exceeding 5 $\mu$m (Kong et al 1999) indicates mean free paths of at least a few micrometres. This is in good agreement with theoretical predictions of 10 $\mu$m or more (White and Todorov 1998). Similarly, high conductances in semiconducting nanotubes (Rosenblatt et al 2002, Javey et al 2003) indicate mean-free-paths of at least hundreds of nanometres. These results have also been corroborated by electrostatic force microscopy (EFM) as described in section 3.2.2.

Only a small number of publications have attempted to predict the mobility in semiconducting nanotubes. McEuen et al (1999) made general arguments that backscattering would be greatly suppressed in metallic nanotubes relative to semiconducting nanotubes: interband backscattering is suppressed by a symmetry in metallic nanotubes which is not present in semiconducting nanotubes at low doping. However, at high doping semiconducting nanotubes should be expected to have similar conductivity to metallic nanotubes. From this
argument one can make a simple estimate of the mobility in semiconducting nanotubes: assuming that the conductivity approaches the metallic value for an amount of doping that pushes the Fermi level to $E_F = 2\Delta$ (i.e. the nanotube is doped to the second subband), we may estimate $\mu$ from the metallic conductivity and the carrier density $n$ necessary to achieve that doping level. Approximating the Fermi wavevector $k_F$ in the semiconducting nanotube by $k_F$ in the metallic nanotube at $E_F = 2\Delta$, and using the relationship between the carrier density $n$ and $k_F$ (for a single 1D subband: $n = \frac{2k_F}{\pi}$), we get

$$n = \frac{4k_F}{\pi} = \frac{8\sqrt{3}\Delta}{h v_F}. \quad (1)$$

Using the expression $\Delta = 0.7/d \text{ eV nm}^{-1}$ from above and the Fermi-velocity $v_F = 8.1 \times 10^5 \text{ m s}^{-1}$ (Cobden et al 1998) of metallic tubes we get $nd = 2.9$. Using the mean free path $l = 3 \text{ \mu m}$ of a metallic nanotube we can use $\mu = G_{\text{max}}/ne$ obtaining mobilities of $\mu = 10 000$–$50 000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ for commonly observed tube diameters in the range between 1 and 5 nm. This mobility is high, but very reasonable considering the mobilities of 15 000 and 20 000 $\text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ for holes and electrons, respectively, in graphite (Dresselhaus et al 1996).

Pennington and Goldsman (2003) have used a semiclassical model to more carefully investigate electron–phonon coupling in semiconducting nanotubes. From this model they derive the electron drift velocity for various diameters of zigzag nanotubes, as shown in figure 4 plotted versus applied electric field. The mobility may be obtained from $\mu = v_d/E$; clearly $\mu$ increases with nanotube diameter, yielding a maximum mobility of 120 000 $\text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ for a (59, 0) nanotube.

Pennington and Goldsman also found a strong dependence of the mobility on nanotube diameter. The low-field scattering time was found to be roughly proportional to diameter, $\tau \sim d$. The effective mass at the band bottom is inversely proportional to diameter $m^* \sim 1/d$, so the mobility $\mu = et/m^* \sim d^2$. It is likely that this relationship fails for large nanotube
diameters when the subband spacing $\Delta$ becomes smaller than the temperature or some disorder strength. Still, this analysis suggests that larger diameter nanotubes may be optimal for high-mobility applications.

At higher carrier velocities, it is expected that more channels will be open for scattering, and the carrier mean free path will be reduced. This is evident in the calculations of Pennington and Goldsman (see figure 4), where a peak in the drift velocity is seen as a function of electric field. Such a reduction in mean free path has been observed in metallic nanotubes (Yao et al 2000) and interpreted as due to the efficient emission of $\sim 160$ meV zone-boundary phonons. Park et al (2003) used a scanned contact probe to verify that the mean free path for phonon scattering at low velocities in a metallic nanotube is a few micrometres, while at high velocities it is shortened to $\sim 10$ nm.

2. Nanotube synthesis and device fabrication

2.1. Nanotube synthesis

There are two fundamentally different types of method of growing carbon nanotubes for device fabrication. The first class of methods tends to produce large amounts of nanotubes, usually as nanotube-containing soot, from which the nanotubes are deposited onto substrates for device fabrication after being purified. The second class of methods synthesizes the nanotubes directly on a substrate by first depositing some kind of catalyst and then exposing it to carbon-containing feedstock gas.

The methods of the former kind were the first available procedures for nanotube synthesis. The initial discovery of multiwalled nanotubes (Iijima 1991) was made using an arc-discharge between two carbon electrodes. The earliest methods for the production of SWNTs (Bethune et al 1993, Iijima and Ichihashi 1993) also used an arc-discharge method. The first method used for the production of large quantities of CNTs (Thess et al 1996) was laser-ablation; in this method a carbon target doped with a catalyst consisting of powdered Ni and Co is heated to $1000^\circ C$ in vacuum and then bombarded with laser pulses resulting in a high yield of SWNTs with a narrow diameter distribution.

More recently the so-called HiPCO method (Bronikowski et al 2001) has further increased the quantities in which nanotubes can be produced. The acronym HiPCO stands for high-pressure decomposition of CO gas. Here CO gas with a small amount of added Fe(CO)$_5$ is sprayed through a nozzle into a reactor at 1050$^\circ C$ and 30 atm. The Fe(CO)$_5$ provides the metal to form tiny catalyst droplets while the CO is the feedstock material.

All these methods produce relatively large quantities of SWNTs that can be of high quality with few defects if the synthesis parameters are properly optimized. However, one common feature is the contamination of the products with amorphous carbon (‘soot’) and leftover catalyst. Furthermore the nanotubes tend to occur in thick bundles of several tens to several hundreds of nanotubes. For the fabrication of electronic devices (see section 2.2) the nanotubes need to be separated from each other and cleaned from the amorphous carbon residue. Although there are many ways of cleaning the tubes, the most prominent method is that described in Liu et al (1998), which purifies the nanotube soot using nitric acid, the resulting nanotubes are usually fairly short (not more than a few micrometres) and often contain a significant number of defects due to the harsh cleaning procedures.

For electronics devices it is often desirable to investigate longer nanotubes and—especially if one is interested in the intrinsic properties of nanotubes—highly defect-free nanotubes. For this type of application growth methods producing CNTs directly on substrates (Dai et al 1996a, Fonseca et al 1997, Kong et al 1998a) are often better suited. In particular Dai et al
have done much pioneering work on such CVD (chemical vapour deposition) methods. Initially the catalyst for the CVD process was prepared by a rather time-consuming and complicated method of impregnating $\text{Al}_2\text{O}_3$ nanoparticles with Fe, Co, Ni or Mo or their compounds; those particles were then deposited on the substrates.

Hafner et al (2001) simplified the process of preparing the catalyst substantially by simply dipping the substrates into a solution of $\text{Fe(NO}_3\text{)}_3$ in 2-propanol and then into hexane, forcing the catalyst to precipitate onto the substrate creating nanometre-sized clusters. Another interesting route was introduced in Kong et al (1998b): by using standard electron-beam lithography techniques it was possible to create a mask to deposit catalyst only in predefined locations.

Once the catalyst is prepared using the appropriate method, the substrates with the catalyst on top are placed in a furnace heated to between 600 and 1000 °C, depending on the particular recipe. After that, feedstock gas (e.g. CH$_4$, CO or C$_2$H$_4$), often together with an inert carrier gas (e.g. Ar, or N$_2$) and H$_2$ to reduce catalyst-compounds, are flowed through the furnace. Due to the high temperatures and the strongly reducing environment during the growth process not all substrates are suitable for nanotube CVD. While thick thermal SiO$_2$ on Si is the most common substrate for CVD growth of nanotubes, a variety of other substrates have been used including high-$\kappa$ dielectrics (e.g. strontium titanate (Kim et al 2004)) and metals (e.g. molybdenum (Franklin et al 2002)).

Recently CVD-methods have been improved to allow for the growth of nanotubes of lengths of several hundred micrometres to over a millimetre (Kim et al 2002b, Huang et al 2003). Figure 5 shows examples of CNTs grown with CVD methods following various recipes.
2.2. Nanotube device fabrication

Although there were some earlier attempts at contacting CNTs (Dai et al 1996b, Langer et al 1996) the first publications showing the successful contacting of single SWNTs (or bundles of a few nanotubes) were (Bockrath et al 1997, Tans et al 1997). In both cases the authors used an organic solvent (dichloroethane or acetone) to produce a suspension of nanotubes and deposited this suspension on a degenerately doped Si-chip capped with SiO$_2$. In Bockrath et al (1997) alignment markers on the chips made it possible to locate the nanotubes with atomic force microscopy (AFM) and then contact them by using electron-beam lithography to deposit Cr/Au contacts onto the nanotubes. Here the Cr is a thin (3 nm) layer that promotes adhesion of the thicker (50 nm) gold, which forms the actual current leads. In Tans et al (1997) the nanotube suspension is deposited on a chip with pre-patterned Pt contacts. The same method was used to fabricate the first nanotube FETs using a semiconducting CNT (Tans et al 1998), again with Pt contacts and (Martel et al 1998) with Au contacts. Figure 6 shows a schematic of a typical nanotube device and an AFM image of such a device. It has a source and a drain contact touching the nanotube and employs the highly doped Si substrate as a so-called back-gate that allows the nanotube electronic structure to be influenced by an applied gate voltage. The gate is insulated from the nanotube by the SiO$_2$ layer. While the use of nanotubes deposited from suspension allows for both depositing contacts onto the nanotubes and depositing the nanotubes on top of the contacts, CVD methods that grow CNTs directly on the substrate require the contacts to be put on top of the nanotubes after locating them. Besides the use of AFM for locating the CNTs it is also possible to use field-emission scanning electron microscopy (FE-SEM) (Brintlinger et al 2002) which cuts the time needed for locating nanotubes by a factor of ten.

Recent improvements have allowed local gate contacts either beneath the nanotubes (Bachtold et al 2001) or on top of the nanotubes (Javey et al 2002, Wind et al 2002) which makes it possible to have several nanotube devices on one chip and address them separately.

Several studies have tried to optimize the material used for the contacts. From the initial choices of material (Cr/Au or Pt) only the Cr/Au contacts have been used widely. It has been found that a wetting layer of Ti (Zhang and Dai 2000) allows deposition of smooth films of many metals onto carbon nanotubes because Ti, especially when annealed, forms titanium carbide at the interface with the nanotube. For this reason Ti/Au contacts are another frequently used combination of contact materials. Many publications investigating Schottky barriers between a nanotube and its contacts (e.g. Appenzeller et al 2002, Heinze et al 2002 have employed this kind of contact). Pd is another material investigated in Zhang and Dai (2000) that wets nanotubes well. It has been used in a recent publication (Javey et al 2003) to produce NT-FETs with Ohmic contacts (here and elsewhere in this paper we use ‘Ohmic contacts’ to mean contacts with zero or negative Schottky barrier).
3. Experiments on single-walled carbon nanotubes

3.1. Transport in single-walled carbon nanotube field-effect transistors

Despite considerable research effort the details of electronic transport in semiconducting SWNTs are not yet clear. The initial publications showing experimental data from NT-FETs already disagreed upon the explanation of the transistor behaviour. While Tans et al (1998) developed a model similar to a BARITT-diode (Sze 1981), which basically consists of two back-to-back Schottky barriers, the description initially given by Avouris and co-workers (Martel et al 1998) is based on the model of a conventional MOSFET with a diffusive channel (Sze 1981). From the characteristics of a NT-FET (analogous to a p-channel MOSFET, see figure 2(b)) it is clear that the charge carriers have to be holes. However, in the model presented in Tans et al (1998) the holes are injected into the nanotube from the contacts, whereas according to Martel et al (1998) they are intrinsic to the nanotube. Martel used the gate-voltage-dependent conductance of his NT-FETs to calculate a mobility of $\sim 20 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, much lower than the mobility of p-silicon ($\sim 450 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$). As shown below, this mobility estimate was incorrect, due to the influence of Schottky barriers at the contacts. Recently Ohmic contacts to semiconducting nanotubes have allowed the intrinsic transport properties to be probed.

3.1.1. Ohmic or Schottky contacts? While researchers had been able to produce devices from metallic nanotubes that were close to the fundamental limit of resistance for nanotubes ($6.5 \text{ k} \Omega$) (Liang et al 2001), until recently the lowest resistances reported for semiconducting nanotubes were some 100 $\text{k} \Omega$, with 1 M$\Omega$ being more typical, suggesting the presence of some kind of barrier, either in the semiconducting nanotube itself (McEuen et al 1999), or at the contacts. Furthermore, several experiments corroborated the idea of Schottky contacts directly. Freitag et al (2001) found that an AFM tip with applied positive bias-voltage placed near one of the contact of a NT-FET with Cr/Au contacts had much stronger influence on the device behaviour than a tip placed elsewhere on the device. Bachtold et al (2001) measured the behaviour of Au-contacted CNTs on top of Al/Al$_2$O$_3$ microstrips acting as gates, with good agreement of the device characteristics with theoretical models based on the Schottky-contact picture developed in Léonard and Tersoff (2000) and Odintsov (2000). Similarly, Derycke et al (2002) and Martel et al (1998) found evidence for barriers at the contacts of their NT-FETs. Gathering all this information Heinze et al (2002) presented theoretical predictions about the behaviour of NT-Schottky-barrier transistors which are then found to be followed well by Ti/Au-contacted devices by Appenzeller et al (2002).

Figure 7 shows an overview of the results of Appenzeller et al. Figure 7(a) describes the subthreshold behaviour for a NT-FET, showing the source–drain current $I$ as a function of applied gate voltage $V_g$ for three different source–drain voltages $V_{sd}$. The current depends on applied $V_{sd}$ as expected for a Schottky-barrier transistor. The subthreshold swing, which is defined as $S = (d \log G/dV_g)^{-1}$ (mV/decade), is shown in the inset in figure 7(a). For conventional MOSFETs it is expected to be proportional to the temperature (Sze 1981): $S \approx 2.3 k_B T/e$, where $k_B$ is Boltzmann’s constant and $e$ the electron charge. For a Schottky-barrier FET in which tunnelling through the Schottky barrier dominates the current, $S$ is expected to be largely temperature independent (Appenzeller et al 2002, Heinze et al 2002) as observed here (at least below 200 K), though a weak temperature dependence can result from thermally assisted tunnelling (Appenzeller et al 2004). Finally the scaling of $S$ with gate-oxide thickness is different for Schottky-barrier FETs and conventional FETs. Figure 7(b) shows how the values for $S$ obtained by Appenzeller et al and other groups follow the expected behaviour for Schottky-barrier transistors much better than the behaviour expected
Figure 7. Schottky barriers in NT-FETs from Appenzeller et al (2002). (a) Subthreshold swing $S$ in the $I-V_g$ curve for a back-gated device on SiO$_2$ under different bias voltages and at different temperatures (inset). The temperature independence (below 200 K) of $S$ and the bias-voltage dependence of the current in the subthreshold regime indicate the presence of Schottky barriers. (b) Scaling of $S$ with effective oxide thickness (oxide thickness corrected for dielectric constant of the materials used) for back-gated devices of different lengths, on different dielectrics and prepared by different groups. The scaling agrees with a Schottky barrier model but not with a traditional MOSFET model. (c) Band structure schematic for a Schottky barrier NT-FET (a) without and (b) with applied bias voltage. (Adapted from Appenzeller et al (2002) with permission from the authors.)

for traditional MOSFETs. All this evidence taken together suggests that many NT-FETs are Schottky-barrier transistors with a band structure similar to the one shown schematically in figure 7(c) (also adapted from Appenzeller et al (2002)).

However, it is not clear that all NT-FETs should have Schottky barriers at the electrodes. It was shown (Léonard and Tersoff 2000) that the work function difference between the metal electrode and nanotube is dominant in determining whether a Schottky barrier is present; in the one-dimensional nanotube interface dipoles cannot be completely screened (Léonard and Tersoff 2002). Thus the choice of a large work-function metal should provide Ohmic contact to the valence band. The work function of nanotubes is about 4.5 eV (Tans et al 1998), so a metal with work function greater than $\sim 4.5$ eV + $E_g/2$ should provide Ohmic contact to the valence band.

Palladium is a promising material with a high work function that at the same time wets nanotubes (Zhang and Dai 2000). Javey et al (2003) fabricated NT-FETs with Ohmic contacts by depositing Pd-contact pads onto CNTs and annealing the devices in Ar. Figure 8(a) shows a diagram of these Pd-contacted NT-FETs, while figure 8(b) shows $I-V_g$ curves of such a device and the inset shows the signature of Fabry–Pérot oscillations measured at low temperatures, analogous to those seen in metallic nanotubes for very low contact resistances.
Figure 8. NT-FET with Ohmic palladium contacts from Javey et al (2003). (a) Device image with Ti/Au bonding pads and Pd NT contacts. (b) Device $G-V_g$ curves showing device conductance up to 0.5$G_0$. The inset shows Fabry–Pérot type resonances at low temperature, which indicates highly transmissive contacts (Liang et al 2001). (c) By exposure to H$_2$ gas the work function of the Pd contacts is lowered, decreasing the device conductance and creating Schottky barriers (d). (Adapted from Javey et al (2003) with permission from the authors.)

(Liang et al 2001). The nanotube conductance shown here reaches close to one half of the theoretical limit, suggesting almost barrier-free contacts. Figure 8(c) shows how the device conductance can be significantly lowered by exposing the FET to hydrogen gas. Since H$_2$ gas is known to lower the work function of Pd (Mandelis and Christofides 1993) this creates barriers at the contacts. Finally figure 8(d) shows a schematic of the band structure of a NT-FET with no, little and large Schottky barrier after various degrees of H$_2$ exposure. Pure Au has also been shown to form an Ohmic contact with an NT-FET (Yaish et al 2004), and likewise annealed (Yaish et al 2004) or as-deposited (Dürkop et al 2004) Cr/Au contacts may also form contacts without Schottky barriers. To date the nature of the nanotube/metal interface still remains a subject of intense interest, and it is likely that the complete story is not yet known.

3.1.2. Measuring the mobility in a semiconducting nanotube transistor. While recent research has shown that NT-FET behaviour may be understood in terms of Ohmic or Schottky contacts, the nature of conduction in the nanotube channel is still not understood. Especially in devices with strong Schottky barriers the assumption of ballistic conduction serves well to describe device properties, which are dominated by the said Schottky barriers (Heinze et al 2002). The same assumption is reasonable for short ($\sim$300 nm) devices with Ohmic contacts, but not necessarily for longer ones (Javey et al 2003). Apparently semiconducting nanotubes show a transition from ballistic conductance to diffusive conductance at a length scale of
a few micrometres. It is not yet known what scattering mechanisms govern this regime of diffusive conductance. One of the parameters that allows us to draw conclusions about the details of the conductance is mobility.

The dependence of the mobility on temperature, applied voltages and other physical quantities provides information about scattering mechanisms, charge density, impurities and many materials parameters. For comprehensive reviews of mobility in semiconductors see for example Sze (1981), Schrader (1998), Seeger (2002). There are several different methods of measuring mobility. The most common method measures the so-called Hall-mobility. This method, however, cannot be used in 1D systems such as carbon nanotubes. Here one has to use methods that are based on the behaviour of devices, especially NT-FETs: following the Drude model the mobility is given by $\mu = \sigma / q$. In a 1D system the conductivity is given by $\sigma = GL$, where $G$ is the device conductance and $L$ its length. The charge density $q$ can be calculated from the capacitance per length $c_g$ between the device and the gate controlling the device and the applied gate voltage. Allowing for a non-zero threshold this gives $q = c_g(V_g - V_{th})$. We also assume that the gate capacitance is much smaller than the quantum capacitance of the nanotube (see below), such that the quantum capacitance may be neglected. The relationship ignores thermally activated carriers, which may be significant for small $(V_g - V_{th})$ and high temperature. Most importantly, this relationship assumes $c_g$ and $V_{th}$ do not vary appreciably along the length of the channel, which requires $L \gg t$, the dielectric thickness. Few nanotube devices studied in the literature satisfy this last criterion.

From this approach we get the expression for the mobility that is closest to the intrinsic mobility of a nanotube:

$$\mu = \frac{L}{c_g} \frac{G}{V_g - V_{th}}.$$  \hspace{1cm} (2)

Here $G$ should be the conductance of the channel only. It is appropriate to approximate $G$ by the total device conductance here only if the channel conductance is much smaller than the contact conductance, i.e. this formula provides meaningless results in the case of Schottky-barrier transistors, in which the contact conductance is small. This formula for the mobility is analogous to what is known as the effective mobility in conventional FETs (Schrader 1998). It is only applicable if $V_{th}$ can be determined unambiguously. If this is not the case one can still calculate the so-called field-effect mobility:

$$\mu_{FE} = \frac{L}{c_g} \frac{\partial G}{\partial V_g}.$$  \hspace{1cm} (3)

Equations (2) and (3) are equivalent only in the case that $G$ is linearly proportional to $V_g - V_{th}$. However, in many materials, $G$ is sub-linear in $V_g$, and equation (3) typically underestimates the mobility.

The third method of measuring the mobility is not as reliable as the first two, but since it probes different aspects of the device behaviour it is useful to independently verify the results obtained with the first two formulae. This approach measures the saturation mobility $\mu_{sat}$ using the saturation current at high bias-voltages:

$$\mu_{sat} = \frac{2L}{Bc_g} \frac{I_{d,sat}}{(V_g - V_{th})^2}.$$  \hspace{1cm} (4)

The main problem with this approach is the presence of the body factor $B$, which accounts for dependence of $V_{th}$ on the position along the device. Even for conventional MOSFETs $B$ is not
Figure 9. Ultralong NT-FET. (a) FE-SEM image of a 325 $\mu$m long device. The scale bar is 100 $\mu$m long. (b) Conductance $G$ as a function of gate voltage $V_g$ curves for this device at different temperatures. The hysteresis in the curves is discussed in section 5. Only the decreasing-$V_g$ portion of the data is used for further analysis in this figure and figure 10. (c) $G^2$ versus $V_g$ for this device. Straight lines indicate $G \sim (V_B - V_{th})^{1/2}$ and are used to calculate $V_{th}$. (d) Subthreshold behaviour of this device. The temperature dependence of the subthreshold swing $S$ is shown in the inset (see also Dürkop et al 2004).

very well understood, much less for carbon nanotubes. Whenever this formula is applied to NT-FETs below, it is assumed that $B$ is unity. Normally the saturation mobility underestimates the mobility (Schroder 1998).

Due to the requirements to have channel length much greater than dielectric thickness, and negligible Schottky barriers at the contacts, there are few published results of measurements of mobility in nanotubes. Recently Dürkop et al (2004) were able to fabricate very long ($L > 300 \mu$m) semiconducting nanotube FETs in which the channel resistance dominated the transport through the device. Figure 9(a) shows such a device. The length $L = 325 \mu$m between contacts, and the diameter $d = 3.9$ nm. Figure 9(b) shows the conductance of the device as a function of gate voltage. As can be seen from figure 9(c) the conductance $G$ of this device follows the empirical relationship $G^2 \sim (V_{th} - V_B)$, which allows fitting to determine the threshold voltage $V_{th}$.

It is notable that the conductance at $V_B = -10$ V exceeds 1.4 $\mu$S at room temperature. Assuming zero contact resistance, i.e. $G = G_{wire}$, this conductance corresponds to a 1D conductivity $\sigma$ of $4.6 \times 10^{-8}$ S cm (finite contact resistance would imply greater $\sigma$). If this nanotube is single-walled, or multi-walled with the current largely carried by the outer wall at low bias (Collins et al 2001), then the electronic mean-free path $l$ is given by $\sigma / 2 G_0 = 2.9 \mu$m, where $G_0$ is the conductance quantum, approximately 77.5 $\mu$S. Finite contact resistance would imply a larger $\sigma$ and, thus, a larger $l$. This analysis already suggests that the measured resistance is the nanotube resistance; if contact resistance (e.g. from Schottky barriers) was dominant this would imply a much longer mean-free path.
Figure 10. Mobility calculations for the same device as in figure 9. (a) Intrinsic mobility according to equation (1). (b) Field-effect mobility according to equation (2). (c) $I-V$ curves in the saturation regime at 1.5 K. The saturation current $I_{sat}$ has been determined from the intersection of the Ohmic behaviour for low $V_{sd}$ and the saturation behaviour. (d) Saturation current as a function of applied gate voltage, showing the $V_g^2$ behaviour expected from traditional MOSFETs (see also Dürkop et al 2004).

The subthreshold behaviour (figure 9(d)) provides evidence that the channel resistance dominates the device resistance over the entire range of measured gate voltage; i.e. the contacts are Ohmic, or the Schottky-barrier resistance is negligible over the range of gate voltages probed. This is also consistent with the observation that the relation $G^2 \sim (V_{th} - V_g)$ holds over a wide range of $G$, i.e. the same behaviour governs the device from high conductance to turn-off.

The mobility of the device in figure 9 may then be calculated using the above methods. Figure 10 summarizes the results of these calculations; all mobilities calculated here are hole mobilities. The capacitance between the device and the gate which is necessary for any determination of mobility was calculated using a commercial computer simulation. For a device with 3.9 nm diameter the capacitance is 190 fF cm$^{-1}$. This value is substantially lower than the commonly used analytical expression for a conducting cylinder above a plane completely embedded in a dielectric (Javey et al 2002). Strictly speaking, the total gate capacitance is given by $C_{g,tot} = (c_g^{-1} + c_Q^{-1})^{-1}$, where $c_g$ is the electrostatic gate capacitance, and the quantum capacitance $c_Q = \epsilon^2 D(E)$, where $D(E)$ is the density of states, and is approximately 4 pF cm$^{-1}$ in nanotubes (Guo et al 2002, Rosenblatt et al 2002). The quantum capacitance corrects for the fact that the electrochemical potential in the nanotube shifts with
added charge. Since the two contributions to the gate capacitance add inversely, the smallest capacitance limits the total capacitance. In the devices discussed here, $c_g \ll c_Q$ and hence $c_{g,\text{tot}} \approx c_g$. Quantum capacitance is, however, important in devices with thin high-$\kappa$ gate dielectrics (Javey et al 2002, Kim et al 2004).

Figure 10(a) shows the intrinsic mobility calculated according to equation (1). As expected from the power-law behaviour of the conductance it shows a power-law dependence itself, reaching values of over 100 000 cm$^2$ V$^{-1}$ s$^{-1}$ for low gate voltages. The field-effect mobility according to equation (2) is shown in figure 10(b). Like in conventional MOSFETs (Schroder 1998), it peaks for low gate voltages. For this device the peak is at 79 000 cm$^2$ V$^{-1}$ s$^{-1}$ at room temperature. Finally, figures 10(c) demonstrates the saturation behaviour of the device $I$–$V$ curves, and figure 10(d) shows the square-law behaviour of equation (3) (i.e. $I_{\text{sat}} \sim V_g^2$) used to extract the saturation mobility of 55 000 cm$^2$ V$^{-1}$ s$^{-1}$. Because of the length of this device, the electric field is less than 300 V cm$^{-1}$, much lower than the fields at which significant mobility decreases are expected (Pennington and Goldsman 2003). For comparison, the Ohmically contacted, much shorter devices investigated in Javey et al (2003) show a saturation mobility of 4000 cm$^2$ V$^{-1}$ s$^{-1}$; however, it is possible that contact resistance plays a role here, and it is also not clear that the devices are in the low-field limit.

The mobility values in semiconducting carbon nanotubes exceed those of other materials and devices when measured at room temperature. The highest intrinsic mobility at room temperature of any semiconductor is the electron mobility in InSb, 77 000 cm$^2$ V$^{-1}$ s$^{-1}$ (Hrostowski et al 1955) (measured as Hall-mobility). Typical field-effect mobilities of Si devices are around 1000 cm$^2$ V$^{-1}$ s$^{-1}$ (Takagi et al 1994). The semiconductor with the highest hole mobility is PbTe with 4000 cm$^2$ V$^{-1}$ s$^{-1}$ (Sze 1981).

3.2. Scanned probe experiments

While transport experiments as described in the previous section can provide much information about the properties of a nanotube as a whole, they cannot test for local variations along the nanotube, for example to test for the influence of defects on the electronic structure or the shape of the potential drop along a nanotube. On insulating substrates, which are prerequisite for the fabrication of nanotube-devices, it is virtually impossible to use techniques like STM that rely on electric current for their feedback mechanisms. The methods of choice for imaging nanotube devices, therefore, are techniques derived from AFM.

3.2.1. Scanned gate microscopy. In scanned gate microscopy (SGM) a voltage is applied to a conducting AFM tip (see figure 11(a)). Thus, the tip—without touching the nanotube—acts as a local gate to the devices being imaged (Bachtold et al 2000, Tans and Dekker 2000). The signal that is recorded is the current through the device. This technique is very well suited for probing the local response of a nanotube device to gate voltage. Metallic nanotubes are found to be insensitive to local gating (Bachtold et al 2000) (as they are to global gating), although local defects in metallic nanotubes may be sensitive to the local gate (Bockrath et al 2001). Freitag et al (2001) used this technique to investigate NT-FETs. Their results showed that nanotubes were much more sensitive to gating near the positive contact than in the centre of the nanotube, suggesting the presence of a Schottky barrier at the contacts. Similar results were presented by Radosavljević et al (2002) showing the presence of contact barriers in n-type NT-FETs.

The second type of information that can be gained from SGM measurements shows the influence of disorder on the electronic structure of a nanotube. Measurements performed by Tans and Dekker (2000) and Bachtold et al (2000) show that the response of the nanotube
Figure 11. (a) Scanned gate microscopy (SGM): a gate voltage $V_{\text{tip}}$ is applied to the AFM tip. The current through the device at a given bias voltage $V_{\text{bias}}$ is recorded as a function of the tip position. The back-gate is used with a constant voltage to put the device into a certain region of its $I-V_g$ characteristic. (b) Alternating-current electrostatic force microscopy (AC-EFM): while the tube is biased with $V_{\text{bias}}$ at a frequency $\omega$ and gated with a fixed $V_g$ the tip is used to probe the local electric field around the tube. Using a lock-in technique the tip amplitude at $\omega$, proportional to the local AC potential of the nanotube, is recorded as a function of the tip position.

3.2.2. Electrostatic force microscopy. Electrostatic force microscopy (EFM) images the electrostatic potential around conducting materials. Figure 11(b) shows a schematic of the setup to an applied positive gate voltage is not uniform along the nanotube, suggesting potential variations along the device (Tans and Dekker 2000). Freitag et al (2002) were able to estimate the magnitude of the potential variations as 20–50 meV for ‘strong’ defects. Weaker defects could also be observed using a large positive voltage applied to the tip (Freitag et al 2002, Kalinin et al 2002) (a technique termed scanning impedance microscopy or ‘SIM’). Bockrath et al (2001) were able to identify defects in metallic nanotubes from their reaction to SGM.

Figure 12(a) shows an AFM image of a CVD-grown nanotube of diameter 2.0 nm and length 5 $\mu$m between electrodes. Figure 12(b) shows a scanned gate microscopy image of this nanotube. A series of spots are seen along the nanotube, corresponding to areas which are sensitive to the tip gate. The maximum variation in resistance when the tip is over the nanotube is approximately 500 $\Omega$, corresponding to changes in transmission probability of less than 10%. These changes are significantly smaller than the many M$\Omega$ resistance variations seen by Bachtold et al in semiconducting nanotubes grown by laser ablation, indicating smaller disorder in these CVD-grown nanotubes (Fuhrer et al 2001).
used for AC-EFM, a variation of this technique. While the setup is similar to SGM, the quantity measured in this case is the force acting on the AFM tip due to the electric fields around the devices under investigation. By showing the details of the potential drop along a biased nanotube this technique provides information about the character of the nanotube conductance (Bachtold et al 2000). Using EFM techniques to image various laser-ablation-synthesized nanotube devices at room temperature Bachtold et al found no voltage drop along a micrometre-length metallic nanotube except for the contact areas. For semiconducting nanotubes, however, in addition to the voltage drops at the contacts, large (tens of MΩ) resistive barriers were also seen.

Figure 12(c) shows a line trace of the voltage along the CVD-grown semiconducting nanotube shown in figure 12(a) as determined by EFM. The voltage drop along the nanotube is roughly linear, with a magnitude of 9.2 kΩ μm⁻¹, indicating a mean-free path of 700 nm at \( V_g = 0 \) V (gate-voltage dependence of the conductivity was not studied). This observation is reasonably consistent with the mean-free path calculated above and in Dürkop et al (2004), and those of Javey et al (2003) who observed ballistic conduction in 300 nm Ohmically contacted devices, but diffusive conduction in 3 μm devices.

4. Charge detection with carbon nanotubes

The high mobility of carbon nanotubes suggests high sensitivity in applications where charge detection is required, for example in a memory cell, in which the charge on a floating gate is detected by a transistor, or a chemical sensor, in which chemisorption of a target species
produces a charge detected by a transistor. Charge detection with an individual NT-FET has been explored by constructing a crude floating-gate memory device where charges are stored in traps in the SiO₂ gate dielectric. Charge is reversibly injected and removed from the dielectric by applying a moderate voltage (≤10 V) across the dielectric between nanotube and substrate. It is found that in this type of NT-FET memory discrete charge states corresponding to differences of a single electronic charge are observed, and can be written, read and erased at temperatures up to 100 K, with changes in current of more than 50 nA. A device consisting of a single NT-FET containing a single defect is discussed in section 4.4; in this device a regular series of discrete charge states allows a quantitative analysis of the capacitances of the nanotube and gate to the charge trap.

4.1. Devices and fabrication

The carbon nanotubes discussed in this section were synthesized via chemical vapour deposition following the methods outlined in section 2.1; more details may be found in Fuhrer et al (2002). Source and drain Cr/Au electrical contacts to the nanotube were fabricated via electron-beam lithography; the conducting Si substrate acts as a gate electrode. Room temperature and low temperature electrical measurements were carried out with the samples mounted on a cryostat in flowing helium gas.

4.2. Hysteresis and memory

Figure 13(a) shows an atomic force microscope topograph of the NT-FET described in this section; the nanotube has a length \( L = 4.8 \, \mu \text{m} \) and a diameter \( d = 2.7 \, \text{nm} \) as determined from the height profile of this image. Figure 13(b) shows the drain current \( I \) as a function of gate voltage \( V_g \) with \( V_{sd} = 500 \, \text{mV} \) applied to the source electrode. As the gate voltage is swept back and forth between +10 and −10 V, a large hysteresis is evident in the \( I-V_g \) curves; the threshold gate voltage \( V_{th} \) at which the nanotube begins to conduct is shifted by more than 6 V. Figure 13(c) demonstrates that this hysteresis may be used as the basis of a stable memory at room temperature. Here the state of the device is read at \( V_{sd} = 500 \, \text{mV}, \ V_g = -1 \, \text{V} \), and written and erased with pulses of the gate voltage to ±8 V. A current of >1 \( \mu \text{A} \) may be switched. After an initial slow decay (≈50 s) the current remains constant; the hold time of the memory exceeds 5000 s.

From the linear portion of the \( G(V_g) \) curve we determine the field-effect mobility (equation (3)). Using the gate capacitance determined directly from low-temperature Coulomb blockade measurements, \( C_g = 54 \, \text{aF} \) and the slope \( dG/dV_g = -2.2 \, \mu \text{S} \, \text{V}^{-1} \) determined from the linear portion of the \( I-V_g \) curves in figure 13(b), we calculate the field-effect mobility of holes to be ≈9000 cm² V⁻¹ s⁻¹. This mobility is significantly smaller than that found above in section 3.1.2, and likely indicates a significant role of the contacts in the resistance of this shorter device. Still, this compares very favourably with conventional FETs.

The shift in the threshold voltage of the NT-FET indicates a reconfiguration of the charge environment of the transistor under an applied gate voltage. Such a reconfiguration of charge can happen in one of two ways: either charges present in the system move in the gate field, or new charges are injected into the system from the conducting electrodes or channel. These two mechanism result in an opposite ‘sign’ of the hysteresis loop, i.e. positive gate voltage increases the threshold voltage in the case of charge injection, but decreases the threshold voltage in the case of mobile charges. This indicates that charge injection is responsible for the hysteresis in these nanotube devices. Because of the device geometry, it is posited that the source of injected charge is the nanotube itself, rather than the gate or source/drain contacts,
Figure 13. NT-FET as a memory device. (a) AFM image of the device. (b) $I(V_g)$ behaviour at room temperature and applied bias of 500 mV; a pronounced hysteresis is evident. (c) By applying $V_g$ pulses the device can be switched from between the two branches of the hysteresis loop resulting in a change in device current. The current is stable for minutes after writing (see also Fuhrer et al 2002).

because the electric field is much higher at the nanotube than at the electrodes or Si/SiO$_2$ interface. An estimate of the electric field at the nanotube/SiO$_2$ interface of 0.3 V nm$^{-1}$ at $V_g = 10$ V indicates that charge injection into the SiO$_2$ would reasonably be expected; the electric field is comparable to the breakdown electric field for SiO$_2$ of $\sim 0.25$ V nm$^{-1}$. (Note that ‘breakdown field’ is typically quoted as breakdown voltage divided by dielectric thickness and does not include the dielectric constant $\kappa$, so a breakdown field of 1.0 V nm$^{-1}$ for SiO$_2$ with $\kappa \approx 4$ is equivalent to an internal electric field of 0.25 V nm$^{-1}$.)

Note that the nature of the charges in SiO$_2$ is not elucidated by these experiments. The charges may be related to the so-called anomalous positive charge observed in MOS capacitors, whereby stressing of the capacitor was observed to cause formation of positive charge centres which could subsequently be filled or emptied by electron injection (Fischetti et al 1982a, 1982b). Other researchers have also observed hysteresis in NT-FETs. Radosavljević et al (2002) observed similar effects as those reported here, with similar interpretation, i.e. that the hysteresis results from electrons injected into SiO$_2$ from the nanotube. Bradley et al (2003) intentionally caused hysteresis by coating NT-FETs with polymer containing mobile ions. Kim et al (2003) found that the hysteresis could be affected by high-temperature annealing or coating the nanotube with a polymer such as poly(methyl methacrylate), and proposed that the hysteresis was due to charge trapping by surface-bound water molecules near the nanotube. Water and water-related species are known to affect the anomalous positive charge, but different groups have reported conflicting results on the exact role of water in charge trapping in SiO$_2$ (Fischetti et al 1982b). A similar non-volatile memory based on charge storage in a thick SiO$_2$ gate dielectric was reported in organic thin-film transistors; in this parallel-plate geometry
Figure 14. Single-electron memory operation. (a) Current–gate voltage $I_d(V_g)$ curves of the device shown in figure 12 at a temperature of 20 K and bias voltage 500 mV. Two branches of the $I_d(V_g)$ curves are evident, with discrete switching between the branches occurring in certain intervals of $V_g$. The branches correspond to two charge states of the dielectric differing by one electronic charge. (b) Memory operation of the device. The current level is read at $V_g = -2.25$ V, and gate voltage pulses to $V_g = -1.5$ V (−3.0 V) are used to write (erase) the memory (see also Fuhrer et al 2002).

$V_g = 100$ V ($E \approx 0.03$ V nm$^{-1}$) was sufficient to polarize the SiO$_2$ dielectric (Katz et al 2002). The same sign hysteresis was observed as in our case, also suggesting injection of charge into the dielectric.

4.3. Single-electron memory

At lower temperatures, discrete behaviour in the hysteresis is seen; the current as a function of gate voltage is observed to fall on one of a set of discrete curves. Switching between the curves is stochastic, but appears within certain ranges of $V_g$. The discrete $I–V_g$ curves are identified with discrete charge states of the dielectric differing by a single electronic charge. Figure 14(a) illustrates the $I–V_g$ characteristics of the device shown in figure 13 at 20 K. Here $V_g$ is swept back and forth over a particular region (between −1.3 and −3 V) to emphasize one large discrete hysteresis loop; other hysteresis loops could be observed with much smaller widths by scanning different ranges of gate voltage. This behaviour is typical of NT-FETs, most often only one or two large hysteresis loops may be observed, with switching occurring near the threshold gate voltage.

A single electronic charge added to the gate electrode would shift the threshold of the nanotube transistor by an amount $\Delta V_{th} = e/C_g \approx 3$ mV. This is much smaller than the observed threshold shift in figure 14(a), approximately 50–150 mV. The effect of a single trapped charge near the nanotube is exaggerated for several reasons. First, the length of nanotube affected by the trapped charge is much shorter than the entire nanotube, and thus the threshold shift should be greater. Second, the effect of trapped charge is exaggerated by the device being operated at high bias (near pinchoff) which would lead to greater sensitivity to
charges near the source electrode. The fact that the spacing in gate voltage between the two curves in figure 14(a) varies significantly, becoming larger nearer the threshold (more positive gate voltage) argues that pinchoff is important; indeed, at the upper end of the gate voltage range the source–drain bias (500 mV) is comparable to the difference between the gate voltage (−1.5 to −3.0 V) and the threshold voltage (∼−1.1 V), and pinchoff should be significant.

Figure 14(b) demonstrates memory operation of the device at 20 K. The drain current is measured at a gate voltage of −2.25 V and a source–drain bias of 500 mV. The drain current is switched between high and low states by application of a gate voltage pulse of ±0.75 V (relative to the −2.25 V baseline). The difference in current between high and low states is approximately 60 nA. The memory is stable for at least 100 s.

4.4. Detecting single electrons with a single nanotube defect

Charge sensitivity has been explored more closely in a second NT-FET memory (Kim et al 2002a). In this device, a single defect was detected using scanned-gate microscopy; this defect was much more sensitive to gate voltage, and the sensitivity of this defect allowed a regular series of charge states in traps in the dielectric to be detected. The regular shifts in gate voltage and write voltage allow an estimation of the trap charging energy to be made.

Figure 15(a) shows the geometry of this nanotube device. A 3 µm long semiconducting nanotube is contacted by the source and drain electrodes separated by 1.85 µm. The nanotube has a diameter 3 nm as determined from the height profile of this AFM image. SGM (see section 3.2.1) was used to investigate the local electric field sensitivity of the nanotube. Figure 15(b) shows a SGM image of the device with 4 V applied to the AFM tip. The nanotube appears darker (larger resistance change) than the surrounding substrate, indicating that the entire nanotube is semiconducting, i.e. gate-voltage sensitive. One particularly dark region of the nanotube is evident, located about 1.3 µm from the left contact. A line trace of the SGM image along the nanotube (figure 15(d)) shows a sharp peak at this spot with a FWHM of ∼30 nm and a height approximately five times greater than elsewhere along the nanotube. The spot has no corresponding topographic feature in the AFM image. This spot is interpreted as due to a single atomic-scale defect in the nanotube; similar spots of large gate-voltage dependence have been observed in metallic nanotubes (Bockrath et al 2001) and were similarly interpreted.

Figure 15(c) presents the drain current (I) versus gate voltage (V_g) characteristics of the device measured at 5 K. As in section 4.3, discrete I(V_g) curves are seen, with hysteretic switching between the curves occurring at specific ranges of gate voltage. The discrete curves correspond to the charge state of the dielectric differing by a single electronic charge. Since the gate voltage dependence of the conductance of the device is dominated by a single defect as observed in figures 15(a) and (b), it is assumed that the trap site is strongly coupled to this defect, and changes in the potential of this defect due to charges at the nearby trap sites result in the various curves seen in figure 15(e). Because the device behaviour is dominated by a single charge trap and a single defect region acting as the electrometer, a regular series of I(V_g) curves is seen corresponding to differences in charge at the trap of one electron.

Figure 15(c) shows a probable schematic of this device. The charge trap (either at the SiO_2 surface or in the SiO_2 bulk) is coupled capacitively to the defect in the nanotube channel through a capacitance C_{TC}, and to the gate through a capacitance C_{GT}. There is also an additional capacitance between the gate and channel, C_{FC}. In addition, a non-linear resistor must be present between the nanotube channel and the charge trap to allow charging and discharging of the defect. Such a device structure has been considered in detail in Yano et al (1999), where it was shown that this structure forms the basis of a single electron memory.
Figure 15. Single atomic-scale defect in a semiconducting nanotube. (a) AFM image of the device; the large white areas are Cr/Au contacts, the thin white line is the nanotube, with diameter 3.0 nm and length 1.85 µm between electrodes. (b) Scanned-gate microscopy (SGM—see figure 11) image of the device. (c) Schematic of the device (see text). (d) SGM shows the resistance change as a function of position along the nanotube; i.e. a line trace through image (b). The resistance change at the point marked with the arrow in (a), (b) and (d) is ~5 times larger than in the rest of the nanotube. The point marked with the arrow is identified as an atomic-scale defect in the nanotube. (e) Drain current $I$ as a function of gate voltage $V_g$ for the nanotube device at a temperature of 5 K. Five discrete $I(V_g)$ curves are seen, with no intermediate values of $I$. Four hysteresis loops are evident in the figure. Several sweeps back and forth over $V_g$ ranges 5–7.8, 3.5–6.8, 2.5–6 and 1–5 V are shown with the respective ranges highlighted in green, red, blue and pink, respectively; these ranges were chosen to sweep out the four hysteresis loops. $\Delta V_{th}$ indicates the shift in $V_g$ between the $I(V_g)$ curves; $\Delta V_{th} \approx 200$ mV. $\Delta V_w$ indicates the shift in $V_g$ between the jumps between $I(V_g)$ curves; $\Delta V_w \approx 1$ V. (f) $I$ versus $V_g$ at a temperature of 200 K plotted as scattered points. Random switching between curves was observed as $V_G$ was swept back and forth between 3 and 8 V.
The $I(V_g)$ characteristics of a transistor coupled to a charge trap in this way will show discrete curves, separated in gate voltage by an amount

$$\Delta V_{th} = \frac{eC_{TC}}{(C_{TC} + C_{GC})C_{GC}}$$

(5)

where $e$ is the electronic charge and $C_{GC}$ is the total gate capacitance given by

$$C_{GC} = C_{FC} + \frac{C_{TC}C_{GT}}{C_{TC} + C_{GT}}$$

(6)

Periodic switching events will occur between curves with a spacing

$$\Delta V_n = \frac{e}{C_{GT}}$$

(7)

and the hysteresis width $\Delta V_h$ (the distance in gate voltage between the switching events between two charge states of the trap when increasing gate voltage and decreasing gate voltage) is determined by the nonlinear characteristics of the resistor $R_{TC}$ between channel and charge trap (Yano et al 1999). If we make the simplifying assumption that $R_{TC} = 0$ for voltages less than some critical voltage $V_c$, and finite for larger voltages, then $\Delta V_h$ is given by

$$\Delta V_h = \frac{2V_c(C_{TC} + C_{GT})}{C_{GT}} - \Delta V_w.$$ 

(8)

For this device, the threshold shift is $\Delta V_{th} \approx 200$ mV, the write voltage periodicity $\Delta V_w \approx 1.0$ V, and the hysteresis width $\Delta V_h$ is roughly 1.8 V (see figure 15(e)). These three voltages alone are insufficient to determine the four unknowns in equations (5)–(8): $(C_{TC}, C_{TC}, C_{GT}$ and $V_c)$. One of the capacitances, $C_{FC}$, is estimated from an estimate of the size of the defect region in the channel. From the SGM line trace (figure 15(b)), the defect region is approximately 30 nm in extent (or perhaps smaller, this distance is comparable to the resolution of the image). This distance can be interpreted as the screening length in the nanotube at the defect. Estimating the gate capacitance per length of the nanotube transistor as approximately 100 fF cm$^{-1}$ obtained from Coulomb-blockade measurements of similar devices, this gives a capacitive coupling $C_{TC} \approx 0.30$ aF between the defect region and gate. Solving for the other capacitances gives $C_{TC} \approx 0.14$ aF and $C_{GT} \approx 0.16$ aF. The total capacitance of the charge trap $C_{TT} = C_{TC} + C_{GT} = 0.30$ aF. From this value a rough estimate of the size of the charge trap can be made. The self-capacitance of a sphere of radius $r$ is given by $C_{self} = 4\pi\varepsilon_0\varepsilon_r r$, where $\varepsilon_0$ is the permittivity of free space and $\varepsilon$ the dielectric constant of the medium in which the sphere is embedded, $\sim 4$ for SiO$_2$. Setting $C_{TT} = C_{self}$ gives $r \approx 7$ Å, reasonable for a vacancy site or complex of dangling bonds in the SiO$_2$ dielectric.

The total capacitance of the charge trap gives a Coulomb charging energy $E_c = \frac{e^2}{C_{TT}} \approx 530$ meV. The Coulomb blockade should be effective up to a temperature $k_B T$ of the order of $0.1 E_c$, or approximately 620 K. Figure 15(f) shows the $I(V_g)$ characteristics for the device at a temperature of 200 K in the form of a scatter plot. The $I(V_g)$ data are observed to fall on a set of discrete curves, but random telegraph switching between the curves was observed at all gate voltages. The data in figure 15(f) are taken at a bandwidth of 10 Hz; from the fact that the curves corresponding to single electron differences at the charge trap are well resolved, (signal to noise ratio $> 3$) we estimate that the nanotube defect electrometer has a sensitivity better than $0.1$ e Hz$^{-1/2}$ at 200 K. The observation of discrete $I(V_g)$ curves indicates that the Coulomb blockade of the charge trap is indeed still effective at 200 K, but the random telegraph switching indicates that the nonlinear resistance $R_{TC}$ between channel and trap has become leaky at elevated temperatures, allowing the charge on the trap to fluctuate. Using equation (8) to find the critical voltage for $R_{TC}$, we find $V_c = 430$ mV, which is still much greater than the thermal energy at 200 K. However, we notice that the hysteresis widths become narrower as the temperature increases, probably due to a finite activated component of $R_{TC}$ at low voltages.
5. Conclusions

Recent advances have allowed the fabrication of long, Ohmically contacted semiconducting carbon nanotube transistors. These devices have been used to measure the intrinsic carrier mobility in the nanotube channel; the mobility exceeds 100,000 cm² V⁻¹ s⁻¹ at room temperature and low doping, higher than any other known semiconductor at room temperature. Scanned-probe experiments on semiconducting nanotubes help to elucidate the disorder present in various types of nanotubes. Large conduction barriers are often observed in solution-deposited semiconducting nanotubes; however in CVD nanotubes grown on-chip, the degree of disorder is apparently lower. Electrostatic force microscopy has been used to directly verify long mean free paths in CVD-grown semiconducting nanotubes, consistent with transport experiments. The results indicate that semiconducting nanotubes should be an excellent material for a number of semiconductor applications, especially in high-speed transistors where mobility is crucial.

Semiconducting nanotube field-effect transistors have been used to inject and detect electrons trapped in the SiO₂ gate dielectric. Single-electron sensitivity is seen, and unusual devices enable single-electron detection up to 200 K, with a sensitivity of better than 0.1 e Hz⁻¹/². The results suggest that semiconducting nanotubes may find applications as exquisite sensors of e.g. chemical or biochemical species, in which a chemical signal is translated into charge. Single molecule detection appears feasible with such a device.

Acknowledgments

This research was supported by ARDA and the Office of Naval Research through grant N000140110995, the Director of Central Intelligence Postdoctoral Research Fellowship Program, and the National Science Foundation through grant DMR-0102950. The authors would like to thank the coauthors of the original works reviewed here: Todd Brintlinger, Raymond Chan, Yung-Fu Chen, Enrique Cobas, Manu Forero, Stephanie Getty, Paul McEuen and Alex Zettl. The authors also thank Gary Pennington and Neil Goldsman for useful conversations.

References

http://xxx.lanl.gov/abs/cond-mat/0309641
Radosavljević M, Freitag M, Thadani K V and Johnson A T 2002 Nano Lett. 2 761–4
Seeger K 2002 Semiconductor Physics: an Introduction (Berlin: Springer)
Wallace P R 1947 Phys. Rev. 71 622–34
Yao Z, Kane C L and Dekker C 2000 Phys. Rev. Lett. 84 2941–4