NEHRU COLLEGE OF ENGINEERING AND RESEARCH CENTRE
(NAAC Accredited)
(Approved by AICTE, Affiliated to APJ Abdul Kalam Technological University, Kerala)

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

COURSE MATERIALS

EC 402:NANOELECTRONICS

VISION OF THE INSTITUTION

To mould true citizens who are millennium leaders and catalysts of change through excellence in education.

MISSION OF THE INSTITUTION

NCERC is committed to transform itself into a center of excellence in Learning and Research in Engineering and Frontier Technology and to impart quality education to mould technically competent citizens with moral integrity, social commitment and ethical values.

We intend to facilitate our students to assimilate the latest technological know-how and to imbibe discipline, culture and spiritually, and to mould them in to technological giants, dedicated research scientists and intellectual leaders of the country who can spread the beams of light and happiness among the poor and the underprivileged.

ABOUT DEPARTMENT

- Established in: 2002
- Course offered : B.Tech in Electronics and Communication Engineering

  M.Tech in VLSI
- Approved by AICTE New Delhi and Accredited by NAAC
- Affiliated to the University of Dr. A P J Abdul Kalam Technological University.

DEPARTMENT VISION

Providing Universal Communicative Electronics Engineers with corporate and social relevance towards sustainable developments through quality education.

DEPARTMENT MISSION

1) Imparting Quality education by providing excellent teaching, learning environment.

2) Transforming and adopting students in this knowledgeable era, where the electronic gadgets (things) are getting obsolete in short span.

3) To initiate multi-disciplinary activities to students at earliest and apply in their respective fields of interest later.

4) Promoting leading edge Research & Development through collaboration with academia & industry.

PROGRAMME EDUCATIONAL OBJECTIVES

PEOI. To prepare students to excel in postgraduate programmes or to succeed in industry / technical profession through global, rigorous education and prepare the students to practice and innovate recent fields in the specified program/ industry environment.

PEO2. To provide students with a solid foundation in mathematical, Scientific and engineering fundamentals required to solve engineering problems and to have strong practical knowledge required to design and test the system.

PEO3. To train students with good scientific and engineering breadth so as to comprehend, analyze, design, and create novel products and solutions for the real life problems.

PEO4. To provide student with an academic environment aware of excellence, effective communication skills, leadership, multidisciplinary approach, written ethical codes and the life-long learning needed for a successful professional career.

PROGRAM OUTCOMES (POS)

Engineering Graduates will be able to:

1. **Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM SPECIFIC OUTCOMES (PSO)

PSO1: Ability to Formulate and Simulate Innovative Ideas to provide software solutions for Real-time Problems and to investigate for its future scope.

PSO2: Ability to learn and apply various methodologies for facilitating development of high quality System Software Tools and Efficient Web Design Models with a focus on performance optimization.

PSO3: Ability to inculcate the Knowledge for developing Codes and integrating hardware/software products in the domains of Big Data Analytics, Web Applications and Mobile Apps to create innovative career path and for the socially relevant issues.

COURSE OUTCOMES

EC 402

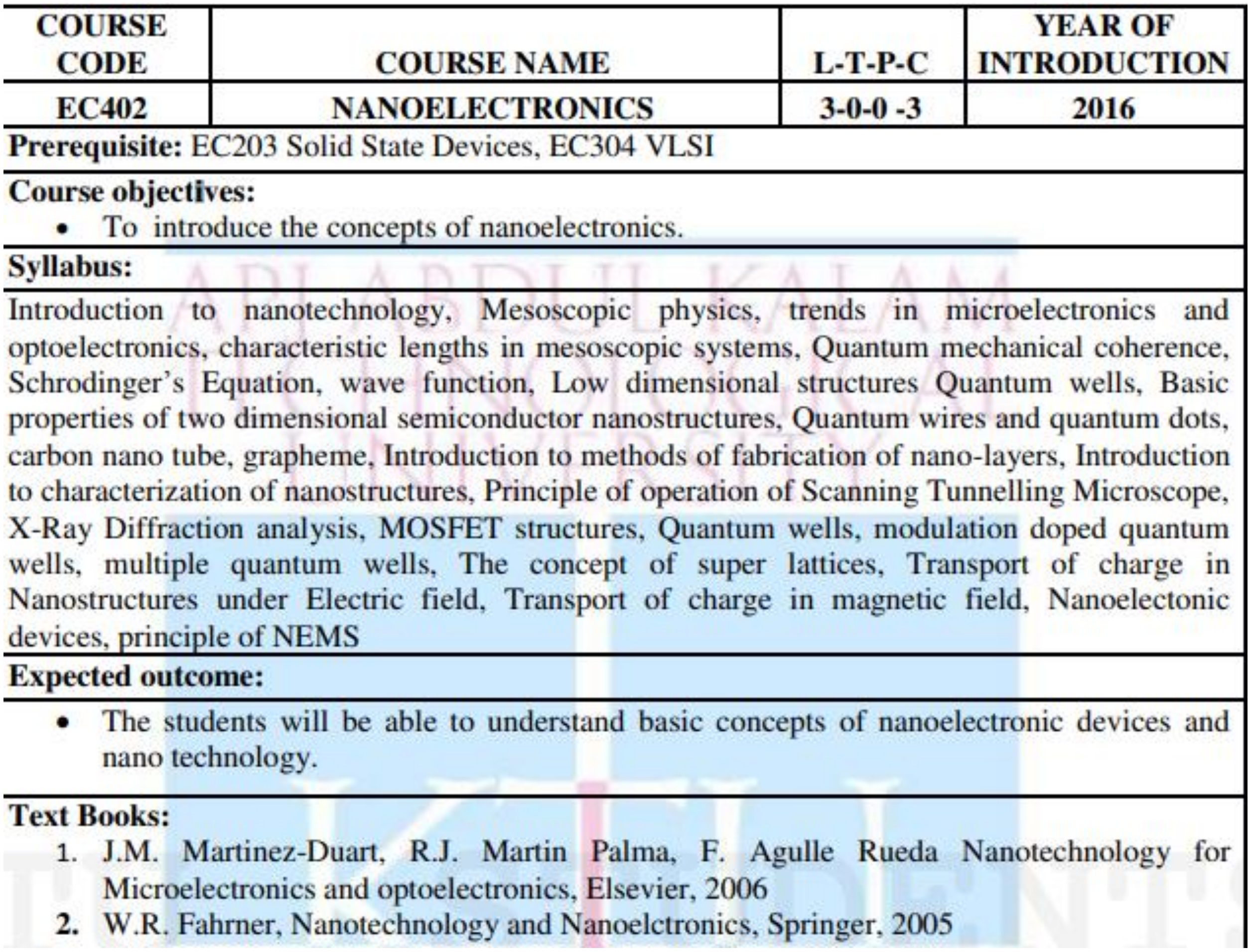
SUBJECT CODE: EC 308	
COURSE OUTCOMES	
C402.1	The students will be able to understand basic concepts of nanoelectronic materials
C402.2	The students will be able to understand basic concepts of fabrication methods
C402.3	To student get the knowledge about microscopic methods
C402.4	To student get the knowledge about quantum wells modulation dopped quantum wells
C402.5	The students will be able to understand about quantum transport nano structures
C402.6	The students will be able to understand basic concepts hetrostructure

MAPPING OF COURSE OUTCOMES WITH PROGRAM OUTCOMES

CO'S	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
C402.1	3	2	3		3		I			1		3
C402.2	3	2	3				I			1		3
C402.3	3		3		3		1			2		3
C402.4		2	3							2		
C402.5	3		2		2		2			2		2
C402.6	3	2	2		2		3			2		2
C402	3	2	1		2		2			2		3

CO'S	PSO1	PSO2	PSO3
C308.1	3	2	3
C308.2	3	2	3
C308.3	2	2	2
C308.4	2		2
C308.5			2
C308.6			
C308	3	3	2

SYLLABUS



COURSE CODE	COURSE NAME	L-T-P-C	YEAR OF INTRODUCTION
EC402	NANOELECTRONICS	3-0-0 -3	2016

Prerequisite: EC203 Solid State Devices, EC304 VLSI

Course objectives:

- To introduce the concepts of nanoelectronics.

Syllabus:

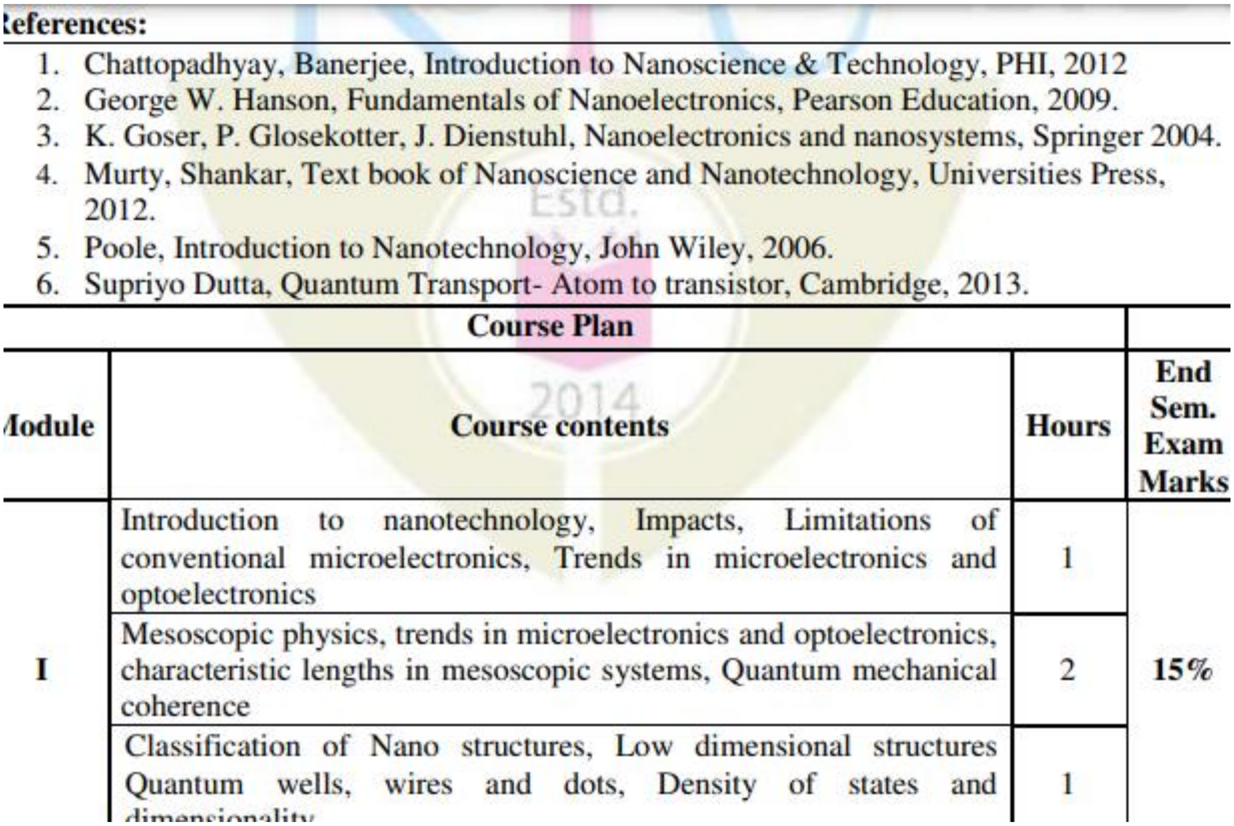
Introduction to nanotechnology, Mesoscopic physics, trends in microelectronics and optoelectronics, characteristic lengths in mesoscopic systems, Quantum mechanical coherence, Schrodinger's Equation, wave function, Low dimensional structures Quantum wells, Basic properties of two dimensional semiconductor nanostructures, Quantum wires and quantum dots, carbon nano tube, grapheme, Introduction to methods of fabrication of nano-layers, Introduction to characterization of nanostructures, Principle of operation of Scanning Tunnelling Microscope, X-Ray Diffraction analysis, MOSFET structures, Quantum wells, modulation doped quantum wells, multiple quantum wells, The concept of super lattices, Transport of charge in Nanostructures under Electric field, Transport of charge in magnetic field, Nanoelectonic devices, principle of NEMS

Expected outcome:

- The students will be able to understand basic concepts of nanoelectronic devices and nano technology.

Text Books:

1. J.M. Martinez-Duart, R.J. Martin Palma, F. Agulle Rueda Nanotechnology for Microelectronics and optoelectronics, Elsevier, 2006
2. W.R. Fahrner, Nanotechnology and Nanoelctronics, Springer, 2005

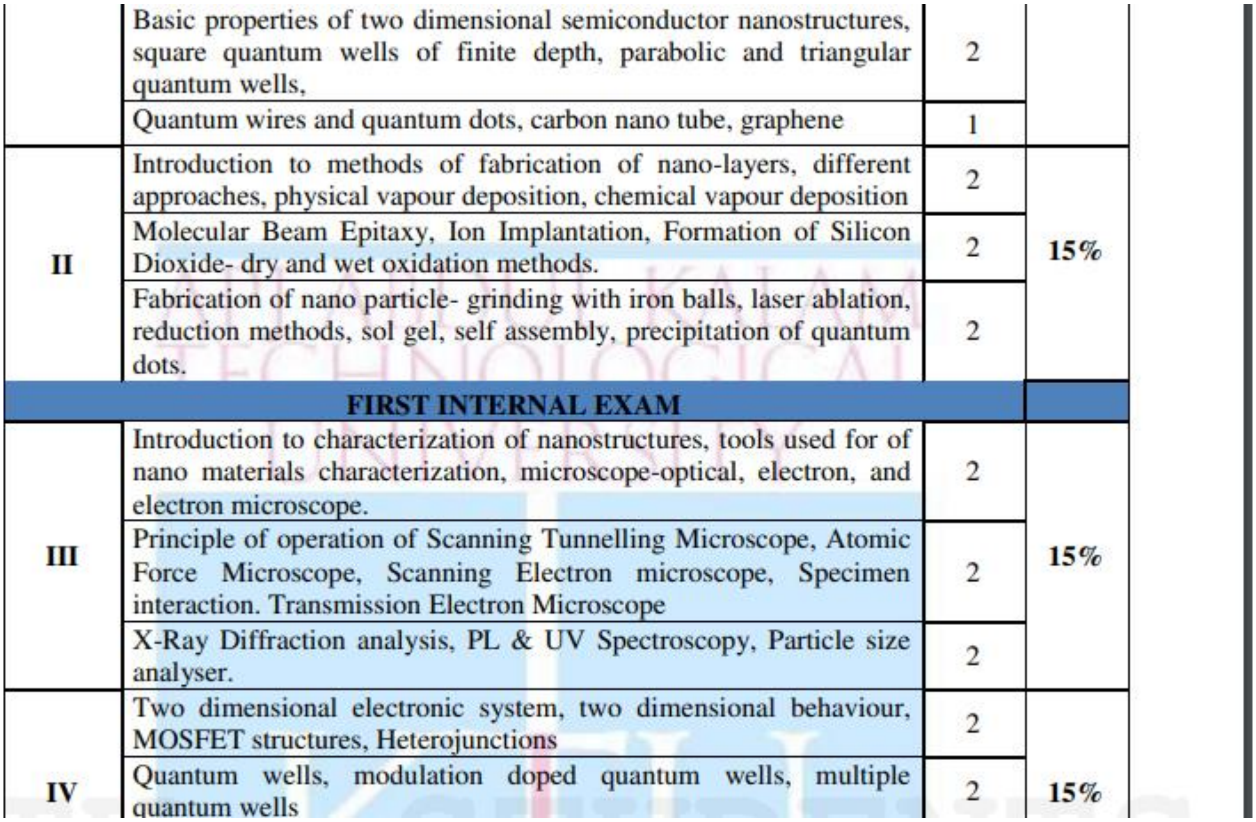


eferences:

1. Chattopadhyay, Banerjee, Introduction to Nanoscience & Technology, PHI, 2012
2. George W. Hanson, Fundamentals of Nanoelectronics, Pearson Education, 2009.
3. K. Goser, P. Glosekotter, J. Dienstuhl, Nanoelectronics and nanosystems, Springer 2004.
4. Murty, Shankar, Text book of Nanoscience and Nanotechnology, Universities Press, 2012.
5. Poole, Introduction to Nanotechnology, John Wiley, 2006.
6. Supriyo Dutta, Quantum Transport- Atom to transistor, Cambridge, 2013.

Course Plan

odule	Course contents	Hours	End Sem. Exam Marks
I	Introduction to nanotechnology, Impacts, Limitations of conventional microelectronics, Trends in microelectronics and optoelectronics	1	15%
	Mesoscopic physics, trends in microelectronics and optoelectronics, characteristic lengths in mesoscopic systems, Quantum mechanical coherence	2	
	Classification of Nano structures, Low dimensional structures Quantum wells, wires and dots, Density of states and dimensionality	1	



	Basic properties of two dimensional semiconductor nanostructures, square quantum wells of finite depth, parabolic and triangular quantum wells,	2	
	Quantum wires and quantum dots, carbon nano tube, graphene	1	
II	Introduction to methods of fabrication of nano-layers, different approaches, physical vapour deposition, chemical vapour deposition	2	15%
	Molecular Beam Epitaxy, Ion Implantation, Formation of Silicon Dioxide- dry and wet oxidation methods.	2	
	Fabrication of nano particle- grinding with iron balls, laser ablation, reduction methods, sol gel, self assembly, precipitation of quantum dots.	2	
FIRST INTERNAL EXAM			
III	Introduction to characterization of nanostructures, tools used for of nano materials characterization, microscope-optical, electron, and electron microscope.	2	15%
	Principle of operation of Scanning Tunnelling Microscope, Atomic Force Microscope, Scanning Electron microscope, Specimen interaction. Transmission Electron Microscope	2	
	X-Ray Diffraction analysis, PL & UV Spectroscopy, Particle size analyser.	2	
IV	Two dimensional electronic system, two dimensional behaviour, MOSFET structures, Heterojunctions	2	15%
	Quantum wells, modulation doped quantum wells, multiple quantum wells	2	

IV	Two dimensional electronic system, two dimensional behaviour, MOSFET structures, Heterojunctions	2	**15%**
	Quantum wells, modulation doped quantum wells, multiple quantum wells	2	
	The concept of super lattices Kronig - Penney model of super lattice.	2	
V	Transport of charge in Nanostructures under Electric field - parallel transport, hot electrons, perpendicular transport.	2	**20%**
	Quantum transport in nanostructures, Coulomb blockade	2	
	Transport of charge in magnetic field - Effect of magnetic field on a crystal. Aharonov-Bohm effect, the Shubnikov-de Hass effect, the quantum Hall effect.	3	
VI	Nanoelectonic devices- MODFETS, heterojunction bipolar transistors	1	**20%**
	Resonant tunnel effect, RTD, RTT, Hot electron transistors	2	
	Coulomb blockade effect and single electron transistor, CNT transistors	2	
	Heterostructure semiconductor laser	1	
	Quantum well laser, quantum dot LED, quantum dot laser	2	
	Quantum well optical modulator, quantum well sub band photo detectors, principle of NEMS.	2	
	END SEMESTER EXAM		

Question Paper Pattern

The question paper shall consist of three parts. Part A covers modules I and II, Part B covers modules III and IV, and Part C covers modules V and VI. Each part has three questions uniformly covering the two modules and each question can have maximum four subdivisions. In each part, any two questions are to be answered. Mark patterns are as per the syllabus with 70% for theory and 30% for logical/numerical problems, derivation and proof.

QUESTION BANK

MODULE I				
Q:NO:	**QUESTIONS**	**CO**	**KL**	**PAGE NO:**
1	List out the disadvantages of conventional microelectronics	CO1	K2	2
2	Explain the basic properties of a two dimensional semiconductor nano structure	CO1	K3	4
3	List out the limitations of conventional microelectronics	CO1	K3	3
4	Explain in detail electronic properties of grapheme	CO1	K3	37
5	Explain the different characteristic lengths in a mesoscopic system?	CO1	K2	7
6	Differentiate between parabolic and triangular quantum well. Explain the features	CO1	K2	26
7	Explain about square quantum well of finite size	CO1	K2	27
8	Explain in detail about carbon nano tubes.	CO1	K2	35
9	Explain about low dimensional structures	CO1	K2	32
10	Analyze parabolic quantum well	CO1	K2	28
MODULE 2				
1	Explain Ion Implantation technique	CO2	K1	38

2	Compare PVD & CVD	CO2	K2	39
3	Analyze Precipitation of quantum dots	CO2	K2	42
4	Analyze molecular beam epitaxy	CO2	K2	43
5	Analyze the different approaches for the formation of nanolayers	CO2	K2	41
6	Discuss about molecular beam epitaxy	CO2	K2	42
7	Explain sol-gel process for fabrication of nano-particles	CO2	K2	43
8	Draw the neat diagram laser ablation method and explain its detail	CO2	K2	47
9	Explain any two physical vapour deposition methods in detail	CO2	K2	48
10	Discuss the properties of nanoparticles	CO2	K2	47
11	Analyze chemical vapour deposition	CO2	K2	49
MODULE III				
1	Differentiate optical microscope and Electron microscope	CO3	K1	56
2	Explain the working principle of Atomic Force Microscope.	CO3	K3	58
3	Draw the neat diagram of photoluminescence spectroscopy and explain its working.	CO3	K2	60
4	Explain about transmission electron microscopy	CO3	K3	59
5	Draw the neat diagram of photoluminescence spectroscopy and explain its working.	CO3	K2	61
6	Explain about transmission electron microscopy	CO3	K2	63
7	Explain with neat diagram different types of specimen interactions taking place in a sample during SEM.	CO3	K2	64
8	Explain the working of XRD analyzer and how it can be used to analyze a crystal.	CO3	K3	66
9	Explain the basic principles of scanning probe microscopy in detail	CO3	K2	69
10	Describe about electron microscopy. Explain SEM	CO3	K3	66

MODULE IV				
1	Differentiate MQW and super lattices	CO4	K3	74
2	Explain about Multiple quantum wells with neat diagrams.	CO4	K2	76
3	Discuss about Bloch theorem.	CO4	K2	78
4	Explain high electron mobility transistor	CO4	K2	75
5	Discuss about important characteristics of hetero structures	CO4	K2	73
6	Explain the working of a resonant tunneling hot electron transistor, with the aid of energy band representation	CO4	K2	77
7	Discuss about parallel transport.	CO4	K3	78
8	Illustrate about perpendicular transport	CO4	K3	79
9	Discuss about QHE and FQHE	CO4	K2	81
10	State and explain Landauer formula	CO4	K2	82
11	Derive quantized conductance with suitable schematic diagram	CO4	K1	83
12	Derive Landauer formula and Landauer Buttiker formula	CO4	K2	84
MODULE V				
1	Explain about resonant tunneling	CO5	K3	85
2	Discuss the important properties of nanotubes.	CO5	K2	86
3	List the benefits of using carbon nanotubes for AFM chips	CO5	K3	88
4	Discuss the principle of stimulated emission	CO5	K3	90
5	Write a short notes on about NEMS	CO5	K2	95

6	MODFETs are known as ‘high electron mobility transistor’. Justify with the help of energy band diagram.	CO5	K3	96
7	Discuss MOSFET structure	CO5	K3	76
8	Explian Kronig-Penney model	CO5	K2	86
9	Explain Brillouin zone.What do you mean by folding zones.	CO5	K2	96
MODULE VI				
1	Compare The Aharonov-Bohn effect and The Shubnikov-De Haas effect	CO6	K3	97
2	Describe about effect of magnetic field on a crystal.	CO6	K2	99
3	Explain electron-photon scattering mechanism in parallel transport.	CO6	K2	100
4	Explain the quantum transport in nanostructure	CO6	K2	101
5	Derive the equation of Aharonov-Bohm effect and explain in detail	CO6	K3	103
6	Discuss in detail about carbon nanotube transistors	CO6	K2	105
7	With neat diagram explain about quantum dot lasers	CO6	K2	106
9	Discuss about resonant tunnel effect .With neat diagram explain RTDs	CO6	K3	107

APPENDIX 1

CONTENT BEYOND THE SYLLABUS

S:NO;	TOPIC	PAGE NO:
1	OPPORTUNITIES AND RISK OF NANOTECHNOGIES	114

.

.

MODULE I
INTRODUCTION TO NANOTECHNOLOGY

1.1 DEFINITION

The design, characterization, production, and application of structures, devices, and systems by controlled manipulation of size and shape at the nanometer scale (atomic, molecular, and macromolecular scale) that produces structures, devices, and systems with at least one novel/superior characteristic or property.

1.2 THE SIGNIFICANCE OF THE NANOSCALE

A nanometer (nm) is one thousand millionth of a meter. For comparison, a red blood cell is approximately 7,000 nm wide and a water molecule is almost 0.3nm across.

People are interested in the nanoscale – which we define to be from 100nm down to the size of atoms (approximately 0.2nm) – because it is at this scale that the properties of materials can be very different from those at a larger scale. We define nanoscience as the study of phenomena and manipulation of materials at atomic, molecular and macromolecular scales, where properties differ significantly from those at a larger scale; and nanotechnologies as the design, characterisation, production and application of structures, devices and systems by controlling shape and size at the nanometer scale.

The properties of materials can be different at the nanoscale for two main reasons:

First, nanomaterials have a relatively larger surface area when compared to the same mass of material produced in a larger form. This can make materials more chemically reactive (in some cases materials that are inert in their larger form are reactive when produced in their nanoscale form), and affect their strength or electrical properties.

Second, quantum effects can begin to dominate the behavior of matter at the nanoscale - particularly at the lower end - affecting the optical, electrical and magnetic behavior of materials. Materials can be produced that are nanoscale in one dimension (for example, very thin surface coatings), in two dimensions (for example, nanowires and nanotubes) or in all three dimensions (for example, nanoparticles).

1.3 IMPACTS OF NANO TECHNOLOGY

The **impact of nanotechnology** extends from its medical, ethical, mental, legal and environmental applications, to fields such as engineering, biology, chemistry, computing, materials science, and communications.

Below are 10 ways nanotechnology impacts our lives on a daily basis.

1. Faster, smaller, and more powerful computers that consume far less power, with longer-lasting batteries. Circuits made from carbon nanotubes could be vital in maintaining the growth of computer power, allowing Moore's Law to continue.

2. Faster, more functional, and more accurate medical diagnostic equipment. Lab-on-a-chip technology enables point-of-care testing in real time, which speeds up delivery of medical care. Nanomaterial surfaces on implants improve wear and resist infection.

3. Nanoparticles in pharmaceutical products improve their absorption within the body and make them easier to deliver, often through combination medical devices. Nanoparticles can also be used to deliver chemotherapy drugs to specific cells, such as cancer cells.

4. Improved vehicle fuel efficiency and corrosion resistance by building vehicle parts from nanocomposite materials that are lighter, stronger, and more chemically resistant than metal. Nanofilters remove nearly all airborne particles from the air before it reaches the combustion chamber, further improving gas mileage.

5. Nanoparticles or nanofibers in fabrics can enhance stain resistance, water resistance, and flame resistance, without a significant increase in weight, thickness, or stiffness of the fabric. For example, “nano-whiskers” on pants make them resistant to water and stains.

6. Water filters that are only 15-20 nanometers wide can remove nano-sized particles, including virtually all viruses and bacteria. These cost-efficient, portable water treatment systems are ideal for improving the quality of drinking water in emerging countries.

7. Carbon nanotubes have a variety of commercial uses, including making sports equipment stronger and lighter weight. For example, a tennis racket made with carbon nanotubes bends less during impact, and increases the force and accuracy of the delivery. Nanoparticle-treated tennis balls can keep bouncing twice as long as standard tennis balls.

8. Most sunscreens today are made from nanoparticles that effectively absorb light, including the more dangerous ultraviolet range. They also spread more easily over the skin. These same nanoparticles are also used in food packaging to reduce UV exposure and prolong shelf life.

9. Many drink bottles are made from plastics containing nanoclays, which increase resistance to permeation by oxygen, carbon dioxide, and moisture. This helps retain carbonation and pressure and increases shelf life by several months.

10. A huge variety of chemical sensors can be programmed to detect a particular chemical at amazingly low levels, for example, a single molecule out of billions. This capability is ideal for surveillance and security systems at labs, industrial sites, and airports. On the medical front, nanosensors can also be used to accurately identify particular cells or substances in the body.

1.4 LIMITATIONS OF CONVENTIONAL MICROELECTRONICS

MICROELECTRONICS

Microelectronics is that area of electronics technology associated with electronics systems built from extremely small electronic parts or elements. Most of today's computers, weapons systems, navigation systems, communications systems, and radar systems make extensive use of microelectronics technology.

The research and development industry has presently been relying on the silicon based technology which was working at micro scale. The miniaturization of this technology is governed by the famous Moore’s law. But with the advent in time, silicon based technology has reached its maximum limits of the reduction in size.

- The main obstacles are in the way:
- The limits of the lithography techniques,
- The rising cost of fabrication and the size of the transistor
- Contamination occurs during high temperature treatment of silicon
- The use of the silicon in device manufacturing is going through the fundamental limitations including the tunneling current and the sub threshold current which lead to the high power consumption and circuit failure.

Ultimately the goal of the scaling down process is to build an individual transistor which is smaller, faster, cheaper and consumes less power. Unfortunately, the scaling down of the lithographically patterned transistors could not continue forever, but the nanoelectronics may be able to continue this scaling down process.

1.5 TRENDS IN NANOELECTRONICS AND OPTOELECTRONICS

The evolution of microelectronic devices is influenced by factors such as growing demand in memory capacity of integrated circuits, high transmission data speed, optical communications, etc. This requires electronic devices with faster speed operation and smaller size as is the case of silicon integrated circuits, in particular dynamic random access memori (DRAMs). Figure shows the reduction of feature size of metal-oxide-semiconduct (MOS) transistors, as well as the number of bits per chip for the period 1970-2000 [1 For example, a 256 M-bit DRAM contains about 10^9 transistors with a feature size L close to 100 nm. For structures with these dimensions, transport can still be treated classically but we are already at the transition regime to quantum transport.

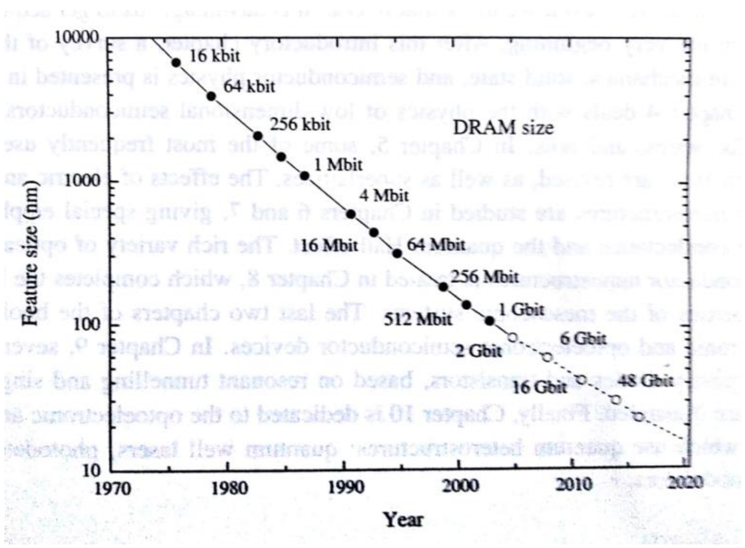


Fig 1.1 : Evolution of the minimum feature of Si DRAM

Today it believed that present silicon technology will evolve towards feature sizes still one order of magnitude lower, i.e. L~10 nm; but below this size, transistors based on new concepts like single electron transistors, resonant tunnelling devices, etc. will have be developed. The operation of this new kind of devices has to be described by the concepts of mesoscopic and quantum physics. It is interesting to remark that quantum effects show up in III-V devices for larger feature sizes, as a consequence of the smaller value the effective mass, and therefore larger value of the de Broglie wavelength In the near future, and due to the growing demands of calculus from industries like communications, information, military, space, etc. microelectronics will be replaced nanoelectronics since the feature size of electronic devices will be reduced to about 10nm. Although there has been an impressive advance at the device and technology level during the last decade, the progress in the development of new system architectures dealing with billions of transistors is scarce. In this sense, new architectures for parallel data processing are under current development.

The evolution towards nanoelectronics has been possible because of advances in the deposition of very thin films to form heterostructures in which electrons could be confined to a 2D mesoscopic system. Standard evaporation and sputtering techniques did not produce heterostructures of enough quality. However, during the 19803 both Molecular Beam Epitaxy (MBE) and Metal Organic Chemical Vapour Deposition (MOCVD) became available .In MBE, a film of the desired material is epitaxially grown over a substrate under ultrahigh vacuum conditions (pressure less than 10^{-10} mbar). The materials are evaporated in tubular effusion cells of the Knudsen type at fairly high temperatures. The atoms or molecules emerge from the effusion cells forming a molecular beam that travels in straight lines towards the substrate where the molecules condense.

Very thin films of elemental semiconductors (Si, Ge), III-V (AlGa, AlGaAs, InP), and II-VI compounds (CdTe, PbS) can be deposited by MBE. This technique allows a layer-by-layer growth of films and superlattices as well as the doping of the material from sources such as B, Al, As, etc. Although the deposited films are of very high quality, the technique is somewhat slow and expensive.

MOCVD is the preferred industrial technique to produce semiconductor heterostructures. For instance, GaAs can be grown over a substrate at about 500 °C by the reaction of arsine (AsH_3) gas and trimethyl gallium $(CH_3)_3Ga$ at sub-atmospheric pressures. Diluted dopant gases are also simultaneously introduced; for instance Si n-type doping of GaAs can be obtained by means of SiH4 diluted in hydrogen which is used as carrier gas. This technique allows for simultaneous deposition on several wafers and is used for the commercial production of semiconductor lasers.

In general, mesoscopic systems require the formation of nanostructures in the range close to 100nm, that is, a decrease in specifications of about one order of magnitude in comparison to the state-of-the-art some 20 years ago. Therefore at present we are close to the limits of conventional optical lithography, and other high resolution nanolithographic techniques (electron beam, ion beam, x-ray, etc.) have to be industrially implemented. As for resists, the most commonly used in nanolithography is the positive tone resist known as PMMA (poly-methylmethacrylate). Although the molecular weight of PMMA is close to 10^6, its roughness, once spin coated onto the substrate, is only about 2 nm.

A frequent matter of present discussion is the ultimate limits in device size. taking into account the evolution shown in Figure 1.1. It seems reasonable that the rate of scaling down predicted by Moore's law will have to slow down. It is expected that the limits of further miniaturization, from an industrial and economic point of view, will be reached in about one decade. The technological limits are related to several factors of which we will mention only two. The first has to do with the amount of heat generated by the

power consumed, and which cannot be eliminated because of thermal conductivity limits of the materials and the increasing number of overlayers. The other factor is related to the so-called "parameter spread" in fabrication. For instance many of the electrical parameters of the MOS transistors are set by doping; however, as the size of the region to be doped is decreased to about 0.1 μm^3, the number of doping atoms becomes so low (about 10) that the parameter spread cannot be controlled appropriately.

In addition to the above mentioned technological limits, there are others of fundamental nature, which are called *physical limits*. Although at present integrated devices are still far from these limits, we think it is important to revise them. They are the following:

- ***Thermal limit*:** The energy necessary to write a bit should be several times kT which is the average energy of thermal fluctuations. In CMOS the lowest values contemplated to write a bit should not be smaller than about 2eV, i.e. 100kT at room temperature or ~3 x 10^{-19} J.

- ***Relativistic limit:*** Signals cannot propagate faster than the speed of light.Therefore assuming that the nucleus of a microprocessor has a size of a few cm, it takes 10^{-10} s for the signal to propagate, which corresponds to a frequency of about 10 GHz.
- ***Uncertainty principle***: According to the Heisenberg's uncertainty principle, the energy and time needed to write or read a bit should be related by $\Delta E \times \Delta t \geq h$. To be safe, we ask the product $\Delta E \times \Delta t = 100h$. Since for future circuits ΔE could be as low as 10^{-19} J, we can appreciate that we can approach the quantum limit as the frequency increases.

At present nanoelectronics is moving simultaneously along several directions. One of them is solid state nanoelectronics, usually consists of heterostructures of well-known materials (Si, Si02, III-V compounds), and several types of transistors: heterojunction, single-electron, resonant-tunnelling, ballistic, etc. However, the amount of computational capacity for some tasks like speech and visual recognition is so large that other radically different alternatives are being sought. Some of these alternatives, like superconductivity electronics and spintronics, use fabrication techniques not too different from those employed in present integrated circuit technologies.

Superconducting electronics, proposed in the 1970s, and developed to the prototype stage during the 1980s, is based on the switching properties of the Josephson junction, which consists of two superconducting layers separated by a very thin oxide insulating film that can be tunnelled by superconducting pairs of electrons. The advantages of superconducting electronics are based on the fact that Josephson junctions can operate at high switching speeds (switching times between 1 and l0 ps), the amount of dissipated power is very low and the resistance of interconnect superconducting lines is practically null.

Another technology being pursued is spintronics which exploits the spin orientation of electrons. Electron-spin transistors are built by enclosing a semiconductor layer (base) sandwiched between two ferromagnetic layers (emitter and collector). Electrons acquiring the magnetization state of the emitter can only travel through the collector across the base if their spins are aligned with the magnetization of the collector. These developments can be considered as running parallel to the efforts in magnetoelectronics to develop MRAM memories, based both in the giant magnetoresistance effect and in the magnetic tunnel junctions developed in 1995. Electron-spin transistors have a big potential if successfully integrated with CMOS circuits.

Other radically new alternatives for future nanoelectronics have been proposed. Molecular electronics is based on the different states or configurations that molecules can take, like "trans" or "cis", as well as parallel or antiparallel alignment of the spin of unpaired electrons. The change between states must be fast, consume little energy, should be addressed by some external signal, and should be readable by a probe. If this technology is able to be put to work, we will have the ultimate step in miniaturization, since molecules are much smaller than present feature sizes in integrated circuits. Besides, molecules have the advantage of being capable of self-organization in 3D supramolecular entities, although, after the development of scanning atomic force microscopy, molecules can be, in principle, individually manipulated. Examples of molecules that can be used in molecular electronics are: azobenzene, hydrazobenzene, etc. One advantage of organic molecules, when compared to inorganic ones, is that it is easier to isolate them in single molecular systems. Although in the future it is expected that molecular wires or nanotubes can be developed for contacting molecules, today only metallic or semiconducting electrodes are contemplated. Even with this limitation, the interfacing of molecules to the external world for addressability in the case of large systems seems, at present, an enormous problem.

Lastly, we mention biology-inspired electronics, also called bioelectronics. In trying to copy nature, we are not concerned with the size of the building blocks, since for instance, a neuron is very large for nanotechnology standards. What nanoelectronics would like to imitate of biological neurons is the capabilities in parallel processing as well as their 3D architectures and the topology of the interconnects. This results from the large number of computations, needed for instance in pattern recognition of the visual systems of humans and animals, which has to be performed simultaneously at different sites. In addition to parallel processing, neural networks try to simulate the integration of computing and memory functions, which in CMOS microprocessors are performed separately.

To conclude this section, the present situation in optoelectronics will be shortly considered. Optoelectronic devices, based mainly in III-V direct gap semiconductors, have received a great upsurge since the development of optical fibre communications. In addition, there is at present a tendency to replace, whenever possible, electronic devices by photonic ones. The III-V semiconductors more frequently used are based on AlGaAs-GaAs and GaInAsP-InP heterostructures which cover the 0.8-1.6μm wavelength range. GaN blue lasers for short wavelength applications were developed about ten years ago. In the last two decades, quantum well semiconductor lasers with very low threshold currents and photodetectors are replacing the conventional ones, especially in long distance optical communications. One very interesting type of quantum well lasers, which operate at still lower threshold currents, is based on strained-layer quantum heterostructures.

At present, laser diodes are manufactured in chips by standard integrated circuit technology, coupled to transistors and optical interconnects, constituting the so-called optoelectronic integrated circuits (OEIC). In all cases, the trend in optoelectronic devices is to achieve a high level of integration which implies smaller sizes, but still in the micron range. Here again there is a lot of research in efficient integration architectures.

Perhaps, the greatest advances in optoelectronics based on quantum semiconductor heterostructures can be found in the field of electro-optical signal modulation. In effect, modulators based on the confined quantum Stark effect are several orders of magnitude more effective than their bulk counterparts. This is due to the fact that excitons in quantum wells have much higher ionization energy than in the bulk, and therefore can sustain much higher electric fields.

1.6 CHARACTERISTIC LENGTHS IN MESOSCOPIC SYSTEMS

Mesoscopic physics deals with structures which have a size between the macroscopic world and the microscopic or atomic one. These structures are also called mesoscopic systems, or nanostructures in a more colloquial way since their size usually ranges from a few nanometres to about 100 nm. The electrons in such mesoscopic systems show their wavelike properties and therefore their behaviour is markedly dependent on the geometry of the samples. In this case, the states of the electrons are wavelike and somewhat similar to electromagnetic radiation in waveguides. For the description of the behaviour of electrons in solids it is very convenient to define a series of characteristic lengths. If the dimensions of the solid in which the electron is embedded is of the order of, or smaller than these characteristic lengths, the material might show new properties, which in general are more interesting than the corresponding ones in macroscopic materials. In fact, the physics needed to explain these new properties is based on quantum mechanics. On the contrary, a mesoscopic system approaches its macroscopic limit if its size is several times its characteristic length. Let us next describe some of the most commonly used characteristics lengths in mesoscopic systems.

i. ***De Broglie Wavelength***: It is well known from quantum mechanics that for an electron of momentum p, there corresponds a wave of wavelength given by the de Broglie wavelength:

$$\lambda_B = \frac{h}{p} = \frac{h}{m * v}$$

In Eq. we have substituted p by m∗ v in a semiclassical description, where m* is the electron effective mass. From solid state physics, we know that inside a semiconductor, electrons behave dynamically as if their mass was m*, instead of the mass m_0 of the electron in vacuum. This observation is very important because for many interesting semiconductors, like GaAs or InSb, m* is much smaller than m_o, For instance, for GaAs and InSb, m* is equal to 0.067mo and 0.014mo, respectively

ii. Mean free path

As the electron moves inside a solid, it is usually scattered by interactions with crystal imperfections like impurities, defects, lattice vibrations (phonons), etc. In most cases, these scattering events or "collisions" are inelastic, i.e. the values of energy and momentum of the system after the interaction, differ from the corresponding ones before they interact. The distance covered by the electron between two inelastic collisions is usually called the mean free path $\boldsymbol{l_e}$. of the electron in the solid. If v is

the speed of the electron, then

$$l_e = v\tau_e$$

where τ_e is known as the relaxation time.

iii. Diffusion length

In a mesoscopic system of typical size L, the electrons can move either in the ballistic regime or in the diffusive regime. If the previously defined mean free path to is much larger than L, the particle moves throughout the structure without scattering; this is the so-called ballistic transport regime in which the surfaces usually are the main scattering entities. In hot electron transistors ,electron transport is ballistic and the electrons can reach energies much higher than the ones corresponding to the lattice thermal energy. 0n the other hand, if (l_e « L, transport can be explained as a diffusion process. In this case, the system is characterized by a diffusion coefficient D. In terms of D, the diffusion length L. is defined by

$$L_e = (D\tau_e)^{1/2}$$

where τ_e is the relaxation time. In semiconductor theory, the concept of diffusion length is used very often; for instance, if electrons diffuse within a p-type semiconductor, their concentration diminishes exponentially with distance with a decay length equal to Le.

In the diffusive regime, transport in the mesoscopic systems is usually explained by means of the Boltzmann equation, as in the bulk. On the contrary, in the ballistic regime, the Boltzmann transport model is not valid, and electrons move through the structure essentially without scattering.

iv. Screening length

In extrinsic semiconductors, the dopants or impurities are usually ionized and constitute a main factor contributing to scattering. However, in general we cannot consider that the electrical potential produced by these impurities varies with distance as 1/r . Because of the screening of free carriers by charges of the opposite polarity, the effect of the impurity over the distance is partially reduced. It is found that the variation of the potential is modulated by the term exp(-r/λs,) where λs; is called the screening length and is given by

$$\lambda s= \left(\frac{\in kT}{e^2 n}\right)^{1/2}$$

where e is the electronic charge, ∈ the dielectric constant of the semiconductor, and n the mean background carrier concentration. One should be careful about nomenclature because some authors call λs; the Debye length or the Fermi-Thomas length. In a typical semiconductor, As is in the range 10-100nm, and is an indication of the attenuation of charge disturbances in a semiconductor. From Eq. it is determined that λs; should be much smaller in metals than for semiconductors.

v. Localization length

The localization length can be understood in terms of transport in disordered materials, in which we know from solid state physics that, in addition to Bloch extended states, there can also be localized states. In disordered materials, the electrons move by "hopping" transport between localized states or from a localized state to a bound state.

In order to describe the hopping transport and other mesoscopic properties of the localized states, it is assumed that the electron wave function is described by

$$\boldsymbol{\psi} = \boldsymbol{e}^{\left(\frac{-r}{\lambda_{loc}}\right)}$$

where λ_{loc} is known as the localization length. Evidently, the electrical conductivity of a material will be proportional, among other factors, to the overlap between the wave functions. If the sample dimensions are of the order of λ_{loc}. we can say that our system is mesoscopic.

1.7 QUANTUM MECHANICAL COHERENCE

In a mesoscopic structure of dimensions similar to the electron dc Broglie wavelength λ_B. the behaviour of the electron should be described quantum mechanically, i.e. by using Schrodinger equation. 1f the electron interacts inelastically with a defect, or any impurity, the electrons change their energy and momentum, as well as the phase of their wave function. ***The phase coherence length L_Φ, is defined as the distance travelled by the electron without the carrier wave changing its phase***. Evidently, interference effects in the electron waves should only be observed if the particles move over distances of the order of, or smaller than, L_Φ. Since we are usually interested in mesoscopic systems in the quasi-ballistic regime, in which electrons are practically unscattered, L_Φ , should be a length similar to the inelastic scattering mean free path *le* previously defined. Coherent states can evidently show interference effects. On the contrary, once the coherent states loose their coherence, by inelastic scattering, the corresponding waves cannot be superposed and cannot interfere (in some way, the matter-wave becomes a particle). In mesoscopic physics, the loss of coherence is usually called dephasing. Evidently, the coherence processes are characteristic of mesoscopic systems.

From the above definition of L_Φ, electrons can show interference effects over distances smaller than L_Φ.If electrons with phase Φ1 interfere with electrons of phase **$\Phi 2$**, we know from the wave theory that the amplitude of the resultant wave varies as cos($\Phi 1$-$\Phi 2$) and the amplitudes can add up to each other, or they can be subtracted depending on the phase difference.

1.8 SCHRODINGER WAVE EQUATION

As it was established by Schrodinger in I926. The dual wave-like and particle-like properties of matter are described by the so-called wave function $\boldsymbol{\Phi}(\vec{\boldsymbol{r}},\boldsymbol{t})$. Which is continuous and has continuous derivatives. The wave function is also complex. I.e. It has a real part end an imaginary one. The wave function satisties the second-order. Linear, differential Schrodinger equation,

$$\left[-\frac{h^2}{2m}\nabla^2+\boldsymbol{V}(\vec{\boldsymbol{r}},\boldsymbol{t})\right]\Phi=ih\frac{d\Phi}{dt} \quad \text{......................(1)}$$

Where ∇^2 is the operator $\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$ and V is the potential energy. Which is generally a function of position and possibly of time. Although the function Φ does not have a physical meaning, the product of Φ by its complex conjugate is a real quantity, such that the probability dP of finding a particle in a small volume d V is given by

$$dP=|\Phi^2|dV \text{...(2)}$$

If the potential energy V is not time dependent, we can search for a solution to Eq (1)

$$\Phi(\vec{r},t)=\psi(\vec{r})e^{-j\omega t} \text{..................................(3)}$$

Substituting Eq. (3) in (1) and writing E = hω, we find the time-independent Schroidinger equation

$$\left[-\frac{h^2}{2m}\nabla^2+\boldsymbol{V}(\vec{\boldsymbol{r}})\right]\psi(\vec{r})=E\,\psi(\vec{r})$$ ………………(4)

for the time-independent wave function $\psi(\vec{r})$. Schroidinger's equation can be solved exactly only in a few cases. Probably the simplest one is that of the free particle, as for instance a free electron of energy E and mass m. In this case $\boldsymbol{V}(\vec{\boldsymbol{r}})$ = 0 and the solution of Eq. (1) is easily found to be

$$\Phi=Ae^{i(kx-\omega t)}+Be^{i(-kx-\omega t)}$$ ……………………(5)

Where

$$k=\left(\frac{2mE}{h^2}\right)^{1/2}$$ ………………………………………………(6)

Therefore, the free electron is described by a wave. which according to the de Broglie relation has momentum and energy given. respectively, by

$$p=hk,\qquad E=\frac{p^2}{2m}$$ ……………………………(7)

in general we will assume that the electron travels in one direction, for instance, along the x-axis from left to right. and therefore the coefficient B in Eq. (5) is zero. The wave equation for the free electron can simply be written as:

$$\Phi=Ae^{i(kx-\omega t)}$$ ………………………………………………(8)

Another example in which Schrodinger's equation can be solved exactly is that of the hydrogen atom for which the potential is Coulombic, i.e. V varies with the distance r between the proton and electron in the form l/r. Solving Schrodinger's equation, one gets the well-known expression for the electron energies:

$$En=-\frac{m_re^4}{2(4\pi\varepsilon_0)^2h^2n^2}=-\frac{13.6}{n^2}\,(eV)\ ,\qquad n=1,2,3..$$ ……….(9)

where m, is the reduced proton-electron mass. In solid state physics, the mathematical model of the hydrogen atom is often used, as for example, in the study of the effects of impurities and excitons in semiconductors. Although the equation giving the values of the energy is very similar to Eq. (9), the values of the energy are much smaller, since the dielectric constant of the medium has to substitute the value of the permittivity of vacuum ε_0. For instance, in the case of silicon, the value of the dielectric constant is about $12\varepsilon_0$.

The model of the harmonic oscillator is very often used in solid state physics.. In general terms, we can say that the harmonic oscillator model can be used to describe any system which performs vibrations of small amplitude about an equilibrium point. The allowed energies of the harmonic oscillator can be obtained from Schrodingers's equation and are given by:

$$En = \left(n - \frac{1}{2}\right) h\omega, \quad n = 1,2,3 \ldots \quad \ldots\ldots\ldots\ldots(10)$$

Another potential that allows the exact solution of Schrodinger's equation is the so-called infinite square well potential. which consists of a flat potential of width a surrounded by infinite potentials at its extremes. It is a fairly good approximation to modulation-doped quantum wells, which is the basic structure of many quantum well transistors and lasers.

1.9. QUANTUM WELLS, WIRES, AND DOTS

We have defined a series of characteristic lengths λ which correspond to physical properties of electrons which are size dependent. We have also seen that when the dimensions of the solid get reduced to a size comparable with, or smaller than λ, then the particles behave wave-like and quantum mechanics should be used.

Let us suppose that we have an electron confined within a box of dimensions L x, Ly, Lz. If the characteristic length is λ, we can have the following situations:

i. λ<< in Lx,Ly, LZ
   In this case the electron behaves as in a regular 3D bulk semiconductor.
ii. λ > Lx, and Lx << Ly, Lz In this situation we have a 2D semiconductor perpendicular to the x-axis. This mesoscopic system is also called a quantum well .
iii. λ > Lx, Ly and Lx, Ly < Lz; corresponds to 1D semiconductor or quantum wire. located along the z-axis
iv. (iv) λ >> Lx,Ly, Lz ,In this cease it is said that we have a 0D semiconductor or quantum dot

In general, we say in mesoscopic physics that a solid, very often a semiconductor, is of reduced dimensionality if at least one of its dimensions Li, is smaller than the characteristic length. For instance, if Lx and Ly are smaller than λ. we have a semiconductor of dimensionality equal to one. We could also have the case that λ is comparable, or a little larger, than one of the dimensions of the solid but much smaller than the other two. Then we have a quasi 2D system, which in practice is a very thin film, but not thin enough to show quantum size effects.

1.10. DENSITY OF STATES AND DIMENSIONALITY

Although the density of states (DOS) of physical systems will be derived formally in Chapter 4, in this section we see from a mathematical point of view the consequmces of the dimensionality of the system in the DOS. As we know from solid state physics, most physical properties significantly depend on the DOS function ρ. The DOS function, at a given value E of energy, is detined such that ρ (E)ΔE is equal to the number of states (i.e. solutions of Schrodinger equation) in the interval energy ΔE around E. We also know that if the dimensions

L_i(i = x, y, z) are macroscopic and if proper boundary conditions are chosen, the energy levels can be treated as a quasi-continuous. 0n the other hand, in the case where any of the dimensions Li gets small enough, the DOS function becomes discontinuous. Let us next obtain the DOS function for several low-dimensional solids. First, let us remind that for bulk solids ρ(E) varies with energy in the form $\sqrt{E}$

If each quantum state or Bloch state in a solid is designated by a quantum number k (Bloch state), the general expression for the DOS function should be

$\rho(E) = \sum_k \delta(E - E_k)$ ……………………………………1.8.1

where the quantized energies are given by Ek and δ (E) is the Dirac's delta function. If we take into account the electron spin degeneracy, a factor 2 should also appear in the above expression. Let us recall for simplicity the case of a cubic shaped 3D macroscopic crystalline solid, of edge

L = Na, where a is the lattice constant and N the number of sites along the one-dimensional directions which is supposed to be large. In this case, the eigenstates can be considered as quasi-continuous and the summatory in k of Eq. can be replaced by an integral, i.e.

$\sum_k \rightarrow \frac{L}{(2\pi)^3} \int dk$ ……………………………………………………1.8.2

for the case of a cube of size L and volume V = L^3 .We also know from simple solid state theory that if we assume that the energy E_k only depends on the magnitude of k in a parabolic energy; dependence between the momentum hk, i.e.

$E_k = \frac{h^2k^2}{2m^*}$ ………………………………………..1.8.3.

$\rho_{3D}(E) = \frac{V}{2\pi^2}\left(\frac{2m^*}{h^2}\right)^{3/2}\sqrt{E}$ …………………..1.8.4

We can follow exactly the same procedure for 2D and 1D semiconductors of area A and length L, respectively, reaching the following expressions

$\rho_{2D}(E) = \frac{A}{\pi}\left(\frac{m^*}{h^2}\right)$ ………………………………1.8.5

$\rho_{1D}(E) = \frac{L}{2\pi}\left(\frac{2m^*}{h^2}\right)^{1/2}\frac{1}{\sqrt{E}}$ …………………1.8.6

Some important considerations can already be made: the DOS function in 3D semiconductors is proportional to $\sqrt{E}$, in 2D is constant, and in 1D varies inversely pmpomonal t0 $\sqrt{E}$. This implies in the last case that at the bottom of bands, the DOS plays a very important role. because there is a singularity for E = 0.

Eqns 1.8.5. & 1.8.6. were derived for perfectly 2D and 1D solids. but in the real world a 2D solid. for instance, is really a 3D one where the perpendicular dimension is very short. We will see that since electrons can move almost freely in the (x, y) plane, Eq. 1.8.5. should be written for a quasi-2D solid as

$$\rho_{2D}(E) = \frac{A}{\pi}\left(\frac{m^*}{h^2}\right)\sum_{n_z}\theta(E - E_{n_z})$$

where n_z; refers to the quantization in the contined z-axis and θ is the step function.

Similarly, for a quasi-lD solid or quantum wire along the z-direction

$$\rho_{1D}(E) = \frac{L}{2\pi}\left(\frac{2m^*}{h^2}\right)^{1/2}\frac{1}{\sqrt{E - E_{nx,ny}}}$$

where n_x and n_y, are the quantum numbers for the confined x and y directions. This function looks like a series of peaks, one for each value of $E_{nx,ny}$.

Evidently, in the case of a quasi-OD solid or quantum dot, there is no continuous DOS function, since there is quantization in the three spatial directions. Therefore the DOS function consists of a series of peaks given by

$$\rho_{0D}(E) = \sum_i \delta(E - E_i)$$

Where i=(nx,ny,nz) , the peaks in quantum dots are not perfect δ -functions since there is a broadening effect as a consequence of scattering mechanisms.

1.11 BASIC PROPERTIES OF 2D SEMICONDUCTOR NANO STRUCTORS

One of the most practical two-dimensional semiconductor structures consists of a sandwich of gallium arsenide (GaAs), with a thickness in the nanometre range, surrounded on each side by a semiconductor such as aluminium gallium arsenide (Al_xGa_{1-x},As) of higher bandgap. The bandgap of Al_xGa_{1-x},As (x ~ 0.3) is close to 2.0 eV while that of GaAs; is 1.4 eV. As a consequence, the potential energy profile has the shape of a square well , with a barrier height of 0.4 eV for electrons and 0.2 eV for holes. In reality, the profile of the potential barrier is somewhat more complicated, since the potential varies with atomic distances, which also affects the wave functions. However, in most cases, it is a good approximation to consider the average over a few atomic distances (envelope function approximation). As seen in Figure (a), carrier motion for both electrons and holes is not allowed in the direction perpendicular to the well, usually taken as the z-direction because of the potential walls. However in the other two spatial directions (x, y), parallel to the semiconductor interfaces, the motion is not restricted. is the electrons behave as free electrons.

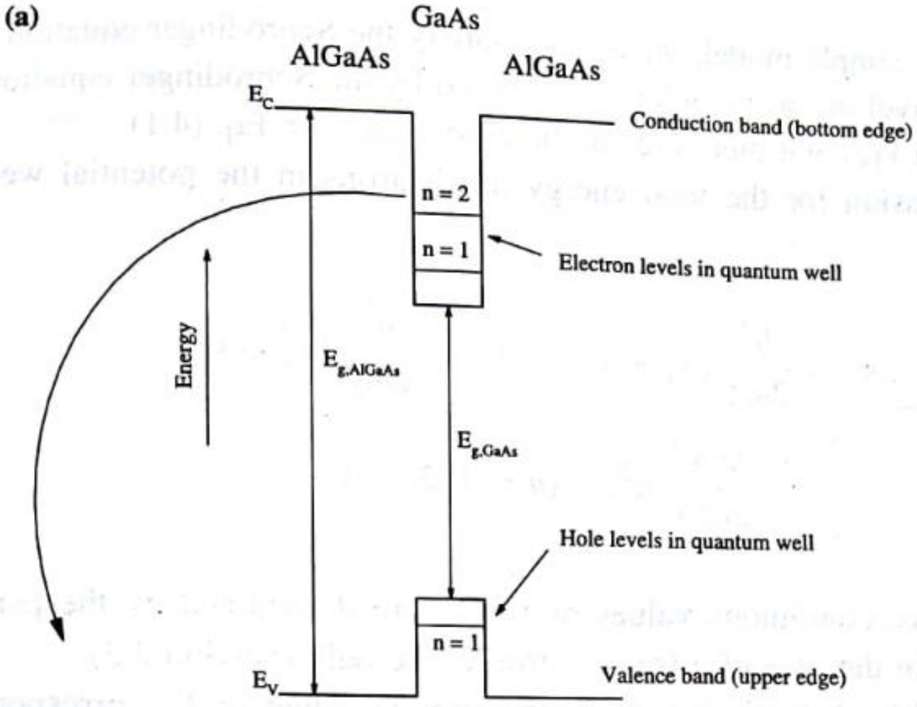


Fig 1.9; (a) AlGaAs-GaAs-AlGaAs Square potential well

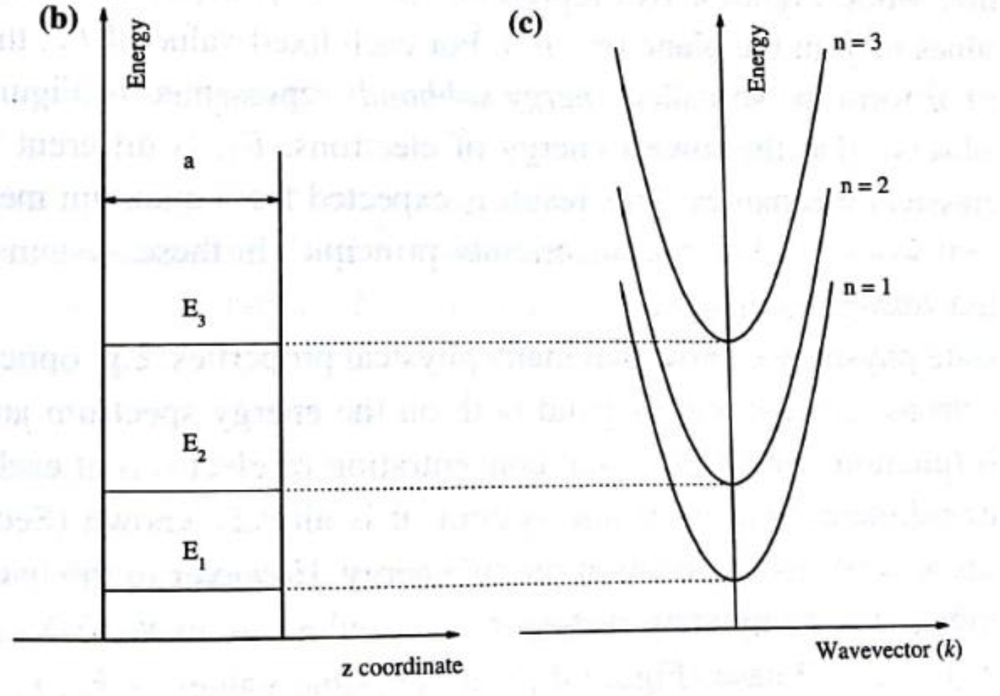


Fig1.9 (b): Energy Levels Fig1.9 : (c) Energy subbands

The behaviour of electrons when their motion is restricted along one direction in the wells of infinite height corresponds to a well-known problem in quantum mechanics, d! so-called particle in a box of infinite wells. It is well known from quantum mechanics that, in the case of infinite potentials barriers, the wave functions and energy levels of the bound electrons are given by

$$\psi_n(z) = \left(\frac{2}{a}\right)^{1/2} sin\left(\frac{\pi n z}{a}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$En = \frac{h^2\pi^2}{2{m_e}^{*}a^2}n^2 \quad , \quad n=1,2,3,\ldots.. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

where ${m_e}^{*}$ is the effective mass of the electrons in the well material for the motion along the

z-direction and a is the width of the well. From Eq. (2) we can derive several important consequences:

(1) In general, quantum size effects will be more easily observable in quantum structures of very small size a, and for materials for which the electron effective mass is as small as possible. In this sense, GaAs nanostructures are very convenient since ${m_e}^{*}$ ~0.067mo, where mo is the free electron mass. This is equivalent to saying that in materials for which the electron mobility or the free electron path are large, quantum effects are easier to observe.

(2) Quantum size effects, which require energy transitions of electrons between levels, are better observed at low temperatures. since the mean thermal energy of carriers is of the order of kT.

As it has been described, the motion of electrons in the quantum well is confined only in one direction, z:, but in the (x, y) planes the electrons behave as in a three-dimensional solid. Therefore the electron wave function is separable as the product of **ψ**x, **ψ**y, **ψ**z, ie.

ψ= **ψ**x **ψ**y **ψ**z ……………………(3)

where, in our simple model, **ψ**x and **ψ**y, satisfy the Schrondinger equation for a free electron, i.e. a travelling wave, while **ψ**z is given by the Schrodinger equation with a Square Well potential V(z) and therefore can be expressed as in Eq. (1).

The expression for the total energy of electrons in the potential well, can then be written as

$$E(kx, ky, n) = \frac{h^2}{2m_e{}^*}\left({k_x}^2+{k_y}^2\right) + En = \frac{h^2}{2m_e{}^*}\left({k_x}^2+{k_y}^2\right) + \frac{h^2\pi^2}{2m_e{}^* a^2}n^2 \text{ ……………………(4)}$$

where the quasi-continuous values of kx, ky, are determined by the periodic boundapy conditions as in the case of a free electron in the bulk.

Figure (b) schematically shows the discrete values of En corresponding to motion in the z-direction while Figure (c) represents the E vs p dependence (remember the p = hk) for values of $\vec{p}$ in the plane (px, py). For each fixed value of En... the values of E as a function of $\vec{p}$ form the so-called energy subbands represented in Figure(c).It is interesting to observe that the lowest energy of electrons, E1, is different from zero, in contrast with classical mechanics. This result is expected from quantum mechanics since a value of E = 0 would violate the uncertainty principle. In these systems the valued E = E1 is called zero-point energy.

From solid state physics we know that many physical properties, e.g. optical absorption. transport of electronic current, etc. depend both on the energy spectrum and the denim of states (DOS) function, which gives the concentration of electrons at each value oi the energy. In a three-dimensional electronic system, it is already known that the DOS depends as a parabolic function on the energy. However in the two-dimensional case, this dependence is completely different. Proceeding as in the 3D case, it can N appreciated that for the 2D case the possible values of kx,

ky, are separated by 2π/L, where L is the dimension of the sample, which has been assumed to be square without loss of generality. The number of states in the k-space within a circular ring limited by the circumferences of radii k and k + dk is therefore:

$n_{2D}(k)dk = \frac{2\pi k dk}{(2\pi/L)^2}$ …………..(5)

And the number of states in k –space per unit area is

$n_{2D}(k) = \frac{k}{2\pi}$ ……………………..(6)

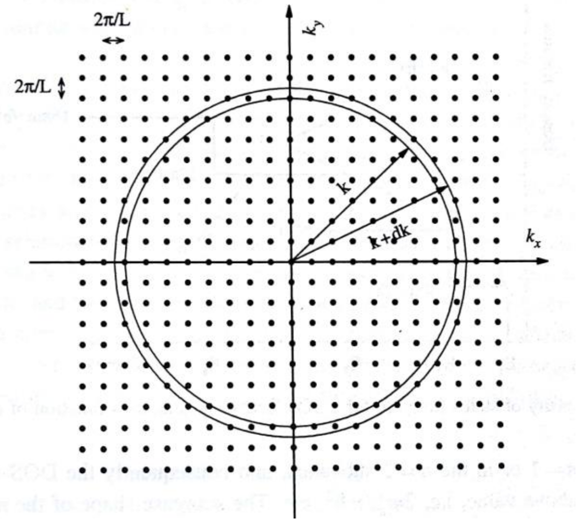


Fig: 1.10 Representation in k space for the states (kx,ky) of a 2D electron system

If we wish to calculate the DOS in energy, we define $n_{2D}(E)$ such that $n_{2D}(E)\,\delta E$ is the number of states in the range δ E. The densities of states in energy and wave vector are related by:

$n_{2D}(E)\,\delta E = n_{2D}(k)\,\delta k$ ……………………………………(7)

where E and k are related by $E = \frac{h^2k^2}{2m_e{}^*}$. Differentiating this expression and taking into account Eq. (6), we have, after substitution. in Eq. (7), and adding a factor 2 which accounts for the spin:

$n_{2D}(E) = 2\frac{k}{2\pi}\frac{\delta k}{\delta E} = 2\frac{k}{2\pi}\frac{m_e{}^*}{h^2k} = \frac{m_e{}^*}{h^2\pi}$ ……………………..(8)

Note that in the 2D case, the DOS function is a constant, independent of energy. Let us show next that the DOS function, for the two-dimensional case, exhibits a staircase-shaped energy dependence (Figure 1.11) in which all the steps are of the same height, but located at energies En given by Eq. (4.2). In effect, from Figure 1.9 (c) it can be appreciated that the interval of energy between 0 and E1 is not allowed. For E such that E1 < E < E2 the electrons will be located in the subband corresponding to n = 1 and the value will be $\frac{m_e{}^*}{h^2\pi}$. For the energy interval between E1; and E3, the electrons can be located either in the n =1 or in the n = 2 subbands, and consequently the DOS function would be twice the above value, i.e. $\frac{2m_e{}^*}{h^2\pi}$, etc. The staircase shape of the $n_{2D}(E)$ function can be directly observed by optical absorption measurements. We have also represented the parabolic 3D case in Figure 4.3, from which it can be appreciated that the differences between the 2D and 3D cases are more discernable for low values of n.

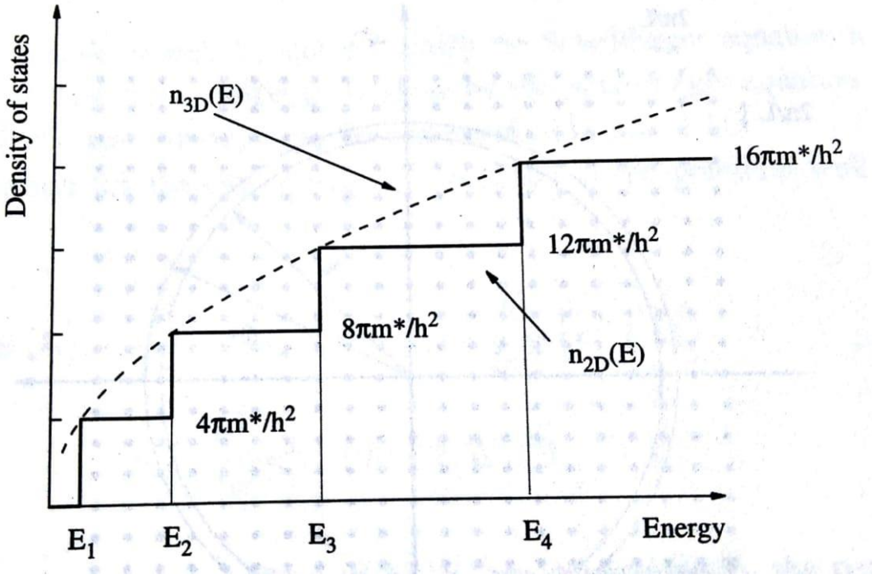


Fig: 1.11 Density states function for 2D electron system as a function of energy

According to the subband in which the electrons are located, their kinetic energy is differently partitioned as a consequence of the relationship expressed in Eq. (4). For instance, for a given energy in the interval E2 to E3, for the same value of energy, the electrons located in subband n = 2 have higher energy in the z-direction, E2, than those in the n =1 subband for which the corresponding value is El. Therefore the energy of motion corresponding to the plane (px , py) should be smaller for the electrons in subband n = 2. Evidently the separation of energy in different “components" (remember that the energy is a scalar function) is a direct consequence of the simple forms adopted for **ψ** and energy in eqns 3 & .4

1.12 SQUARE QUATUM WELL OF FINITE DEPTH

The quantum wells for electrons and holes in GaAs nanostructures surrounded by higher gap AlGaAs,are not of infinite height.In fact, the value of the height of the potential for electrons should coincide with the discontinuity ΔEc that appears at the interface in the conduction bands of AlGaAs and GaAs, which for the above system is of the order of some tenths of an eV. However, it is fairly easy to deduce that for electron energies in the quantum well not too close to the barrier ΔEC (we take the energy as zero at the bottom of the well), the values obtained for the case of infinite wells do not differ too much of those obtained in the case of finite depth wells.

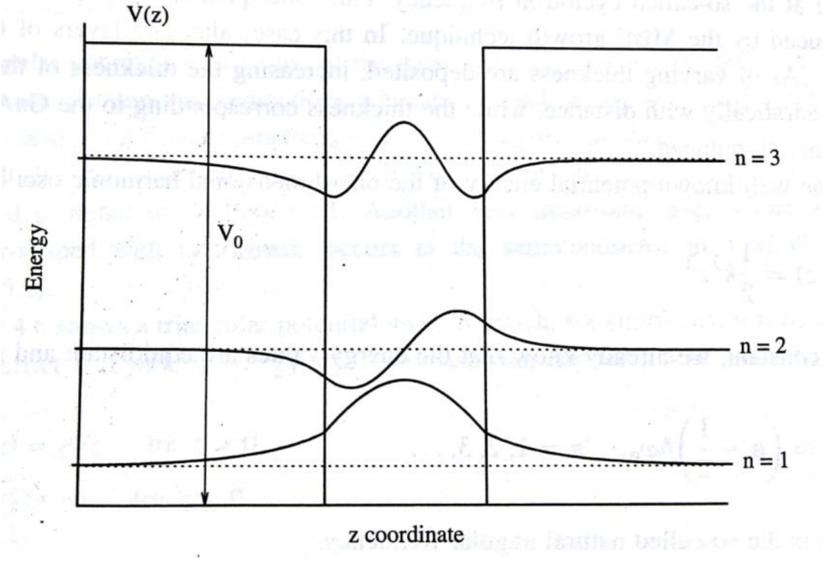


Fig 1.12: Finite potential square well

If we call Vo in Figure 1.12 the height of the finiite square well, it is evident that for states with energy E < Vo we have bound states, i.e. the electrons are trapped inside the well of width a, while for E > Vo, we have continuous propagation states, in which the electrons are free to move from z = -∞ to z = +∞. Since this problem presents inversion symmetry around the centre of the well, this point is chosen as origin for the z-direction. In relation to the bound states, the wave functions inside the well should have the same shape as in the case of the infinite well, i.e. the solutions are symmetric or antisymmetric, and therefore should be sine or cosine functions, respectively. We also know from quantum mechanics that the solutions outside the well, which are obtained from the Schrodinger equation with a potential energy equal to Vo, are exponential decay functions. Therefore the solutions for the wave functions should be linear combinations of the functions:

$$\psi_n(z) = \begin{cases} D\exp(kz)\,, & z > -\frac{a}{2} \\ C\cos(kz), C\sin(kz)\,, & -\frac{a}{2} < z < \frac{a}{2} \\ D\exp(-kz)\,, & z > a/2 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Where

$$k = \left(\frac{2m_e{}^*E}{h^2}\right)^{1/2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$k = \left(\frac{2m_e{}^*(Vo-E)}{h^2}\right)^{1/2} \text{ outside the well } \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Note that in Eqs (2 & 3) we have assumed the same value for the effective electron mass in the barrier and in the well. Usually, this is a good approximation because the barrier material is similar to the well material (e.g. GaAs and AlGaAs with a small Al mole fraction) and because the penetration of electron wave functions into the barriers is small for the lowest states. In order to proceed with the solution of the problem we should next realize that since (z) is a continuous function, therefore, the functions of Eq. (1) must be equal for z = ± a / 2, and the same should hold for their derivatives in the homogeneous effective mass approximation. From this fact a transcendental equation is derived which can be easily solved numerically . From this solution, several important facts are derived. For instance, it can be shown that in a one-dimensional well there is always at least one bound state, independent of how small the value of Vo might be. For the case of weakly bound states, the exponential decay constant k in Eq. (1) is small and therefore the wave function represented in Figure 1.12 penetrates deeply into the barrier region. Evidently, the opposite should be true for strongly bound states, in which the penetration in the energy forbidden region is very small.

1.13 PARABOLIC AND TRIANGULAR QUANTUM WELLS

Parabolic well

The case of the parabolic well is well known in solid state physics since the vibrations of the atoms in a crystal lattice, whose quantification gives rise to phonons, are described in a first approximation by harmonic oscillators. In addition, a magnetic field applied to a two-dimensional electron system gives rise to a parabolic potential, and the electrons oscillate at the so-called cyclotron frequency. Parabolic quantum well profiles can also be produced by the MBE growth technique. In this case, alternate layers of GaAs and $Al_xGa_{1-x}As$ of varying thickness are deposited, increasing the thickness of the AlGaAs layer quadratically with distance, while the thickness corresponding to the GaAs layer is proportionally reduced.

For the well-known potential energy of the one-dimensional harmonic oscillator

$$V(z) = \frac{1}{2}\, k^2 z^2 \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

With k constant

$$En = \left(n - \frac{1}{2}\right) h\omega_o \qquad n = 1,2,3 \ldots. \qquad \ldots\ldots\ldots\ldots\ldots(2)$$

Where ω_o is the natural angular frequency.

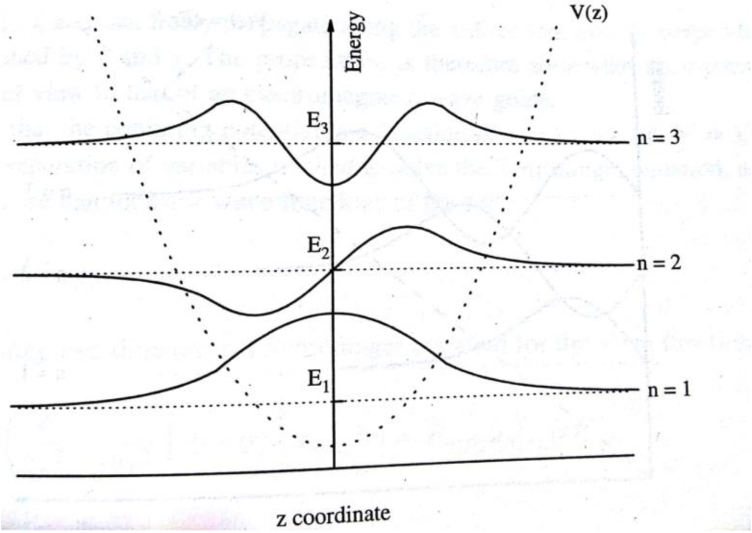


Fig 1.13 : Harmonic oscillator potential well V(z)

Figure 1.13 shows the case of a parabolic well potential and the wave functions of three electron bound states. The wave functions are mathematically expressed in terms of the Hermite polynomials. Note, as in the case of the square well, the symmetric or antisymmetric character of the wave functions and their exponential decay in the forbidden energy zone.

Triangular wells

The triangular potential well is one of the most common geometries, since the potential profile across quantum heterojunctions , such as the well-known modulation doped AlGaAs-GaAs heterojunction, is almost triangular in shape for electrons within GaAs. of all heterostructures, this is probably the most investigated one. Another very important case, where an almost triangular-shaped well is formed, occurs at the semiconductor in a MOS structure.

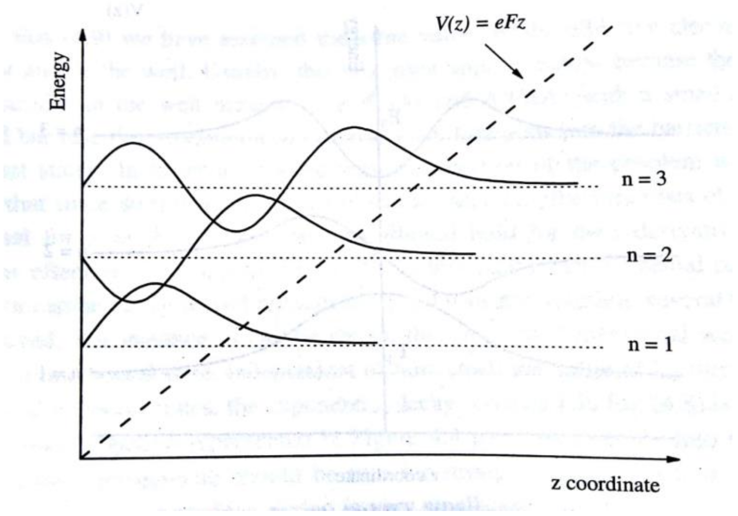


Fig 1.14 : triangular potential well V(z)

Figure 1.14 shows a triangular potential well, in which, for simplicity, it is assumed that the left barrier is infinite in energy and it increases linearly for z > 0:

$$V(z) = e\,Fz, \qquad for\ z > 0$$

$$V(z) = \infty\,, \qquad for\ z \leq 0$$

where e is a constant equal to the electron charge and F is a uniform electric held. As in the other cases, the electron energies and states are found by solving the Schrodinger equation subject to the boundary condition (z = 0) = 0. In this case, the eigenvalues are given in terms of the Airy functions. However, for small values of n, it can be demonstrated by applying the WKB quantum-mechanical approximation, that

$$En \approx \left[\frac{3}{2}\pi\left(n-\frac{1}{4}\right)\right]^{2/3}\left[\frac{e^2h^2F^2}{2m}\right]^{1/3}, \qquad n=1,2,3,\ldots.$$

Figure 1.14 shows the spacing between the energy levels, which get a little closer as n increases, in contrast to the square well where the levels become further apart as n increases (in the parabolic case they were equally spaced). In the same figure, the wave functions are also represented. Observe that as n increases, the wave functions add one more half' cycle. However at difference with the previous parabolic case, the wave functions are neither symmetric or antisymmetric due to the asymmetry of the potential well.

1.15 QUANTUM WIRES

Having considered the two-dimensional electron gas , it is easy to understand that in the one-dimensional electron gas, the electrons should be confined in two directions, (x, y), and can freely propagate along the z-direction, usually perpendicular to the plane defined by x and y. The propagation is therefore somewhat analogous from a formal point of view to that of an electromagnetic wave guide.

Supposing that the confining potential is a function of r = (x , y), ie. V = V (r), and following the separation of variables method to solve the Schrodinger equation, we can look for wave functions of the form:

$$\psi(r) = e^{jk_zz}u(r)\ldots\ldots\ldots\ldots\ldots..(1)$$

With the following 2D Schrodinger equation for wave function u(r)

$$\left[-\frac{h^2}{2m}\left(\frac{d}{dx^2}+\frac{d}{dy^2}\right)+\boldsymbol{V}(\boldsymbol{r})\right]u_{n1n2}(r) = E_{n1n2}(r)\ \ldots\ldots\ldots..(2)$$

Where n1,n2 = 1,2,3.....

The expression for the total energy of electron in the quantum wire should be of the form

$E_{n1n2}(kz) = E_{n1n2} + \frac{h^2 k_z{}^2}{2m_e{}^*}$ ……………………………..(3)

where the last term represents the kinetic energy of the electron propagating along the z-direction.

As an example to get specific expressions for the energy, we consider now the simplest case of a two-dimensional rectangular potential of infinite depth and size ax, ay,. That is:

$$V(x,y) = 0 \qquad 0 < x < ax, 0 < y < ay$$

$$V(x,y) = \infty \qquad x \leq 0\ , \quad x \geq ax, y \leq 0\ , y \geq ay$$

Now energy

$E_{n1n2} = \frac{h^2\pi^2}{2m_e{}^*}\left\{\frac{n_1{}^2}{a_x{}^2} + \frac{n_2{}^2}{a_y{}^2}\right\}$, n1, n2 = 1,2,…….. ………..(4)

Another case, relatively easy to solve in cylindrical coordinates, is that of a wire with circular cross section, in which case the solutions are given in terms of the Bessel functions. Therefore in the case of quantum wires, the energy levels corresponding to the transverse direction are specified by two quantum numbers, and each value $E_{n1,n2}$ is now the bottom Of a parabolic one-dimensional subband in kz space. Observe also that as the electron moves in a narrower wire, the energy corresponding to the $E_{n1,n2}$ levels increases; .

Let us now calculate w density of states for the one-dimensional electron gas. The concentration of states in energy is related to that in wave number by the expression:

$n_{1D}(E)\partial E = n_{1D}(E)\frac{dE}{dk}\partial k = 2n_{1D}(k)\partial k$ …………………..(5)

The factor 2 appears because the wave number could be either positive or negative corresponding to the two directions along the wire. The density of states in k-space per unit length is 1/2π, as seen from considering the one-dimensional version of Figure 1.10 Substituting in Eq. (5) and taking into account that $E = \frac{h^2k^2}{2m_e{}^*}$, we obtain:

$n_{1D}(E) = \frac{1}{\pi h}\sqrt{\frac{2m_e{}^*}{E}}$ ……………………………(6)

Which diverge for E=0

In terms of group velocity vg,

$n_{1D}(E) = \frac{2}{\pi h v_g}$ ……………………………(7)

One interesting result of this equation is that the current in a one-dimensional system is constant and proportional to the velocity and the density of states. The expression (7) for the DOS function in a quantum wire, will have important consequences as, for example, the quantized conductance.

The expression of the total DOS per unit length for a quantum wire can be expressed as

$$n_{1D}(E) = \sum_{n1,n2} \frac{1}{\pi h} \sqrt{\frac{2m_e^*}{E - E_{n1,n2}}} \quad \text{……………………………….(8)}$$

Figure 1.15 represents the DOS for a one dimensional System, which is compared to the parabolic 3D case. Now the DOS diverges at the bottom of the subbands given by the energy values $E_{n1,n2}$ This result will have important consequences in the physical properties of quantum wires.

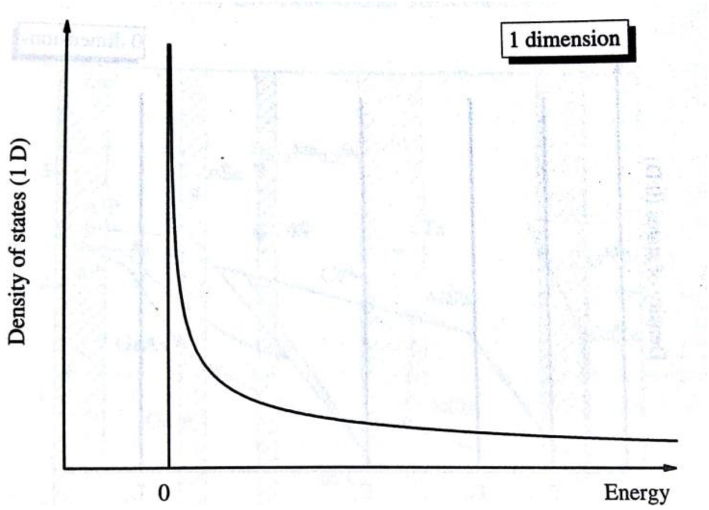


Fig 1.15 : DOS function for 1D electron system, as a function of energy

1.16 QUANTUM DOTS

Quantum dots are often nano crystals with all three spatial dimensions in the nanometer range. Sometimes, as is the case of the II-VI materials, such as CdSe or CdS, the nano crystals can be grown from liquid phase solutions at well-specified temperatures. Conversely they can also be prepared by lithographic etching techniques from macroscopic materials

Although the word “dot” implies an infinitely small size, in practice dots might have a large number of atoms: 10^4-10^6, and still have their three dimensions in the nanometre legion, so that the electron de Broglie wavelength is comparable to the size of the dot. In this case, the wave nature of the electron becomes important. Quantum dots are often referred to as ***artificial***

atoms, because, the spectrum of the energy levels resembles that of an atom. In addition, at least in theory the energy spectrum can be engineered depending on the size and shape of the dot. In analogy to atoms, we can also define an ionization energy, which accounts for the energy necessary to add or remove an electron from the dot. This energy is also called the ***charging energy*** of the dot, in an image similar to the concept of capacitance of a body, in which the addition or subtraction of electric charge is specified by the Coulomb interaction. Therefore the atom-like properties of the quantum dots are often studied via the electrical characteristics. From this point of view, it is very important to remark that even the introduction or removal of one single electron in quantum dots, in contrast to the case of 2D or 1D systems, produces dominant changes in the electrical characteristics, mainly manifested in large conductance oscillations and in the Coulomb blockade effect .

Let us now study the energy spectrum of quantum dots. The simplest case would be that of a confining potential that is zero inside a box of dimensions ax, ay,, and az, and infinite outside the box. Evidently the solution to this well-known problem are standing Waves for the electron wave function and the energy levels are given by

$$E_{n1,n2,n3} = \frac{h^2\pi^2}{2m_e{}^*}\left(\frac{n_1{}^2}{a_x{}^2} + \frac{n_2{}^2}{a_y{}^2} + \frac{n_3{}^2}{a_z{}^2}\right) \quad ; \quad n1, n2, n3 = 1,2,3, \ldots.$$

In contrast to the 2D and 1D cases, now the energy is completely quantized, and, as in the case of atoms, there is no free electron propagation. However, the levels are frequently degenerate, for instance, if two or three of the dimensions of the box are equal.

The case of a spherical dot in which the potential is zero inside the sphere and infinite outside can also be exactly solved, and the solutions are expressed in terms of the Spherical functions. This problem resembles that of a spherically symmetric atom and the energy depends on two quantum numbers, the principal quantum number n, arising from the one-dimensional radial equation, and the angular momentum quantum number *l.*

Since in the case of quantum dots the electrons are totally confined, the energy spectrum is totally discrete and the DOS function is formed by a set of peaks in theory with no width and with infinite height (Figure 1.16).

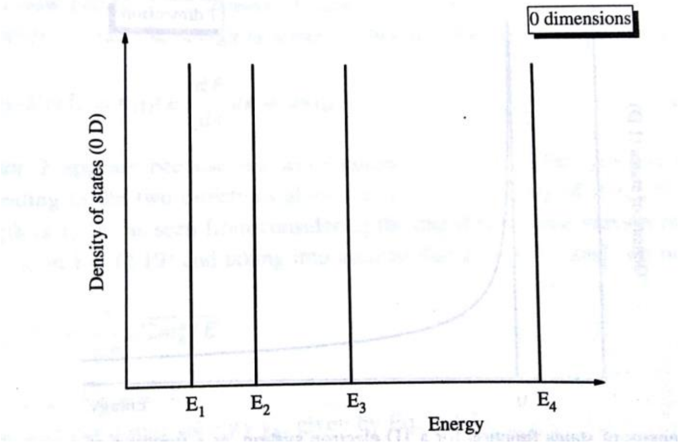


Fig 1.16: DOS function for 0D electron system.

Evidently, in practice, the peaks should have a finite width, as a consequence, for instance, of the interaction of electrons with lattice phonons and impurities.

1.17 CARBON NANOTUBES

One of the greatest impacts on nanoscience and nanotechnology has been made

by the carbon nanotubes. A schematic illustration of carbon nanotubes is presented in Fig, and an atomically resolved image generated using scanning tunnelling microscope.

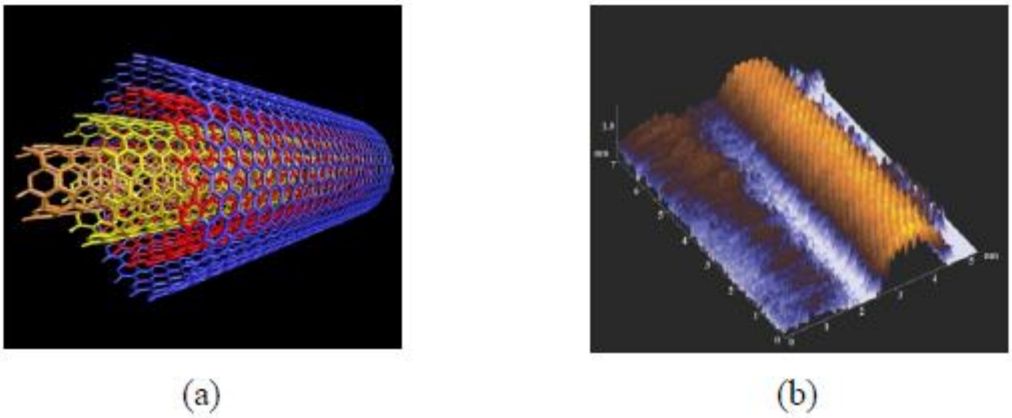

(a) (b)

Fig 1.2 . Carbon nanotube.

- The nanotubes can be single- and multi-walled (which are abbreviated as SWNT and MWNT, respectively). The reason for this is that the carbon nanotubes possess remarkable mechanical, electrical and thermal properties that equal, or even surpass, those of other benchmark materials such as steel, copper, and diamond, respectively. The Young’s modulus of the nanotubes can be as high as 1000 GPa (which is approximately five times higher than steel) while their tensile strength can be up to 63 GPa (around 50 times higher than steel). These

properties, coupled with their low density, give nanotubes huge potential in a range of structural applications. Applications on the nanometer and micrometer scale, such as SWNT-based transistors and chemical sensors are progressing rapidly.

The multi-walled carbon nanotubes consist of several nested coaxial singlewalled tubules. Typical outer and inner diameters of multi-walled carbon nanotube are 2–20 nm and 1–3 nm, respectively. The typical length is 1–100 µm. The intertubular distance is 0.34 nm.

1.17.1 GRAPHENE

Graphene is the basic structural element of carbon nanotubes. It is one-atom thick planar sheet of (*sp*2 bonded) carbon atoms that are densely packed in a honeycomb crystal lattice, as shown in Fig

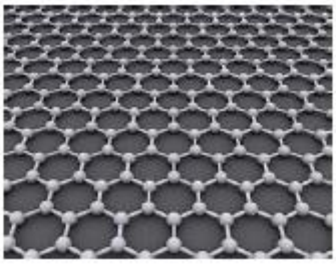


Fig.1.3 Graphene

- The definition of graphene is "a flat monolayer of carbon atoms tightly packed into a two-dimensional honeycomb lattice, which is a basic building block for graphitic materials of all other dimensionalities. It can be wrapped up into fullerenes, rolled into nanotubes or stacked into graphite".
- Graphene is stronger and stiffer than diamond. It, however, can be stretched like rubber. Its surface area is the largest known for its weight.
- The C–C bond length in graphene is ~0.142 nm. The graphene sheets stack to form graphite with an interplanar spacing of 0.335 nm, which means that a stack of $3x10^6$ sheets will be ~1 mm thick.

SWNTs and MWNTs

- The synthesis of C_{60} and other fullerenes stimulated intense interest in the tubular variant of the fullerenes. Iijima (1991) prepared multi-walled nanotubes as shown in Fig

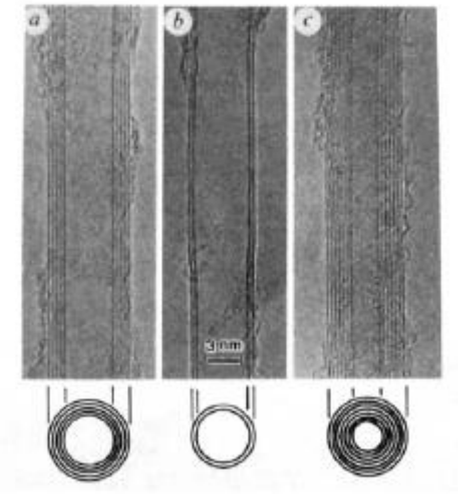


Fig 1.4 .: Electron micrographs of nanotubes of graphitic carbon. The parallel dark lines correspond to the (002) lattice images of graphite. A cross-section of each tubule is illustrated: (a) tube consisting of five graphitic sheets, diameter 6.7 nm, (b) two-sheet tube, diameter 5.5 nm, and (c) seven-sheet tube, diameter 6.5 nm, which has the smallest hollow diameter of 2.2 nm

These nanotubes were grown on the negative end of the carbon electrode used in the dc arc-discharge evaporation of carbon in an argon-filled vessel at 13.3 kPa pressure.

- The tubes grew on certain regions of the electrode. They were made of a single curved sheet of graphite connected at its edge. The electrical energy and the energy associated with the dangling bonds in a graphite sheet sealed the rolled-up sheets in the form of nanotubes.
- Electron microscopy revealed that each needle comprised coaxial tubes of graphitic sheets, ranging in number from 2 up to about 50. On each tube, the hexagons of carbon atom were arranged in a helical fashion about the needle axis. The helical pitch varied from needle to needle and from tube to tube within a single needle. The diameter of the tubes varied between 4 nm and 30 nm, and their length was up to 1 μm.
- The single-walled nanotubes were synthesized in 1993 by Iijima and Ichihashi, and Bethune *et al*. Iijima and Ichihashi used two vertical electrodes in the carbonarc chamber. The anode was a graphitic carbon rod and the cathode was a carbon rod. The latter held a small piece of iron. The chamber was filled with a mixture of methane and argon at low pressure. The carbon discharge arc was generated by running a dc current of 200 A at 20 V between the electrodes. The iron melted and formed a droplet. The iron vapor cooled and condensed into small particles of iron carbide on the electrode. Iron acted as catalyst in the vapor phase in the formation of the single-walled nanotubes. The nanotubes were

found in the sootlike deposits. The diameter of the tubes varied between 0.7 nm and 1.6 nm.

- Bethune *et al.* (1993) synthesized single-walled carbon nanotubes by covaporizing carbon and cobalt in an arc generator. The tubes had about 1.2 nm diameter. The tubes formed a web-like deposit woven through the fullerenecontaining soot, giving it a rubbery texture.
- Thess *et al*. (1996) presented an efficient route for the synthesis of arrays of single-walled nanotubes. Their process involved condensation of a laservaporized carbon⍰nickel⍰cobalt mixture at 1473 K. The yield was more than 70%. X-ray diffraction and electron microscopy showed that these single-walled nanotubes were almost uniform in diameter and they self-organized into 'ropes' which consisted of 100 to 500 nanotubes in a two-dimensional triangular lattice.

The ropes were metallic with a single-rope resistivity of less than 10⍰6 ⍰ m at
300 K.

- The high-pressure-carbon-monoxide (HiPCO) method developed by Nikolaev *et al.* (1999) can produce large quantities of carbon nanotubes. The catalyst for the growth of single-walled nanotube form *in situ* by thermal decomposition of iron pentacarbonyl in a heated flow of carbon monoxide at pressures in the range of 101.325 to 1013.25 kPa, and temperatures between 1073 and 1473 K. The yield of SWNT and the diameter of the nanotubes produced by this process can vary over a wide range, which is determined by the condition and geometry of the flow-cell.
- Nikolaev *et al*. (1999) produced SWNTs of 0.7 nm diameter. The products of thermal decomposition of Fe(CO)5 reacted to produce iron clusters in the gas phase. These clusters acted as nuclei upon which the SWNTs nucleated and grew. The solid carbon was formed through disproportionation of CO (known as the *Boudouard reaction*).

$$CO + CO \rightarrow C(s) + CO_2$$

- This reaction occurred catalytically on the surface of the iron particles. The iron particles promoted the formation of the tube's characteristic graphitic carbon lattice. The rate at which the reactant gases were heated determined the amount and quality of the SWNTs produced. The temperature and pressure had important effects on the yield of the nanotubes. TEM image of the single-walled nanotube produced by this method is shown in Fig.

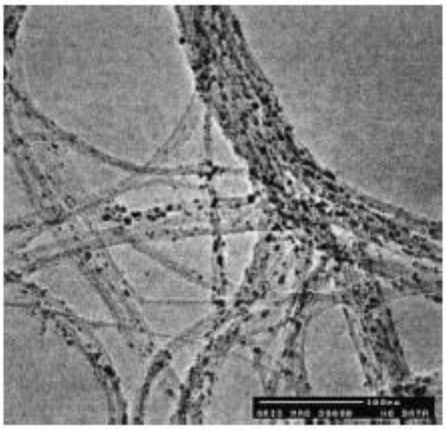


Fig 1.5 : . High-magnification TEM image of single-walled nanotube produced by the HiPCO process

- Ericson *et al*. (2004) produced well-aligned macroscopic fibers of single-walled carbon nanotubes from the concentrated dispersions of SWNTs in 102% sulphuric acid employing a wet-spinning technique. Because of the high-temperature stability of the SWNTs, melt-spinning is not a suitable option and wet-spinning is the only viable approach. The main challenge to the production of neat SWNT fibers is dispersing the SWNTs at high enough concentrations suitable for efficient alignment and effective coagulation.
- However, due to their chemical inertness and strong van der Waals attractions, SWNTs aggregate into ropes with limited solubility in aqueous, organic or acidic media. If a surfactant is used to disperse the SWNTs, there are complications of removing the surfactant from the fiber during coagulation or after processing. In sulfuric acid of concentration greater than 100%, SWNTs form charge-transfer complexes of individual positively charged nanotubes surrounded by a finite number of sulfuric acid anions.
- At very low concentrations, such charged tube-anion complexes behave as Brownian rods. At concentrations greater than 0.03 wt. %, a small amount of dissolved individual tubes coexists with a SWNT spaghetti-phase consisting of seemingly endless swollen ropes of well-aligned positively charged SWNTs intercalated by sulfuric acid anions.
- The SWNTs in the spaghetti are mobile and at a high enough concentration (> 4 wt. %), they coagulate and form ordered domains behaving similarly to nematic liquid crystalline rod-like polymers. The SWNT-acid system is very sensitive to water, and if a very small amount of moisture enters the system, phase separation occurs and discrete needle-like crystal-solvates precipitate. This ordered SWNT dispersion can be extruded and

coagulated in a controlled fashion using conventional fiber-spinning techniques to produce continuous lengths of macroscopic neat SWNT fibers.

- The scanning electron micrographs depicting the aligned macroscopic fibers consisting solely of the single-walled nanotubes are shown in Fig. 1.7 . The morphology of the fibers was strongly dependent on the coagulation conditions.

1.17.2 STRUCTURE OF CARBON NANOTUBES

In the theoretical analysis of carbon nanotubes, the major focus has been on single-walled tubes, cylindrical in shape with caps at each end, such that the two caps can be joined together to form a fullerene. The cylindrical portions of the tubes consist of a single graphene sheet (a single layer of carbon atoms from a three dimensional graphite crystal) that is shaped to form the cylinder.

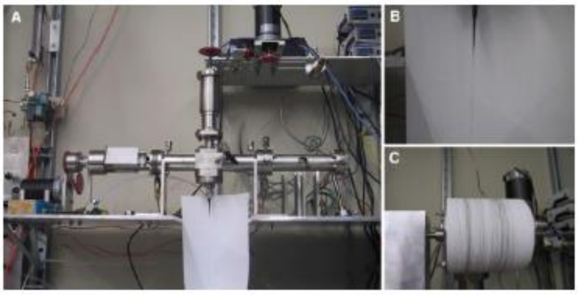


Fig 1.6: The spinning process of single-walled carbon nanotubes (SWNTs) in 102% sulfuric acid: (A) the apparatus for mixing and extruding neat fibers, (B) a jet of SWNT dispersion being extruded from a capillary tube, and (C) a 30 m spool of water-coagulated fiber

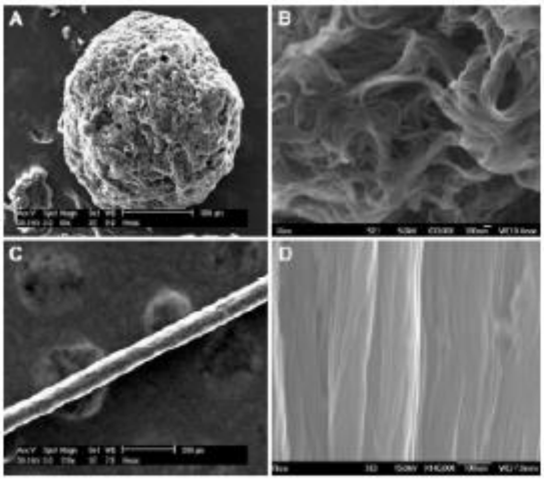


Fig. 1.7: SEM images showing the evolution of purified SWNTs into continuous fiber: (A) SWNTs after the purification process, (B) view inside the purified SWNTs showing the tangled mass of SWNT ropes that are 20–30 nm in diameter, (C) an annealed neat SWNT fiber spun from 8 wt.% dispersion in 102% sulphuric acid and coagulated in water, and (D) higher magnification of the neat fiber surface showing that all the ropes have merged into aligned super-ropes that are 200 nm or larger in diameter

There are two possible high-symmetry structures for nanotubes, known as 'zigzag' and 'armchair'. In practice, it is believed that most nanotubes do not have these highly symmetric forms but have structures in which the hexagons are arranged helically around the tube axis. These structures are generally known as 'chiral', because they can exist in two mirror-related forms.

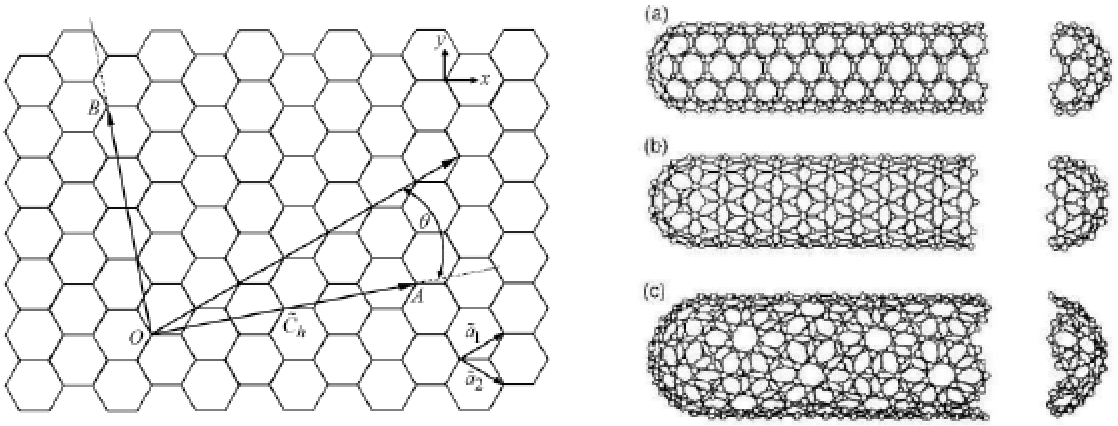


Fig 1.8: Schematic model for single-walled carbon nanotubes with the tubeaxis, *OB*, normal to: (a) θ = π/6 rad direction (an 'armchair' nanotube), (b) θ = 0 direction (a 'zigzag' nanotube), and (c) 0 < θ < π/6 rad (a 'chiral' nanotube)

It is convenient to describe a carbon nanotube in terms of the chiral vector, $\widetilde{C_h}$, and the chiral angle, θ , as shown in this Fig. 1.8. The points, *O*, and A, are crystallographically equivalent on the graphene sheet. These points are connected by the chiral vector, $\widetilde{C_h}$, which is defined in terms of the basis vectors, $\widetilde{a_1}$ and $\widetilde{a_2}$, of the honeycomb lattice.

$$\widetilde{C_h} = n\widetilde{a_1} + \mathrm{m}\widetilde{a_2}$$

where *n* and *m* are integers. The chiral angle, θ , is shown in the figure. *OB* represents the normal to $\widetilde{C_h}$ at point *O*. Depending on the value of the chiral angle, a single-walled carbon nanotube can have three basic geometries: armchair, zigzag and chiral, as illustrated in the Fig 1.8

- The diameter of a carbon nanotube is given by

$$d_t = \frac{|\widetilde{C_h}|}{\pi}$$

The integers *n* and *m* uniquely determine *dt* and θ

- The circumference of the nanotube is given by

$$L = |\widetilde{C_h}| = a(n^2 + m^2 + nm)^{\frac{1}{2}} \qquad 0 \le |m| \le n$$

where *a* is the length of unit vector

Problem 1. Calculate the diameter of a single-walled nanotube if n = m = 5 and the length of the unit vector is 0.246 nm.

$$L = |\widetilde{C_h}| = a(n^2 + m^2 + nm)^{\frac{1}{2}}$$

$$L = .246 \times \sqrt{5^2 + 5^2 + 5 \times 5} = 2.13\text{nm}$$

$$d_t = \frac{|\widetilde{C_h}|}{\pi} = \frac{2.13}{\pi} = 0.678nm$$

- The carbon nanotubes can either have metallic conductivity or can be semiconducting depending on the tube diameter and the chiral angle. Metallic conduction in carbon nanotubes can be achieved without the introduction of doping or defects.
- Ballistic electron transport and electrical conductivity without phonon and surface scattering have been observed in metallic carbon nanotubes. Semiconducting nanotubes have a band gap between their conduction and valence bands, which is proportional to $1/\,dt$.
- It is possible to fabricate a transistor based on a single helical carbon nanotube. This is a step towards molecular electronics

An AFM image of a single nanotube contacting three platinum electrodes is shown in Fig

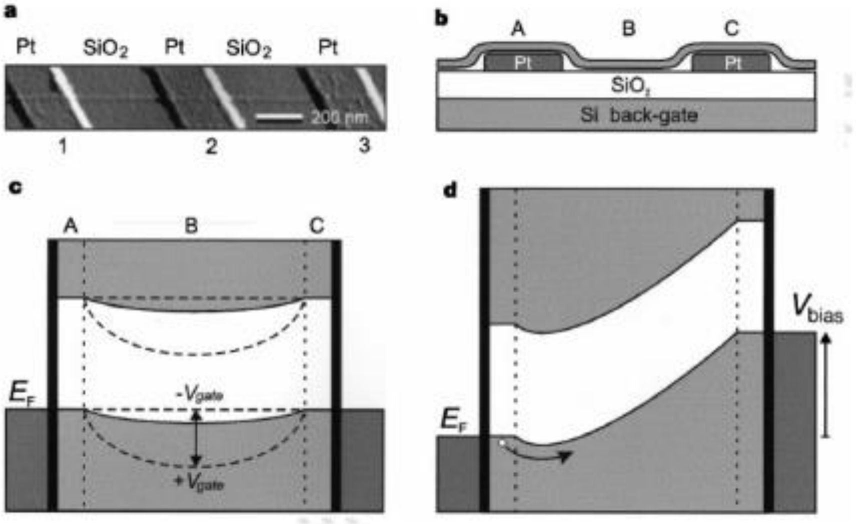


Fig. 1.9 (a) Tapping-mode AFM image of an individual carbon nanotube on top of three Pt electrodes, (b) Schematic side view of the TUBEFET (single carbon nanotube field-effect transistor) device. A single semiconducting nanotube is contacted by two electrodes. The Si substrate, which is covered by a layer of SiO2 300 nm thick, acts as a back-gate, (c) Suggested band diagram of the device. The nanotube with a gap of ~0.6 eV is connected to the leads with Fermi energy *EF* by tunnelling contacts, indicated by the black vertical bars. At A and C (see b), the valence-band edge is pinned to the Fermi energy of the leads. Owing to a difference in work function between the tube and the electrodes, the bands bend towards lower energy in between the electrodes (B). For positive *V*gate the bands bend more strongly, leading to an insulating state. For negative *V*gate the bands flatten, resulting in a metal-like conductance, and (d), Application of a bias voltage results in a suppression of the barrier

The semiconductor Si-substrate covered with a 300 nm layer of thermally grown SiO2 was used as a back-gate. The metallic variety of tubes had linear current I versus voltage (Vbias) curves, and showed no dependence on the gate voltage (Vgate). For the sample shown in the figure, the I versus Vbias curve was slightly nonlinear for Vgate = 0 . When Vgate was increased to positive values, a pronounced gap-like nonlinearity developed around Vbias = 0 .

- The curves exhibited a power-law behavior: $I \propto (V_{bias})^{\propto}$ with $\propto$ lying between 1 and 12. Upon application of a negative Vgate , the I versus Vbias curve became linear with a resistance that saturated around 1 MΩ. This resistance is mainly due to the contact resistance between the tube and the electrodes. Therefore, the device showed controllable semiconductor-to-metal transition in a one dimensional system. The nonlinearity at room temperature and the asymmetric dependence of the conductance on the gate-voltage polarity indicate that the nanotube is semiconducting. The conductance could be modulated by about six orders of magnitude if Vgate was changed by 10 V.
- Both nanowires and single-walled nanotubes can be used to make molecular field-effect transistors. Advanced electronic devices based on carbon nanotubes and various types of nanowires can have a very important role in next-generation semiconductor architectures. However, at present there is lack of a general fabrication method which has held back the development of these devices for practical applications.
- A few assembling strategies have been suggested for devices based on these nanomaterials. Lee *et al*. (2006) used inert surface molecular patterns to direct the adsorption and alignment of nanotubes and nanowires on bare surfaces to form device structures without the use of linker molecules. They have used this method to demonstrate large scale assembly of nanotube- and nanowire-based integrated devices and their applications.

1.17.3 CARBON NANOTUBE COMPOSITE MATERIALS

- Many of the outstanding properties of carbon nanotubes can be best utilized by incorporating the nanotubes into some form of matrix. The exceptional mechanical properties, in particular, have prompted huge interest in the production of composite materials containing nanotubes for structural applications using polymers, ceramics and metals.

- The materials which form composites with carbon nanotubes are mostly polymers. The simplest method for preparing nanotube�polymer composites is mixing the nanotube dispersion with solution of the polymer, and then evaporate the solvent in a controlled way. This method has been used with various polymers such as polyvinyl alcohol, polystyrene, polycarbonate and poly(methyl methacrylate).
- In order to facilitate solubilization and mixing, the nanotubes are often functionalized prior to adding to the polymer solution. For example, acid treatment is used to disperse catalytically produced MWNTs in water, and then nanotube–polymer composite is made by simply mixing one of these dispersions with an aqueous solution of the polymer and casting the mixture as film.
- The solution mixing approach is limited to polymers that freely dissolve in common solvents. An alternative is to use thermoplastic polymers and apply melt processing techniques. One problem associated with this method is that achieving homogeneous dispersions of nanotubes in melts is generally more difficult than with solutions, and high concentrations of tubes are hard to achieve due to the high viscosities of the mixtures.
- An alternative method for preparing nanotube polymer composites is to use the monomer rather than the polymer as a starting material, and then carry out *in situ* polymerization.
- Ceramics have high stiffness and thermal stabilities but relatively low breaking strengths. Incorporating carbon nanotubes into a ceramic matrix produces a composite with both toughness and high-temperature stability. However, achieving a homogeneous dispersion of tubes in an oxide, with strong bonding between tubes and matrix, presents rather more of a challenge than incorporating tubes into a
- Nanotubes have been used to improve the anti-static properties of fuel-handling components and body panels of automobiles. Nanotube-containing sporting equipment (e.g., tennis rackets and high-performance racing bicycles) has been manufactured recently
- It has been shown that considerable improvement in the performance of Li-ion batteries can be achieved by the addition of catalytically-produced MWNTs.

Carbon nanotube reactor

- Carbon nanotubes (CNTs) have well-defined hollow interiors and exhibit exceptionally good mechanical and thermal stability as well as electron conductivity. This opens up possibilities to introduce other matter into the cavities, which may lead to nanocomposite materials with interesting properties or behavior different from the bulk.
- Pan *et al.* (2007) have reported enhancement of the catalytic activity of Rhodium particles confined inside nanotubes for the conversion of CO and H2 to ethanol. The overall formation rate of ethanol inside the nanotubes exceeds that on the outside of the nanotubes by more than an order of magnitude, although the latter is much more accessible.
- The production of ethanol inside a carbon nanotube reactor containing catalytic particles is schematically shown in Fig

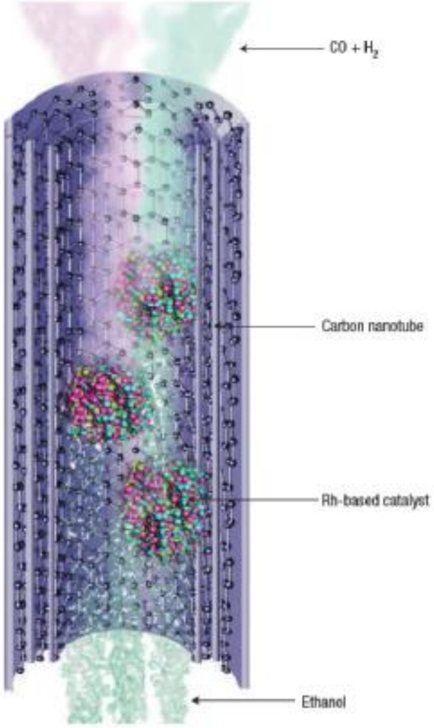


Fig 1.10 : **Schematic diagram showing ethanol production from syngas inside Rh-loaded carbon nanotubes. The black spheres denote carbon atoms, which form the graphene layers of the carbon nanotubes. The streams in light orange and green entering the nanotubes indicate the gas mixture of CO and H2, respectively. The three stacks of small spheres in rose, blue, green and red inside the tubes represent catalyst particles that may comprise more than one component. The streams in light cyan tailing behind the catalyst particles along the axis of the nanotubes represent ethanol**

MODULE II
FABRICATION OF NANO-LAYERS & NANO- PARTICLES

2.1 PRODUCTION OF NANO-LAYERS

2.1.1 Physical Vapor Deposition:

In general, physical vapor deposition (PVD) from the gas phase is subdivided into four groups.

- Evaporation,
- Sputtering,
- Ion Plating,
- Laser Ablation.

The first three methods occur at low pressures.

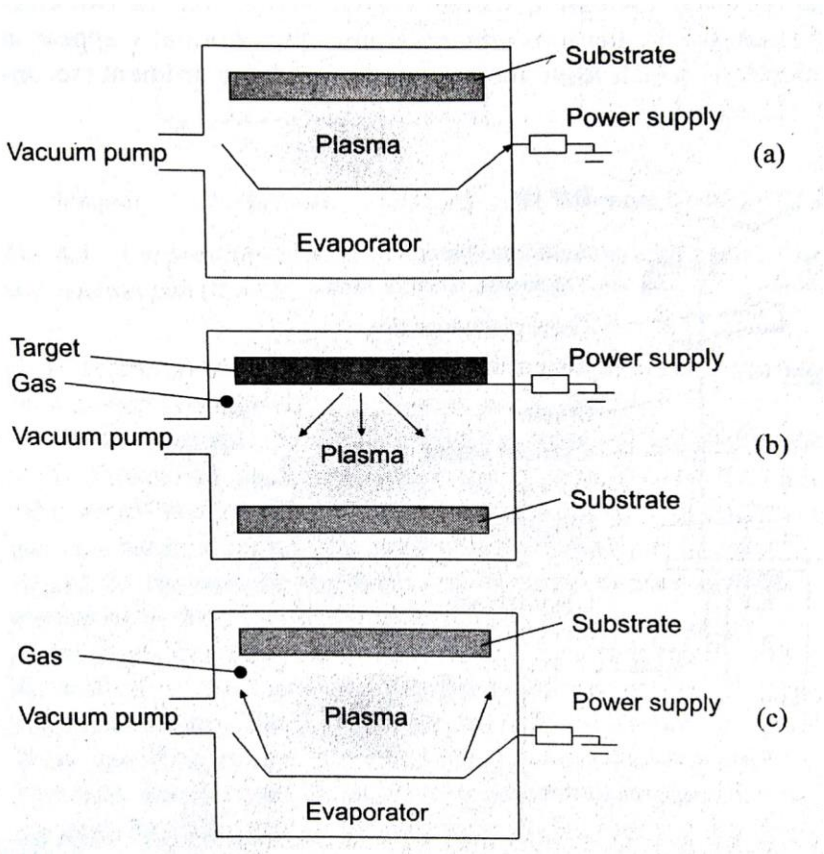


Fig:2.1 Three fundamental PVD (a).Evaporation (b). Sputtering (c.).Ion Implanting

1. Evaporation.

This procedure is carried out in a bell jar as depicted in Fig. 2.2

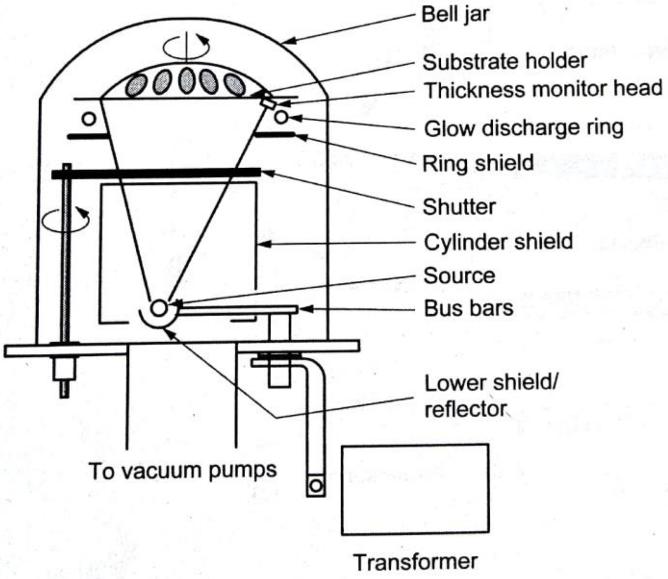
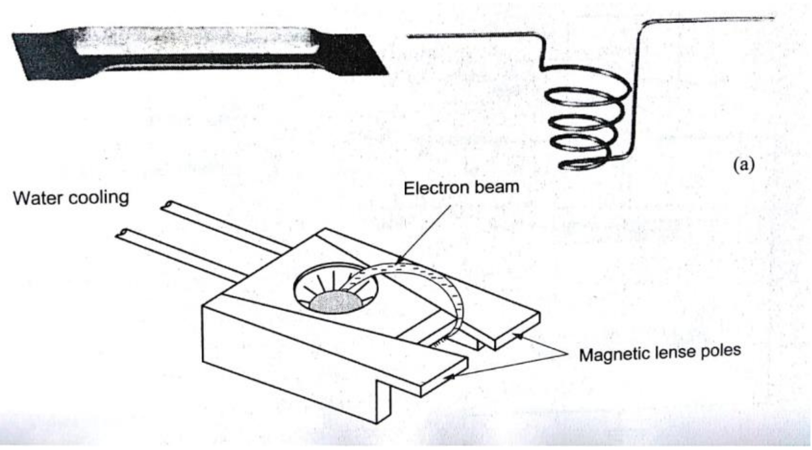
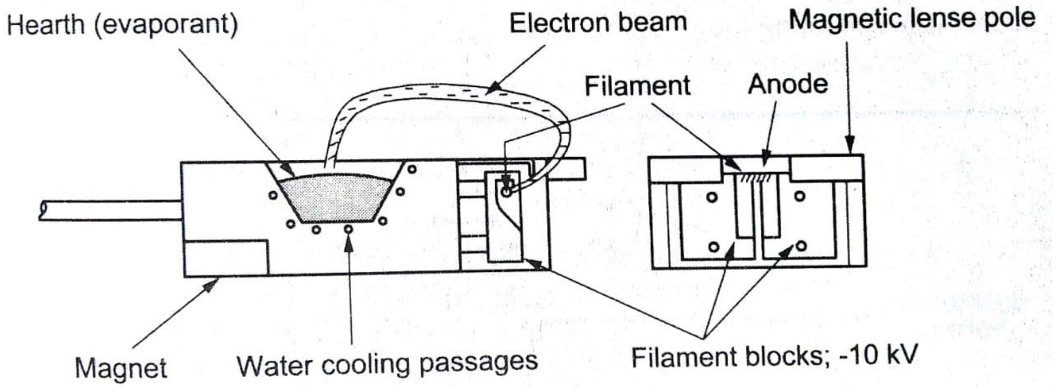


Fig 2.2 Vacuum system for the vaporization from resistance –heated sources.

A crucible is heated up by a resistance or an electron gun until a sufficient vapor pressure develops. As a result, material is deposited on the substrate. Technically, the resistance is wrapped around the crucible, or a metal wire is heated up by a current and vaporized. The electron gun (e-gun) produces an electron beam of, e.g., 10 keV. This beam is directed at the material intended for the deposition on the substrate. The gun's advantage is its unlimited supply of evaporating material and applicability of non-conductive or high-melting materials. Its shortcomings lie in the production of radiation defects, for instance in the underlying oxide coating. Both procedures are more precisely depicted in Fig. 2.3



2.3 (a). Evaporation by means of resistance-heating with a tungsten boat & winding



2.3 (b). Electron Gun

(II) SPUTTERING.

In literature, there is no clear definition of the term sputtering. Generally, an atom or a molecule, usually in its ionized form, hits a solid state (target) and knocks out surface atoms. This erosion is accompanied by a second process, namely the deposition of the knocked out atoms on a second solid state (substrate). The latter process is relevant when forming thin layers.

(a) Glow discharge. In its simplest form, sputtering is achieved by glow discharge with dc voltage. A cross section of the arrangement is schematically represented in Fig. 2.4.

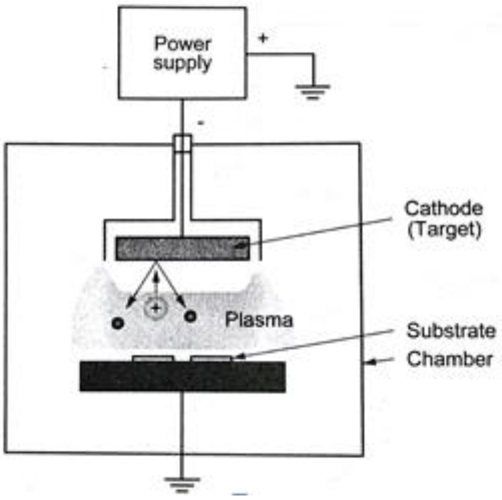


Fig: 2.4 : DC Voltage Sputtering

After mounting the samples on a holder, the chamber is rinsed repeatedly with Ar. Eventually, a constant gas pressure of some 100 mPa is built up. The target, being attached a few centimeters above the substrate, is raised to a negative dc potential from -500 to -5000 V, while both chamber and substrate are grounded. The discharge current requires a conducting target.

When the voltage is slowly increased, a small current flows over the two electrodes. This current is caused by the ions and electrons which normally appear in the gas and by the electrons which leave the target after ion bombardment (secondary electrons). At a certain voltage value, these contributions rise drastically. The final current-voltage curve is shown in Fig. 2.5.

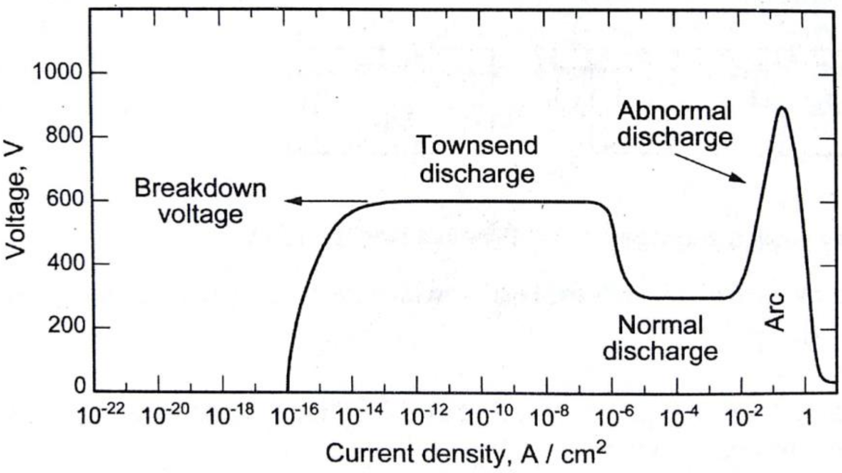


Fig 2.5: Applied Voltage *vs* Discharge current

The first plateau (at 600 V in our example) of the discharge current is referred to as Townsend discharge. Later the plasma passes through the "normal" and "abnormal" ranges. The latter is the operating state of sputtering. A self-contained gas discharge requires the production of sufficient secondary electrons by the impact of the ions 9n the target surface and .conversely, the production of sufficient ions in the plasma by the secondary electrons.

(b) High frequency discharge. When replacing the dc voltage source from Fig. 2.4 with a high frequency generator (radio frequency, RF, generator), target and substrates erode alternately depending on the respective polarity.

But even with these low frequencies, a serious shortcoming becomes apparent: due to the substantially small target surface (compared to the backplate electrode consisting of the bell, the cable shield, etc.) a proportionally large ion current flows if this

backplate electrode is negatively polarized. This would mean that the substrates are covered with the material of the bell, which is not intended.
In order to overcome this shortcoming, a capacitor is connected in series between the high frequency generator and the target, and/or the conducting target is replaced with an insulating one. During the positive voltage phase of the RF signal, the electrons from the discharge space are attracted to the target. They impact on the target and charge it; current flow to the RF generator is prevented by the capacitor. During' the negative half-wave of the RF signal, the electrons cannot leave the target due to the work function of the target material. Thus, the electron charge on the target remains constant.

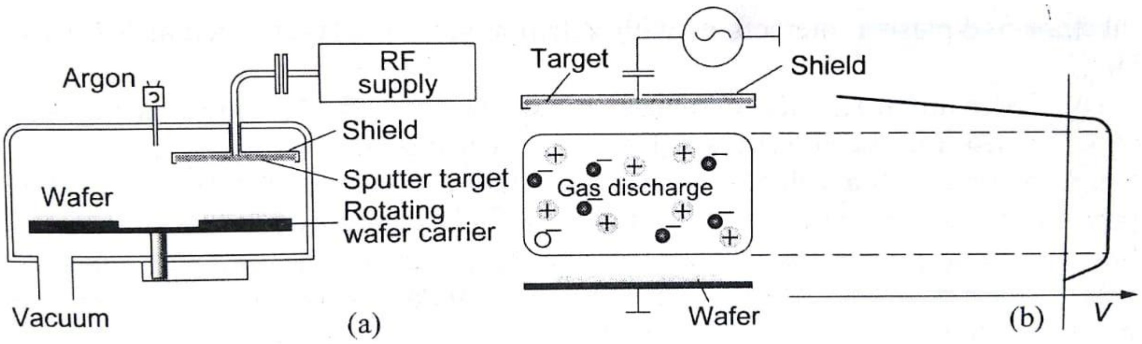


Fig: 2.6 (a) RF sputter System (b) Distribution of potential in an RF Plasma

Due to their mass, the positively charged ions are not capable of following the RF signal with frequencies above 50 kHz. Therefore, the ions are only subjected to the average electrical field which is caused by the electron charge accumulated on the target. Depending on the RF power at the target, the captured charge leads to a bias of 1 000 V or mere and causes an ion energy within the range of 1 keV.

When using a capacitively coupled target, the limitations of the glow discharge can be overcome, i.e., a conducting target is no longer required. Therefore, the number of layers which can be deposited by sputtering is greatly increased.

(III) ION PLATING.

This process is classed between resistance evaporation and glow discharge. A negative voltage is applied to the substrate, while the anode is connected with the source of the metal vaporization. The chamber is subsequently filled with Ar with a pressure of a few Pa, and the plasma is ignited. After cleaning the wafer by sputtering, the e-gun is switched on and the material is vaporized. The growing of the layer on the substrate is improved by the plasma in some properties such as adhesion and homogeneity compared to a sole PVD.

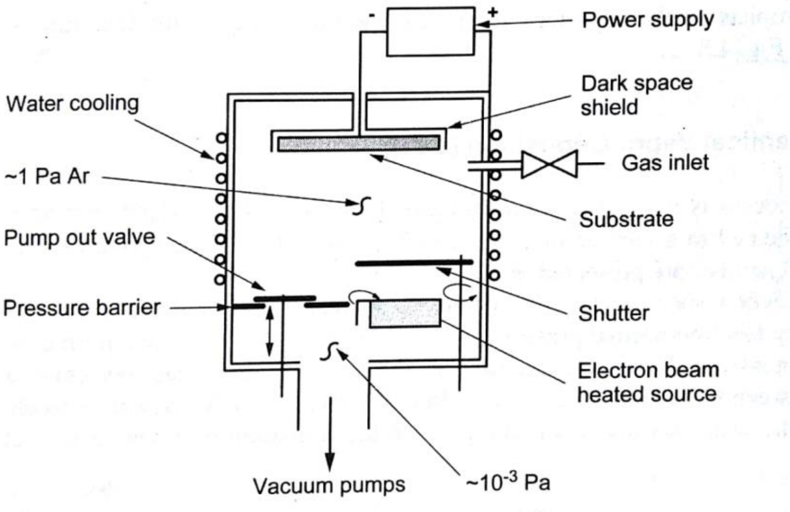


Fig 2.7: Ion Plating system

The advantages of ion plating are higher energies of the vaporized atoms and therefore better adhesion of the produced films. The disadvantage is heating of the substrate and plasma interactions with radiation-sensitive layers such as MOS oxides.

(IV) LASER ABLATION.

The following process data are typical values. A high-energy focused laser beam (100 m], l J / cm^2) is capable of eroding the surface of a target rotating with a velocity of one revolution per second. The material is vaporized on the substrate, and as a result, a film is produced on it at a rate of 0.07 nm/ laser pulse. The growth can be supported by heating the substrate (750 °C) and by chemical reactions (oxygen at 50 Pa). So far, the used lasers are excimer, Nd:YAG, ruby, and $C0_2$ lasers.

Advantages of laser ablation are the deposition of materials of high-melting points, a good control over impurities, the possibility of the vaporization in oxidizing environments, and stoichiometric vaporization. A shortcoming is the formation of droplets on the vaporized layer. A system described in the literature is presented in Fig. 2.8.

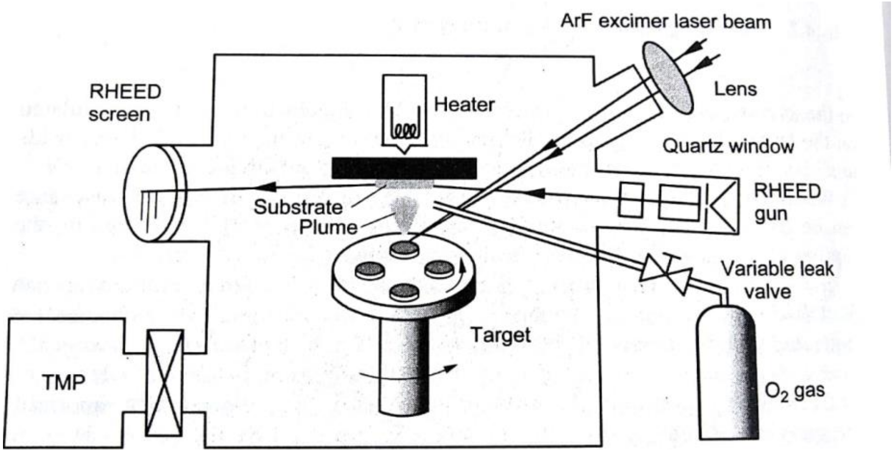


Fig 2.8 : Laser Ablation system under O_2 partial pressure

2.1.2 CHEMICAL VAPOR DEPOSITION (CVD)

The CVD process is performed in an evacuated chamber. The wafer is put on a carrier and heated to a temperature between 350 and 800 °C. Four possible versions of the chamber are presented in Fig. 2.9

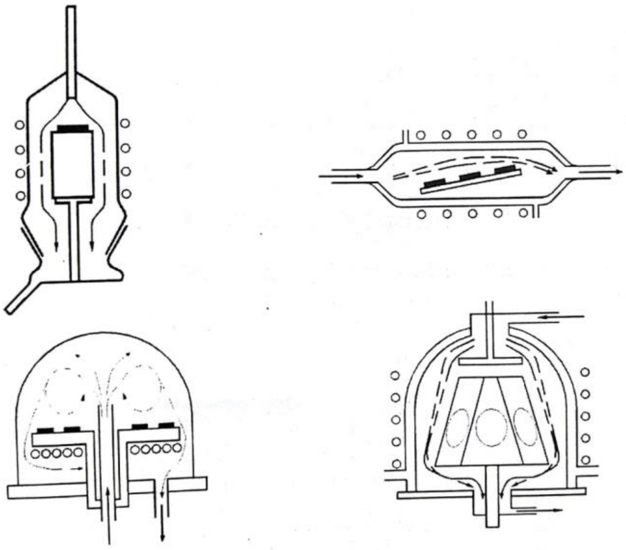


Fig 2.9 : Four versions of CVD chamber

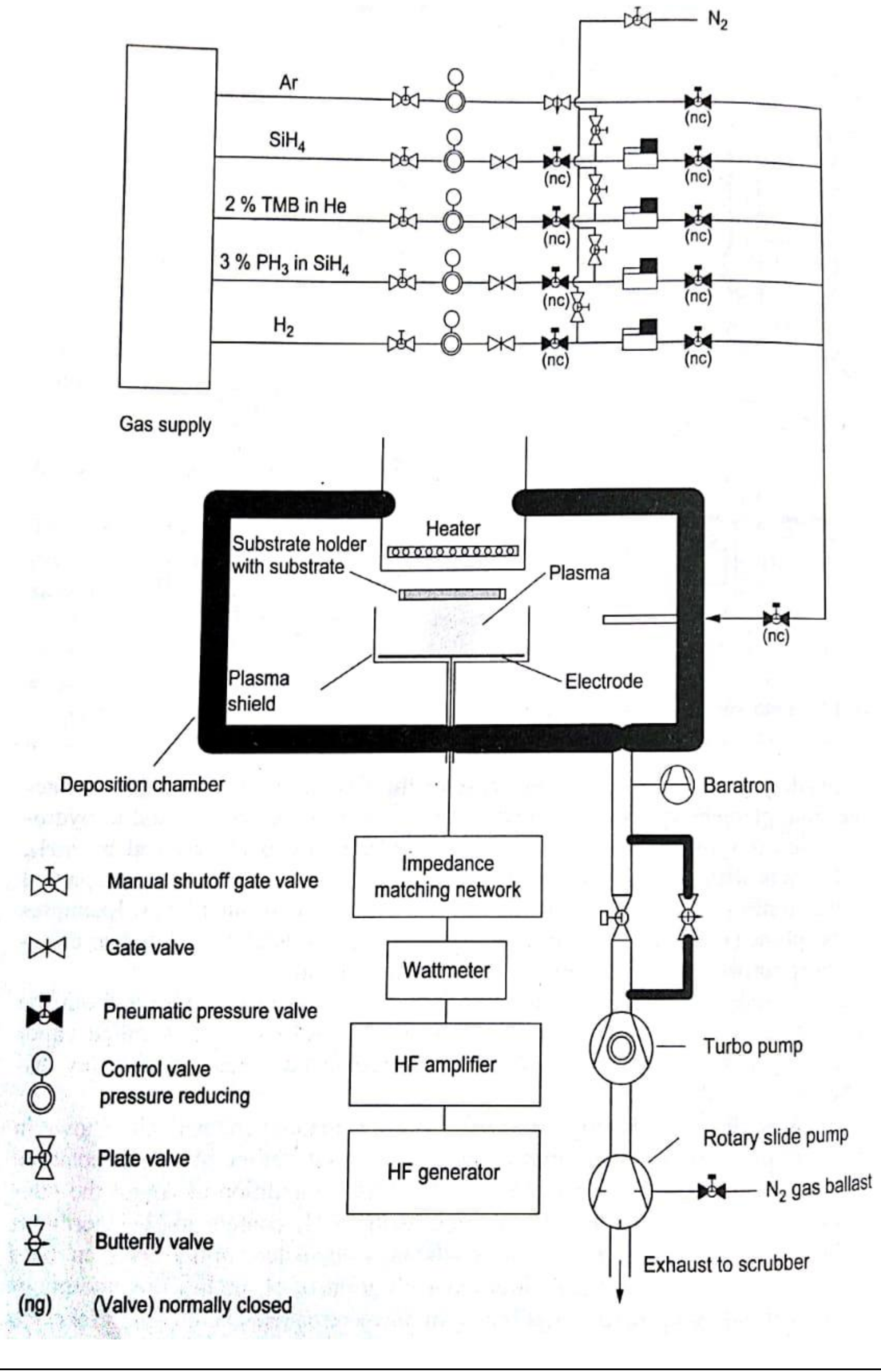


Fig 2.10: Block Diagram of PECVD

One or several species of gases are let in so that a gas pressure is formed between very low and normal pressure. The gas flow hits the wafer at a normal or a glancing incidence. Now a dissociation (in the case of a single gas species) or a reaction between two species takes place. In both cases, a newly formed molecule adheres to the wafer surface and participates in the formation of a new layer.

Let us consider silane (SiH4) as an example of the first case. On impact, it disintegrates into elementary silicon, which partly adheres to the surface, and to hydrogen, which is removed by the pumps. The second case is represented by SiH4, which reacts with N_20 to form SiO_2. The process can of course be accompanied by other types of gases which act as impurities in the deposited layer. Examples are phosphine (PH_3) or diborane (B_2H_6), which also disintegrate and deliver effective phosphorus or boron doping of the deposited silicon.

CVD is a layer structure without the continuation of the underlying lattice. The reverse case is called vapor phase epitaxy.

CVD deposition can be supported by an RF plasma, as schematically shown in Fig. 2.10, an example of an amorphous or micro-crystalline silicon deposition. The major difference to the conventional CVD is the addition of Ar for the ignition of the plasma and of H_2. The degree of the SiH_4 content in 11; determines whether amorphous or microcrystalline silicon is deposited. In the first step, both types are deposited. However, a high concentration of H_2 etches the amorphous portion, and only the microcrystalline component remains. The etching process is even more favoured if higher frequencies (e.g., 110 MHz) other than the usual 13.56 MHz are used. In Fig. 2.11, a typical PECVD system is depicted.

2.1.3 EPITAXY

We are dealing with epitaxy if a layer is deposited on a (crystalline) substrate in such a way that the layer is also monocrystalline. The layer is often referred to as film. In many cases, the film takes 99.9 % of the entire solid state, as in the example of a Czochralski crystal, which is pulled from a narrow seed nucleus. If film and substrate are from the same material, we are dealing with homoepitaxy (e.g., silicon-on-silicon), otherwise with heteroepitaxy (e.g., silicon-on-sapphire). Another distinction is made by the phase from which the film is made: vapor phase epitaxy, liquid phase epitaxy (LPE), and solid state epitaxy. A subclass of vapor phase epitaxy is molecular beam epitaxy (MBE).

The setup of a molecular beam epitaxy (MBE) is depicted in detail in Fig. 2.12.

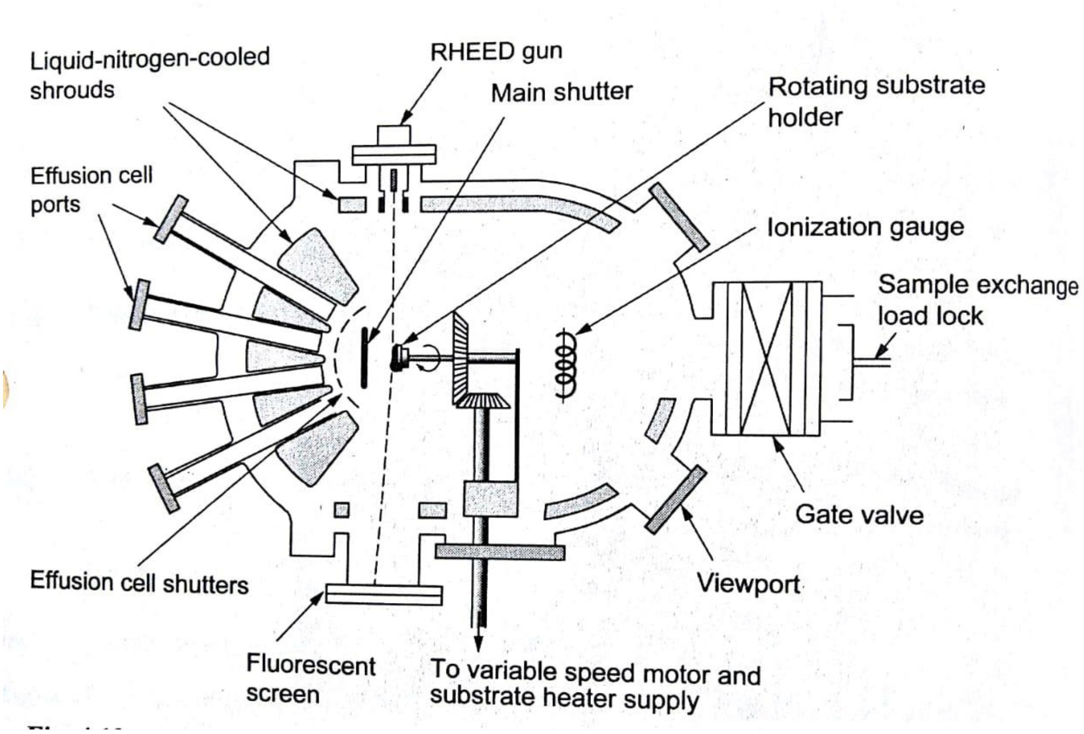


Fig 2.12 : Schematic Structure of MBE

The constituents of the deposited film are contained in mini furnaces as elements, the so-called Knudsen cells, which are discussed below. During heating some vapor pressure develops and an atom beam is emitted, which is bundled by successive apertures. The beam hits the wafer surface to which the atoms remain partially adhered. There, they can react with atoms of a second or third beam, which is also directed towards the wafer surface. A favourable reaction and finally the film deposition depend on the selection of the parameters, i.e., wafer temperature, the ratios of the beam densities, the purity of the surface, etc. As shown in the same figure, the effusion cell can be replaced by an evaporation with the electron gun. The chamber contains many devices for the in situ inspection of the growing layers, for example low energy electron diffraction (LEED), secondary ion mass spectroscopy (SIMS), and Auger and Raman spectroscopy. The quality of the vacuum is controlled by a residual gas analyzer. The effusion (Knudsen) cell is seen in detail in F ig. 2.13.

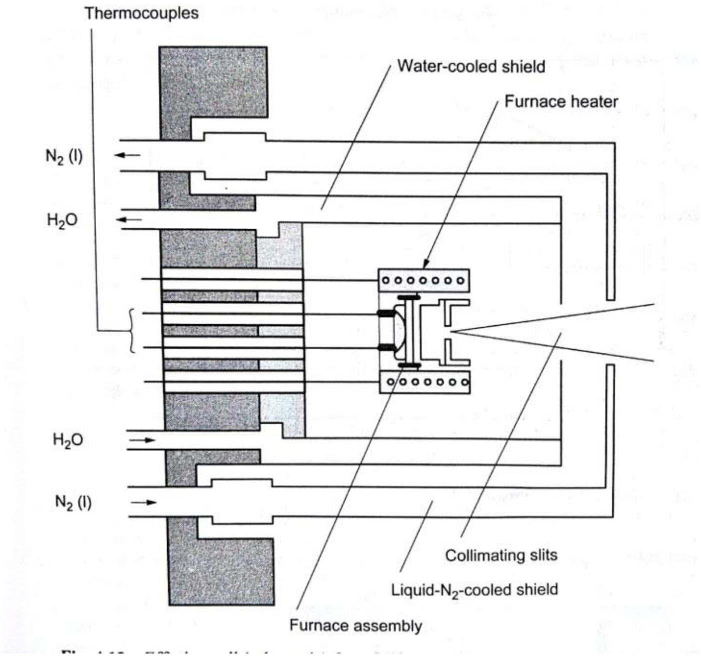


Fig 2.13 : Effusion Cell

The material to be deposited is contained in the innermost cell which is heated up. Its temperature is controlled by a set of thermocouples and resistance heaters. Without mrther measures, the high temperature leads to molecular desorption ' from all warmed up surfaces, to the emission of impurities into the substrate, and in the worst case, to the breakdown of the vacuum. Therefore, a screen cooled with liquid air is installed around the internal cell. Conversely, in order to avoid high thermal flows between furnace and screen, a water-cooled shield is inserted between them.

2.1.4 ION LMPIANTATION

Ion implantation is a doping technique with which ions are shot into a substrate (e.g., a silicon wafer) using an accelerator. The basic principle is presented in Fig. 2.16.

The desired ion species is let in as a gaseous compound through a needle valve (alternatively, solid state Sputtering sources are used). The compound is dissociated and ionized with an electron beam. The arising ions (including the unwanted ones) are pulled out of the source area and brought to an energy of 30 keV by a first, preliminary acceleration (all figures are typical values). Then, the ions pass through a magnetic field filter which is adjusted in such a way that only the desired ion type can run into the acceleration tube. The magnetic field filter is based on the fact that the Lorentz force for a moving charged particle compels a circular path. The radius of the circle depends on the magnetic field B, the velocity v, and the mass m of the particle. More exactly, it depends on m/e (e is the elementary charge). For the desired ion species, i.e., for a given m/e, the magnetic field is adjusted in such a way that the circular path of these particles terminates exactly at the end of the accelerator tube. There, the ions acquire a total energy of 360 keV. This energy can be doubled or multiplied by the use of double or multiple charged ions. However, the ion yield, i.e., the available ion current, is exponentially reduced with the state of charge (ionization state).

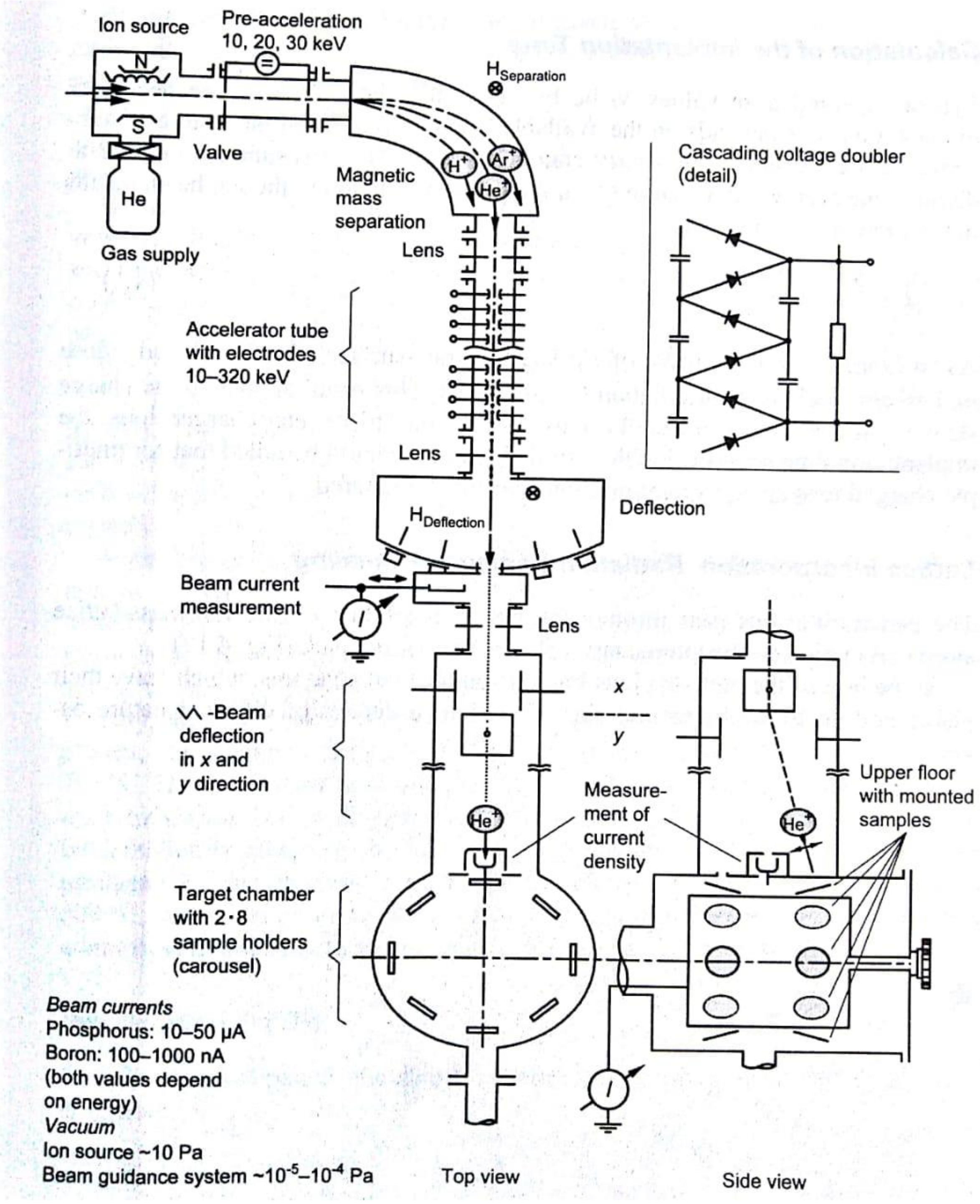


Fig 2.16: Ion implantation Equipment

The beam can be positioned by a combination of an aperture with a Faraday cage. With two capacitor disks each, the beam can be scanned upwards and downwards or left and right. In order to avoid Lissajou figures, horizontally and vertically incommensurable scanning frequencies are chosen. The beam current is measured by an ammeter, which is connected to the substrate holder isolated against ground. The {substrate holder is designed as a carousel, for example, in order to be able to implant several samples without an intermediate ventilation.

2.1.5 FORMATION OF SILICON OXIDE

Thin oxide layers are contained in almost all electronic devices. They appear as gate oxide in MOS transistors or MIS solar cells, field oxide for isolation purposes, antireflection layers in solar cells, or as passivation layers for long-term protection.

We begin with thermal oxidation. A Si wafer is cleaned so that any organic or heavy metal impurity on the surface is removed and the natural oxide is dissolved. Then, the wafer is inserted in a quartz tube heated to a temperature of about 1100 °C. A flow of an oxidizing gas, either pure oxygen (dry oxidation) or nitrogen driven through water (wet oxidation), is maintained. When oxygen penetrates into the substrate, the Si surface reacts with the oxygen and forms silicon dioxide. While this procedure sounds simple, it requires highest cleanliness, which is the critical step for the MOS production.Numerous scientific and technical investigations have focused on the properties of silicon dioxide. Fields of interest are the growth laws, deep levels, capture of charge carriers from the silicon, segregation and rearrangement of the dopant in the neighbouring silicon, masking properties against diffusion and ion implantation, etc. Their in-depth discussion is beyond the scope of this book:

Technical alternatives to thermal oxidation are CVD and PECVD of oxides. These are treated in the section on CVD. There are some technical CVD versions such as TEOS deposition shown in Fig. 2.17.

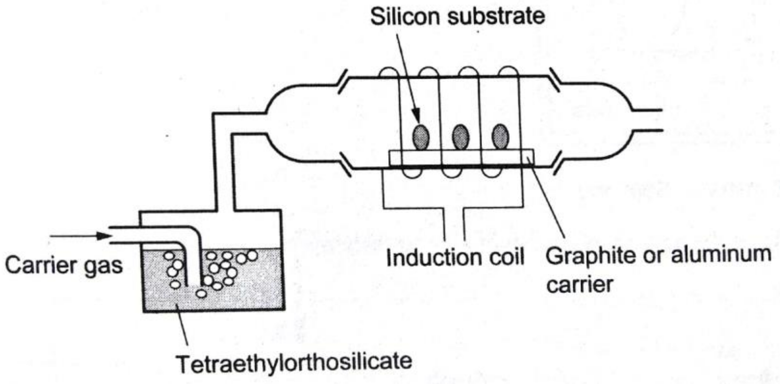


Fig 2.17: TEOS Process

A feed gas (usually nitrogen) is driven through a container filled with tetra ethylorthosilicate (TEOS). TEOS is a liquid at room temperature. Its chemical structure is presented in Fig. 2.18 The enriched nitrogen flows to the wafers where SiOz is deposited on their surfaces. The deposition is maintained at a temperature of about 650 to 850 °C by means of an external induction coil. It should be noted that the silicon in the TEOS is already oxidized, in contrast to the silane process.

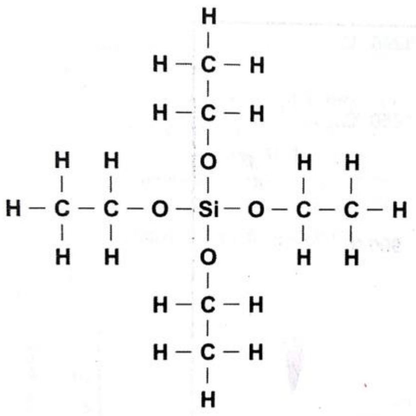


Fig 2.18: Structure of tetraethylorthosilicate (TEOS)

In this state, TEOS finds only limited application because the deposition temperature (>650 °C) prevents its use after metallization. In order to obtain lower deposition temperatures, the application of a more aggressive oxidant, i.e., ozone, is required. After adding some few molar per cent of ozone, the optimum deposition temperature is reduced to 400 °C.

The anodic oxidation is shown in Fig. 2.19. The wafer is immersed into a 0.04 M solution of KNO_3 in ethylene glycol with a small addition of water. Afier mounting it to a holder with a vacuum, it is positively charged, while a platinum disk acts as a backplate electrode.

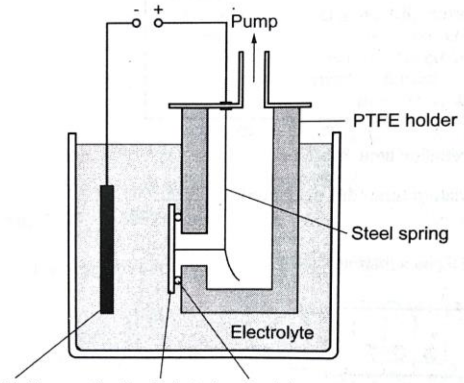


Fig 2.19: Assembly of the anodic oxidation of silicon

The current causes a reaction on the surface of silicon

$$Si + 2H_2O + 4h^+ \rightarrow SiO_2 + 4H^+$$

$$CH_2OH - CH_2OH \rightarrow 2HCHO + 2H^+ + 2e^-$$

$$H_2O \rightarrow \frac{1}{2}O_2 + 2H^+ + 2e^-$$

Technically, anodic oxidation is of little importance. The quality of the oxide is too low, and the process is time-consuming or not compatible with other applications.

Rarely, sol gels are used to manufacture oxide layers. During this process, a suspension of oxide particles in an organic solvent is distributed over the wafer. A centrifuge facilitates homogeneous distribution of the liquid on the wafer. Then, the wafer is baked and the solvent evaporates. The required temperatures range from 500 to 800 °C. The oxide can have the quality of a gate oxide, but no largescale application has been reported so far.

A further manufacturing process is SOI (silicon-on-oxide), i.e., the implantation of oxygen up to the stoichiometric dose and its reaction with silicon to form SiO_2. SOI has gained in importance because of better radiation resistance and heat distribution compared to conventional technologies.

2.2 FABRICATION OF NANOPARTICLES

Nanoparticle is defined as a ball or a ball-like molecule which consists of a few 10 to some 10,000 atoms interconnected by interatomic forces but with little or no relationship to a solid state. A first example is a nano-crystalline Si particle which is embedded into an amorphous matrix.

2.1.1 GRINDING WITH IRON BALLS

First of all, a container is filled with stainless steel balls of a few millimeters m diameter. The material to be crushed is added in the form of a powder of about 50 μm diameter grain size. After filling the container with liquid nitrogen, a rotating shaft grinds the material. The grinding periods are within the range of minutes to some 100 hours. This process is simple; its weakness, however, lies in the fact that the grinding balls contribute to impurities.

2.1.2 GAS CONDENSATION

A typical system is shown in Fig. 2.20. The operation occurs in an evacuated chamber with a pressure of 10-5 Pa. After mounting the raw material on one or more crucibles, it is evaporated thermally, by an electron gun, or by ion sputtering.

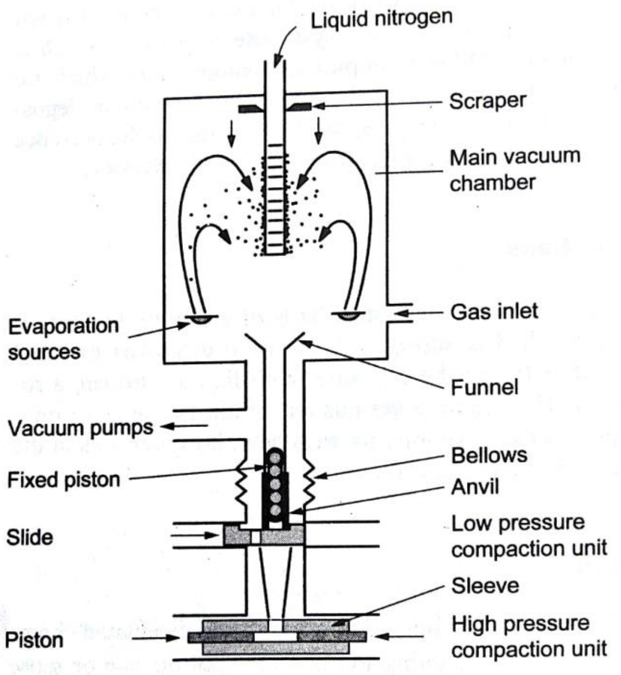


Fig 220: System for inert gas condensation of nanocrystalline power

The evaporated atoms or molecules unite and form particles of different sizes. Finally they are captured with a cold linger from which they are scraped off and collected with a funnel. The particle diameter is usually within the range of 5 to 15 nm.

2.2.3 LASER ABLATION

The raw material is provided as a solid. Its dissolution is achieved by a focused laser beam- similar to the cutting of a metal or a semiconductor. The advantage of this procedure is a 1:1 transfer of the material composition from the raw material to the particles. The system has already been shown in Fig. 2.8.

2.2.4 REDUCTION METHODS

Some metal compounds (e.g., chlorides) can be reduced to elementary metallic nanopowder by the application of $NaBEt_3H$, $LiBEt_3H$, and $NaBH_4$, for example. The reaction equation can be written as (M: metal, Et: ethyl):

$$MCl_x + xNaBEt_3H \rightarrow xNaCl + xBEt_3 + \frac{1}{2}xH_2$$

2.2.5 SOL GELS

A sol (hydrosol) is a colloidal dispersion in liquid. A gel is a jelly-like substance formed by coagulation of a sol into a gel.

The best known example of a so] gel process is probably the production of SiO_2. A catalyst (acid or base) is added to a solution of tetramethoxysilane (TMOS), water, and methanol. Hydrolysis of the Si-OMe (Me: methyl) bonds leads to the formation of Si-OH groups

$$Si(OMe)_4 + 4H_2O \rightarrow Si(OH)_4 + 4MeOH$$

Further dehydration reduces the “$Si(OH)_4$” to SiO_2; gel. If hydrolysis and condensation are completed,a silicon oxide xerogel is formed (in Greek, xerox means dry). During the reaction, the gel reaches a viscosity so low that it can be applied onto a centrifuge and distributed over the wafer. When annealing over 800 °C, homogeneous oxides, comparable to bad MOS gates, can be manufactured. lf doped SiO_2 layers are produced, they can be used as diffusion sources in a subsequent

At moderate and low solidification temperatures, the procedure delivers the so called nanocomposite. By definition, nanocomposites contain nanoparticles of less than 1000 nm in a host matrix. The following nanocomposites have already been manufactured: nano-Co/Mo, Cu, Fe, Ni, Pd, Pt and Ru in $A1_20_3$, SiO_2, TiO_2 and ZrO_2 gels, nano-C, Cu/Ni, Plei, and Pt in silica gel, nano-Ag, Ge, Os, C, Fe, Mo, Pd, Pt, Re, Ru and PtSn in silica gel-xerogel.

2.2.6 PRECIPITATION OF QUANTUM DOTS

Quantum dots are three-dimensional semiconductor materials in or on a matrix. Nano composites from semiconductor materials and the above-mentioned self assembled islands belong to this group. Sometimes it is difficult to differentiate between quantum dots and nano defects. An example is SiO_2 implanted with Ge, which is used for photoluminescence experiments. The earliest descriptions of quantum dots took place with the investigation of semiconductor precipitation in glasses. Precipitation is still used in the manufacturing of CdS, CdSe, CdTe, GaAs, and Si nanocrystallites in silica glasses. The contaminants are added to the melt, and after a further annealing step from 600 to 1400 °C, they form precipitates of controllable size, for instance, 2 nm for CdTe dots in boron silicate glass.

Another procedure is codeposition of quantum dots with thin films. There are several modifications, but the common principle is the production of nanocrystallites in a separate step (by evaporating, laser ablation, sputtering a target, etc.). They are directed towards a substrate which is, however, covered at the same time with a film so that they are included into this film.

Quantum dots can also be manufactured by means of lithography. Since a high resolution is required, electron beam lithography must be employed. The procedure takes place mostly in such a way that the material which is to be converted into quantum dots is deposited on a substrate like GaAs via MBE. The size of the electron beam spot determines the smallest possible size of the quantum dot. Therefore, if the wafer is etched, islands of this size remain. They are further reduced by etching so that quantum dots of a few 10 nm can be produced. Very often, the layers are covered in order to improve the results.

MODULE III
CHARACTERIZATION OF NANOSTRUCTURES

3.1 INTRODUCTION

Optical microscopes are generally used for observing micron level materials with reasonable resolution. Further magnification cannot be achieved through optical microscopes due to aberrations and limit in wavelength of light. In the field of characterization electron microscopy (EM) takes a vital role. Electron microscopes uses electrons instead of photons, because electron have a much shorter wavelength than photons and allow observe matter with atomic resolution.

There are two general types of electron microscopes:

- ***The scanning electron microscope (SEM)*** that scans an electron beam over the surface of an object and measures how many electrons are scattered back, and the transmission electron microscopy (TEM) that shoots electrons through the sample and it measures the electron beam changes. Because of electron scattering in the sample, it works in a vacuum environment and makes the instrument considerably larger and expensive. All electron microscopes work at low pressure and usually in high vacuum chamber to avoid scattering electron in the gas. In environmental electron microscopes, differential pumping systems are used to actually have gases present by the sample together with the electron beam.
- ***Scanning probe microscope (SPM)*** covers the methods where a sharp tip is scanned over a surface in a raster pattern and the interaction with the surface is recorded in each pixel to form an image of the interaction. There are a multitude of methods and interactions in SPM. Broadly speaking, there are three main categories.
    - In ***Scanning Tunneling Microscopy*** (STM), one uses an atomically sharp matellic tip and records the minute tunneling current between the tip and the surface, when the tip is hovering so close to the surface that electrons can move between the surface and the tip.

- In ***Atomic Force Microscopy*** (AFM), a cantilever with a sharp tip-somewhat like the needle of an old record player is scanned over the surface and the topography or surface softness can be recorded.
- In ***Scanning Near- Field Optical Microscopy*** (SNOM) a probe with a small aperture is scanned over the surface collecting the light coming from regions much smaller than the wavelength of the light used.

3.2 SCANNING TUNNELING MICROSCOPY (STM)

The scaning tunneling microscopy (STM) is a type of electron microscope that shows three-dimensional images of a sample. In the STM, the structure of a surface is studied using a stylus that scans the surface at a fixed distance from it. Fig.3.1 shows Schematic view of an STM.

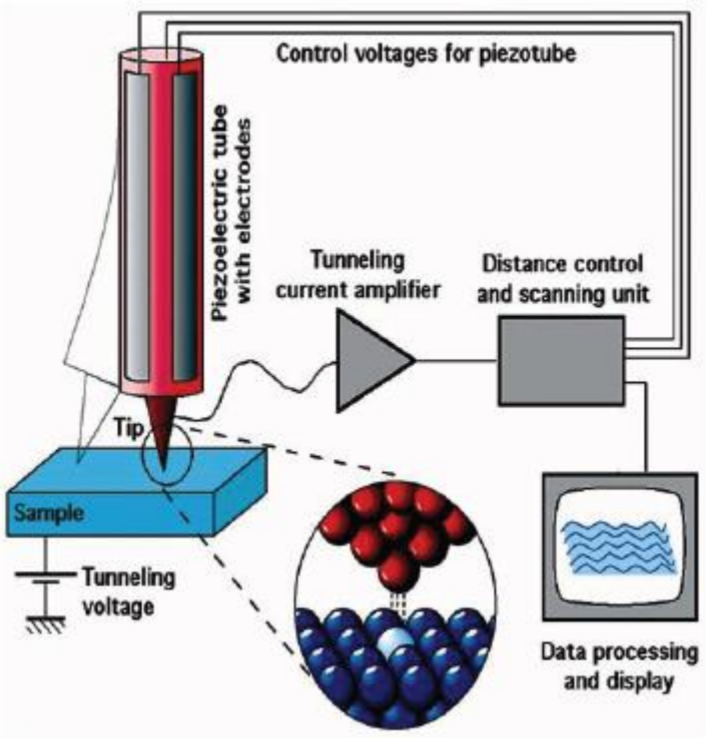


Fig.3.1 Schematic view of a STM

Scanning tunneling microscopy is an instrument for producing surface images with atomic scale lateral resolution, in which a fine probe tip is scanned over the surface of a conducting specimen with the help of a piezoelectric crystal at a distance of 0.5–1 nm, and the resulting tunneling current or the position of the tip required to maintain a constant tunneling current is monitoredThe principle of STM is based on the concept of quantum tunneling. When a conducting tip is brought very near to a metallic or semiconducting surface, a bias between the two can allow electrons to tunnel through the vacuum between them. For low voltages this tunneling current is a function of the local density of states at the Fermi level of the sample. Variations in current as the probe passes over the surface are translated into an image.

For STM, good resolution is considered to be 0.1 nm lateral resolution and 0.01 nm depth resolution. They normally generate images by holding the current between the tip of the electrode and the specimen at some constant value by using a piezoelectric crystal to adjust the distance between the tip and the specimen surface, while the tip is piezoelectrically scanned in a raster pattern over the region of specimen surface being imaged by holding the force, rather than the electric current, between tip and specimen at a set point value. When the height of the tip is plotted as a function of its lateral position over the specimen, an image that looks very much like the surface topography results. STM can be operated in two modes. In constant current imaging, a feedback mechanism is enabled and a constant current is maintained while a constant bias is applied between the sample and tip. As the tip scans over the sample, the vertical position of the tip is altered to maintain the constant separation. An alternating imaging mode is the constant height operation in which constant height and bias are simultaneously maintained.

A variation in current results as the tip scans the sample surface because a topographic structure varies in the tip sample separation. The constant current mode produces a constant directly related to electron charge density profiles, whereas the constant height mode permits faster scan rates. The STM can be used not only in ultra high vacuum but also in air and various other liquid or gas, at ambient and wide range of temperatures. STM can be a challenging technique, as it requires extremely clean surfaces and sharp tips.

The Scanning tunneling microscopy (STM) is capable of acquiring remarkable images on the most extreme scale, easily resolving atomic structure in the right environments. It is apowerful tool in nanotechnology and nanoscience providing facilities for characterization and modification of a variety of materials. STM is successfully used to detect and characterize materials like carbon nanotubes (single-walled and multi-walled) and graphine layer . The instrument has also been used to move single nanotubes or metal atoms and molecules on smooth surfaces with high precision.

3.3 ATOMIC FORCE MICROSCOPE

The atomic force microscope (AFM) is a very powerful microscope invented by Binnig, Quate and Gerber in 1986. Besides imaging it is also one of the foremost tools for the manipulation of matter at the nanoscale. Fig.3.2 shows schematic diagram of AFM

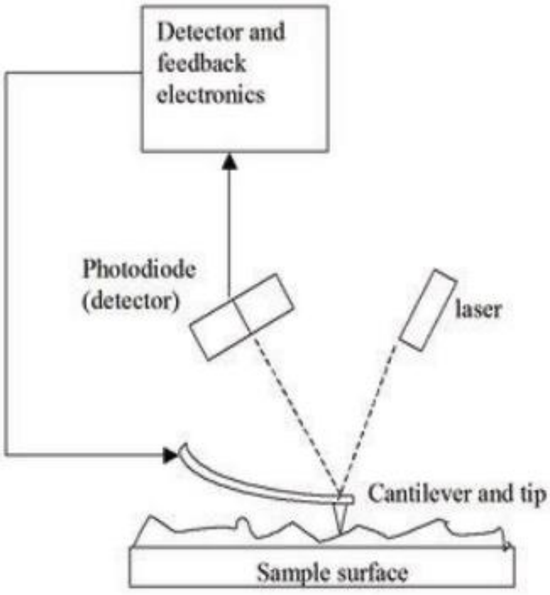


Fig.3.2 Schematic diagram of AFM

The AFM consists of a cantilever with a sharp tip at its end. The tip is brought into close proximity of a sample surface. The AFM measures the forces acting between a fine tip and a sample. The tip is attached to the free end of a cantilever and is brought very close to a surface. Attractive or repulsive forces resulting from interactions between the tip and the surface will cause a positive or negative bending of the cantilever. The force between the tip and the sample leads to a deflection of the cantilever according to Hooke's law.

Typically, the deflection is measured using a laser spot reflected from the top of the cantilever. In this procedure, a flexible force sensing cantilever stylus is scanned in a raster pattern over the surface of the sample. During the scan, it provides the topographic information.

Atomic force microscopy is currently applied to various environments (air, liquid, vacuum) and types of materials such as metal semiconductors, soft biological samples and conductive and non-conductive materials. With this technique size measurements or even manipulations of nano-objects may be performed. AFM provides a number of advantages over conventional microscopy techniques. AFMs probe the sample and make measurements in three dimensions, *x*, *y,* and *z* (normal to the sample surface), thus enabling the presentation of three-dimensional images of a sample surface. This provides a great advantage over any microscope available previously. With good samples (clean, with no excessively large surface features), resolution in the *x-y* plane ranges from 0.1 to 1.0 nm and in the *z* direction is 0.01 nm (atomic resolution).

AFMs require neither a vacuum environment nor any special sample preparation, and they can be used in either an ambient or liquid environment.

In atomic force microscopy a probe consisting of a sharp tip (nominal tip radius is in the order of 10 nm) located near the end of a cantilever beam is raster scanned across the surface of a specimen using piezoelectric scanners. Changes in the tip specimen interaction are often monitored using an optical lever detection system, in which a laser is reflected off of the cantilever and onto a position sensitive photodiode. During scanning, a particular operating parameter is maintained at a constant level, and images are generated through a feedback loop between the optical detection system and the piezoelectric scanners.
There are three scan modes for AFM, depending on the application.
The imaging modes are namely ***contact mode, non contact mode and tapping mode***.

❖ *contact mode*

Fig.3.3. shows schematic diagram for operation of AFM in the contact mode . Contact mode AFM is one of the more widely used scanning probe modes, and operates by rastering a sharp tip, made either of silicon or Si_3N_4 attached to a low spring constant cantilever across the sample.

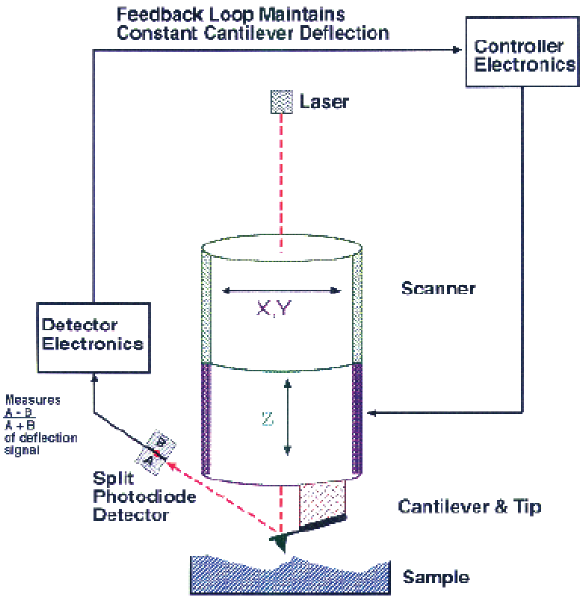


Fig.3.3 Schematic diagram for operation of AFM in the contact mode

An extremely low force approximately 10-9 N, is maintained on the cantilever, thereby pushing the tip against the sample as it raster's. Either the repulsive force between the tip and sample or the actual tip deflection is recorded relative to spatial variation and then converted into an analogue image of the sample surface.

The AFM tip is first brought close to the sample surface, and then the scanner makes a final adjustment in tip sample distance based on a set point determined by the user.

The tip, now in contact with the sample surface through any adsorbed gas layer, is then scanned across the sample under the action of a piezoelectric actuator, either by moving the sample or the tip relative to the other. A laser beam aimed at the back of the cantilever tip assembly reflects off the cantilever surface to a split photodiode, which detects the small cantilever deflections. By maintaining a constanttip sample separation and using Hooke's Law, the force between the tip and the sample iscalculated. Finally, the distance the scanner moves in the *z* direction is stored in thecomputer relative to spatial variation in the *x-y* plane to generate the topographic image of the sample surface.

❖ In the ***non contact mode***, the cantilever is externally oscillated close to its resonance frequency. The oscillation gets modified by the tip-sample interaction forces; these changes in oscillation with respect to the external reference oscillation provide information about the sample's characteristics. The dynamic mode generates lower lateral forces on the sample and is widely used to image biological sample. Non contact mode operation includes frequency modulation and the more common amplitude modulation. In frequency modulation, changes in the frequency of modulation provide information about a sample's characteristics. In amplitude modulation, changes in the oscillation amplitude yield topographic information about the sample. Additionally, changes in the phase of oscillation under tapping mode can be used to discriminate between different types of materials on the surface.

In ***tapping mode*** the cantilever tip is stimulated to vibrations near the resonant frequency (often hundreds of kilohertz) using a piezoelectric crystal. The piezo motion causes the cantilever to oscillate when the tip is not in contact with the surface of a material. The oscillating tip is then moved towards the surface until it begins to tap the surface. During scanning, the vertically oscillating tip alternately contacts the surface and lifts off, generally at a frequency of 50,000–500,000 cycles/s. As the oscillating cantilever begins to intermittently contact the surface, the cantilever oscillation is reduced due to energy loss caused by the tip contacting the surface. The reduction in oscillation amplitude is used to measure the surface characteristics. The tapping mode is less destructive than the contact mode, because the exerted forces are in the pico Newton (pN) range. Tapping mode AFM can be performed on both wet and dry sample surfaces.

Atomic force microscopy mainly used in fundamental and practical research and has important role in the development of nanotechnology. AFM can be used to explore the nanostructures, properties, and surfaces and interfaces of fibers and fabrics. For example, structural characteristics of nanofibre materials, nanolevel surface modification of textile surfaces by plasma can be easily accessed by this sophisticated technique.

The AFM has several advantages over the electron microscope. Unlike the electron microscope which provides a two-dimensional projection or a two-dimensional image of a sample, the AFM provides a true three-dimensional surface profile. Additionally, samples viewed by an AFM do not require any special treatment that would actually destroy the sample and prevent its reuse. While an electron microscope needs an expensive vacuum environment for proper operation, the AFM can work perfectly well in an ambient or even liquid environment. The main disadvantage of AFM as compared to the electron microscope is the image size. The electron microscope can show an area on the order of millimeters by millimeters and a depth of field on the order of millimeters. The AFM can only show a maximum height on the order of micrometres and a maximum area of around 100 by 100 micrometers.

3.4 SCANNING ELECTRON MICROSCOPY (SEM)

Scanning electron microscopy (SEM) is used to inspect the topographies of specimens at very high magnifications. The SEM is a microscope that uses electrons instead of light to form an image.

The scanning electron microscope (SEM) uses a focused beam of high energy electrons to generate a variety of signals at the surface of solid specimens. The signals that derive from electron sample interactions reveal information about the sample including external morphology (texture), chemical composition, crystalline structure and orientation of materials making up the sample.

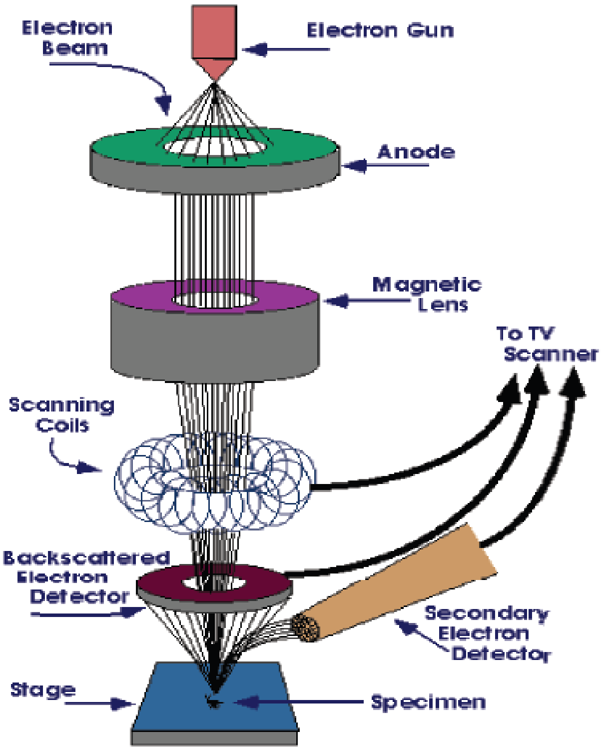


Fig.3.4 Schematic diagram of typical scanning electron microscope

In scanning electron microscope a beam of electrons is produced at the top of the microscope by an electron gun. The electron beam follows a vertical path through the microscope, which is held within a vacuum. The beam travels through electromagnetic fields and lenses, which focus the beam down toward the sample. When the primary electron beam strikes the surface of the sample it is scattered by atoms in the sample. Through this scattering event, the primary beam spreads effectively and fills a teardrop shaped volume, known as the interaction volume, and this extends about 100 nm to 5 µm into the surface. The size of interaction volume depends on the beam accelerating voltage, the atomic number of the specimen and the specimen's density.

Secondary electrons are emitted from interactions in this region, they are then detected, converted to a voltage and amplified to produce an image.

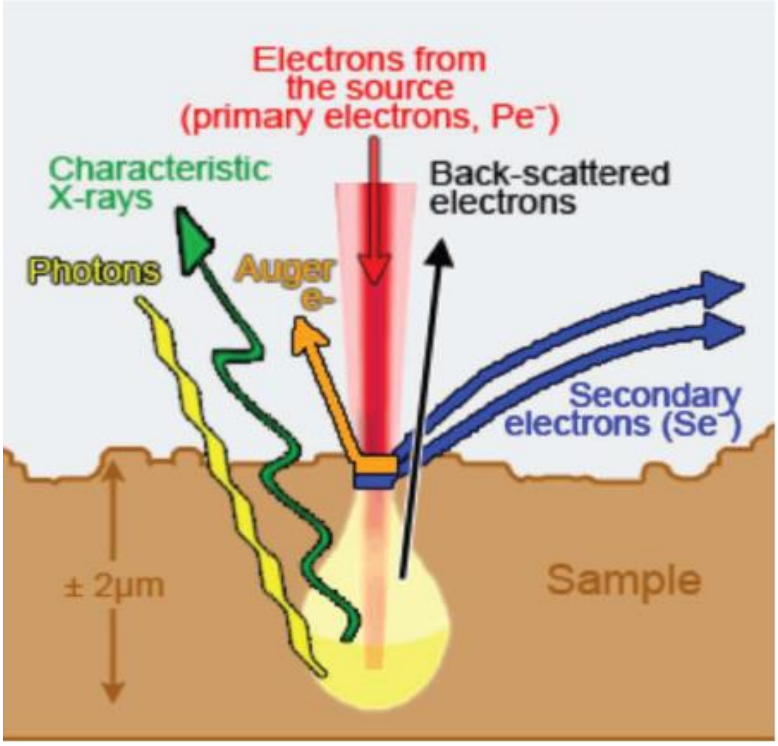


Fig.3.5 Electrons produced in SEM

In SEM, more than one type of signal can be produced and they include secondary electrons, back scattered electrons (BSEs), characteristic X-rays, specimen current and transmitted electrons. Specialized detectors are required for the detection of these signals and they are not usually all present on a single machine. Fig.3.5. shows electrons produced in SEM. The SEM can produce very high resolution images of a sample surface, revealing details about 1-5 nm in size in its primary detection mode i.e. secondary electron imaging. An additional electron interaction in the SEM is that the primary electron collide with an ejected electron from an atom will decay to its ground state by emitting either a characteristics X-ray photon or Auger electron. This characteristics X-ray or Auger electrons have been used for chemical characterization i.e to identify the elemental composition of the sample by a technique known as energy dispersive X-ray .

Back scattered electron images can be used for rapid discrimination of phases in multiphase samples. SEM required with diffracted backscattered electron detectors can be used to examine micro fabric and crystallographic orientation in many materials . SEM requires that the specimens should be conductive for the electron beam to scan the surface and that the electrons have a path to ground for conventional imaging. Non conductive solid specimens are generally coated with a layer of conductive material by low vacuum sputter coating or high vacuum evaporation. This is done to prevent the accumulation of static electric charge on the specimen during electron irradiation. Non conducting specimens may also be imaged uncoated using specialized SEM instrumentation such as the "Environmental SEM" (ESEM) or in field emission gun (FEG) SEM operated at low voltage, high vacuum or at low vacuum, high voltage.

The scanning electron microscope has many advantages over traditional microscopes. The SEM has a large depth of field that allows more of a specimen to be in focus at one time. The SEM also has much higher resolution, so closely spaced specimens can be magnified at much higher levels. As the SEM uses electromagnets rather than lenses, the researcher has much more control in the degree of magnification. All of these advantages, as well as the actual, strikingly clear images, make the scanning electron microscope one of the most useful instruments in research today. The SEM shows very detailed three dimensional images at much high magnifications (up to × 300000) as compared to light microscope (up to × 10000) and can also be used to view dispersion of nanoparticles with a seemingly three dimensional structure. But as the images are created without light waves, they are black and white.

3.5 TRANSMISSION ELECTRON MICROSCOPY (TEM)

The transmission electron microscope (TEM) operates on the same basic principles as the light microscope but uses electrons instead of light. TEMs use electrons as “light source” and their much lower wavelength make it possible to get a resolution a thousand times better than that with a light microscope. With the help of TEM objects can be seen in the order of a few angstroms (10^{-10} m). It can be used to study small details in the cell or different materials down to near atomic levels.

At smaller magnifications TEM image contrast is due to absorption of electrons in the material, due to the thickness and composition of the material. At higher magnifications complex wave interactions modulate the intensity of the image, requiring expert analysis of observed images. Alternate modes of use allow TEM to observe modulations in chemical identity, crystal orientation, electronic structure and sample induced electron phase shift as well as the regular absorption based imaging.

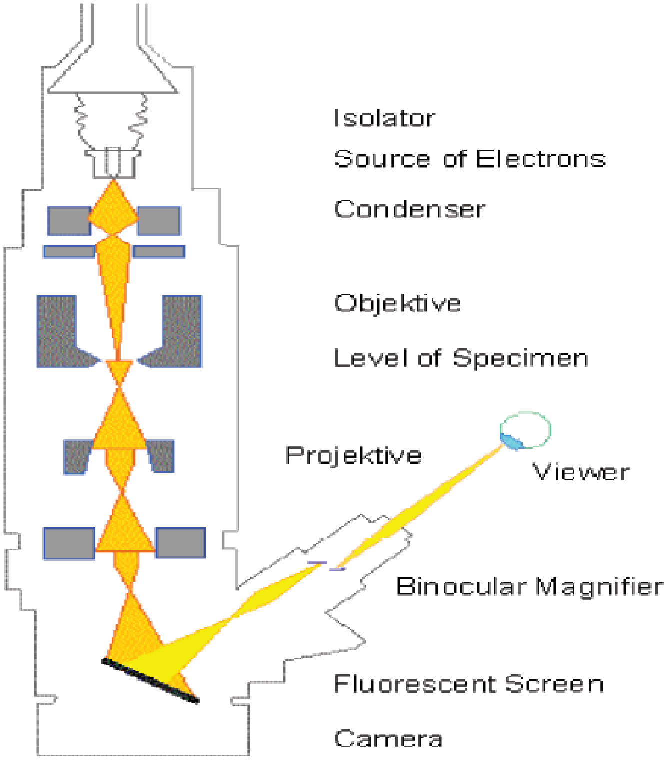


Fig.3.6 Schematic diagram of TEM

Transmission electron microscopy is a microscopy technique whereby a beam of electrons is transmitted through a very thin sample and interacting with the sample as it passes through. The interaction of the electrons are transmitted, magnified and then focused onto a fluorescent screen or it can be detected by a CCD camera. Fig.3.6. shows a schematic diagram of a TEM. The operation of the TEM requires an ultra high vacuum and a high voltage. The electrons are emitted by a source and then it is focused and magnified by a set of magnetic lenses. The image that is formed is either shown on a fluorescent screen or on a monitor and it is printed on photographic film.

In TEM, the crystalline sample interacts with the electron beam mostly by diffraction rather than by absorption. The intensity of the diffraction depends on the orientation of the planes of atoms in a crystal relative to the electron beam; at certain angles the electron beam is diffracted strongly from the axis of the incoming beam, while at other angles the beam is largely transmitted. TEM exploits three different interactions of electron beam specimen like unscattered electrons (transmitted beam), elastically scattered electrons (diffracted beam) and inelastically scattered electrons.

Different types of images are obtained in TEM with the different types of electrons using the apertures properly. As a result, diffraction patterns are shown because of the scattered electrons. If the unscattered beam is selected, we obtain the bright field image or light field images. It is also possible to produce an image from electrons deflected by a particular crystal plane which is known as a dark field image.

TEM forms a major analysis method in a range of scientific fields, in both physical and biological sciences. TEMs find application in cancer research, virology, materials science as well as pollution, nanotechnology and semiconductor research. TEM images reveal the distribution and dispersion of nanoparticles in polymer matrices of nanocomposite fibers, nanocoatings and etc. The extent of exfoliation, intercalation and orientation of nanoparticles can also be visualized using the TEM micrograph. Today's transmission electron microscopes offer resolutions up to 0.1nm at 300kV and probe diameters up to 0.34nm.

3.6 X - RAY DIFFRACTION (XRD)

X - ray diffraction technique is the most common and efficient method for the determination of structure and crystallainity and material identification. XRD is an apt method to examine whether a resultant material has amorphous or crystalline nature. Crystalline phases can be identified by just comparing the interplanar distance 'd' values obtained from XRD data with the fundamental data in Joint Committee on Powder Diffraction Standards (JCPDS).

X - ray diffraction is based on constructive interference of monochromatic X - rays from a crystalline sample. The X - rays, generated by a cathode ray tube are filtered to produce monochromatic radiation, collimated and directed towards the sample. X - ray primarily interact with electrons in atoms, collide and some photons from the incident beam are deflected away from original. The X - rays interfere constructively and destructively producing a diffraction pattern on the detector. The incident X - ray radiation produces a Bragg peak if their reflections from the various planes interfered constructively. The interference is constructive, when the phase shift is a multiple of 2π, this condition can be expressed by Bragg's law.

$$n\lambda = 2d\ sin\theta$$

where n is an integer, λ is the wavelength of incident wave, d is the spacing between the planes in the atomic lattice and θ is the angle between the incident ray and the scattering planes. Schematic diagram of X - ray diffraction is shown in Figure 3.7

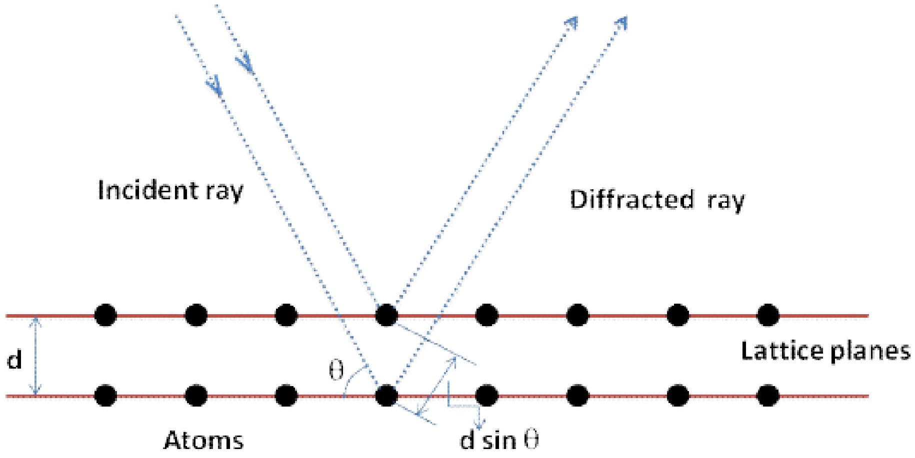


Fig 3.7 Bragg's law

A typical powder X - ray diffractometer consists of a source of radiation, a monochromator to choose the wavelength, slits to adjust the shape of the beam, a sample and a detector. A goniometer is used for fine adjustment of the sample and the detector positions. The goniometer mechanism supports the sample and detector, allowing precise movement. The source X - rays contains several components; the most common being Kα and Kβ. The specific wavelengths are characteristic of the target material (Cu, Fe, Mo, Cr). Monochromators and filters are used to absorb the unwanted emission with wavelength Kβ, while allowing the desired wavelength, Kα to pass through. The X - ray radiation most commonly used is that emitted by copper, whose characteristic wavelength for the Kα radiation is equal to 1.5418 Å. The filtered X - rays are collimated and directed onto the sample as shown in the Figure 3.8 .

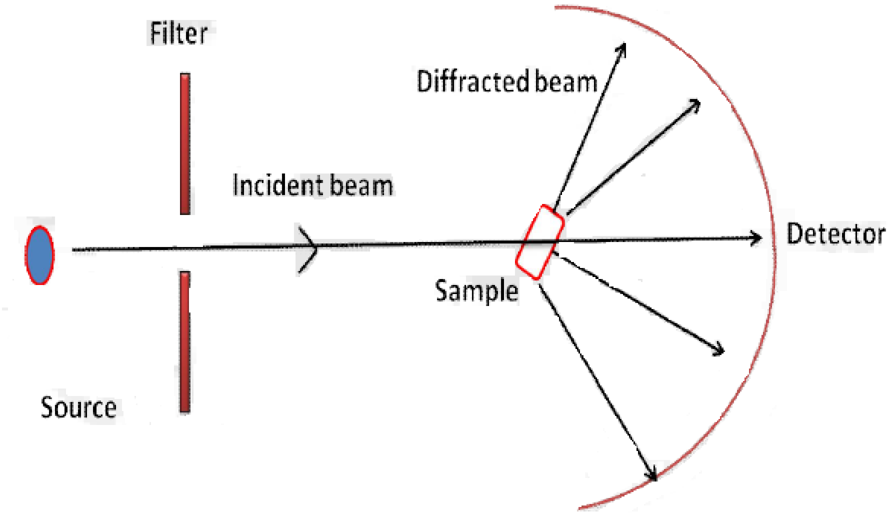


Figure 3.8 Schematic diagram of the diffractometer

When the incident beam strikes a powder sample, diffraction occurs in every possible orientation of 2θ. The diffracted beam may be detected by using a moveable detector such as a Geiger counter, which is connected to a chart recorder.

The counter is set to scan over a range of $2\ \theta$ values at a constant angular velocity. Routinely, a $2\ \theta$ range of 5 to 70 degrees is sufficient to cover the most useful part of the powder pattern. The scanning speed of the counter is usually $2\ \theta$ of 2^{o} min^{-1}. A detector records and processes this X - ray signal and converts the signal to a count rate which is then fed to a device such as a printer or computer monitor. The sample must be ground to fine powder before loading it in the glass sample holder. Sample should completely occupy the square glass well.

3.7 UV - VISIBLE SPECTROSCOPY (UV - Vis)

This refers to absorption spectroscopy in the ultra - violet and visible spectral region. In this region of the electromagnetic spectrum, molecules undergo electronic transition. When sample molecules are exposed to light having an energy ($E = hv$ where 'E' is energy in joules, 'h' is Planck's constant 6.62×10^{-34} J s and 'v' is frequency in Hertz), that matches a possible electronic transition within the molecule, some of the light energy will be absorbed as the electron is promoted to a higher energy orbital. An optical spectrometer records the wavelengths at which absorption occurs, together with the degree of absorption at each wavelength. The resulting spectrum is presented as a graph of absorbance (A) versus wavelength (λ). The optical properties of materials can be studied with the help of UV - Vis spectra.

Principle

The absorbance of light by molecules in the solution is based on the Beer - Lambert law,

$$A = log\frac{I_o}{I} = \varepsilon \times b \times c$$

where, $I0$ is the intensity of the reference beam and I is the intensity of the sample beam, ε is the molar absorbtivity with units of L mol^{-1} cm^{-1}, b = path length of the sample in centimeters and c = concentration given solution expressed in mol L^{-1}

The main components of the UV - Vis spectrometers are a light source, double beams (reference and sample beam), a monochromator, a detector and a recording device. The source is usually a tungsten filament lamp for visible and deuterium discharge lamp for UV measurements. The light coming out of the source is split into two beams - the reference and the sample beam as shown in the Figure 3.9.

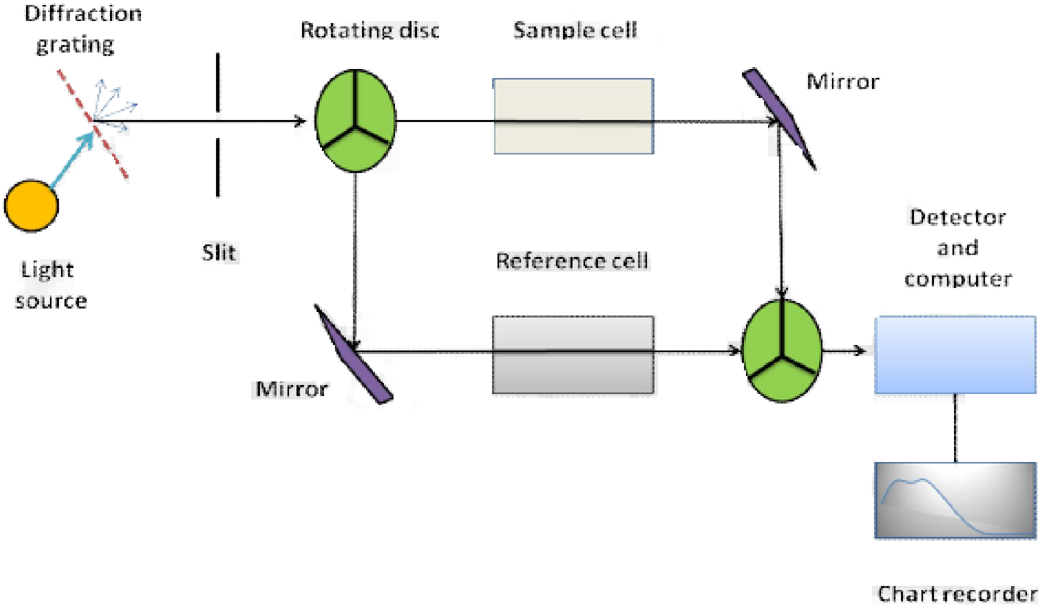


Fig 3.9: Functional Block diagram of a double beam UV – Visible Spectrophotometer

The sample and reference cells are rectangular quartz / glass containers; they contain the solution (to be tested) and pure solvent, respectively. The spectrometer records the ratio between the reference and sample beam intensities. The recorder plots the absorbance (A) against the wavelength (λ).

3.8 PHOTOLUMINESCENCE SPECTROSCOPY (PL)

Photoluminescence spectroscopy is a contactless, non - destructive method to probe the electronic structure of materials. The intensity and spectral content of the emitted photoluminescence is a direct measure of various important material properties, including band gap determination, impurity levels and defect detection, recombination mechanisms.

Principle

Light is directed onto a sample, where it is absorbed and imparts excess energy into the material in a process called photo - excitation. Photo – excitation causes electrons within a material to move into permissible excited states. These electrons return to their equilibrium states, by a radiative process (the emission of light) or by a non - radiative process as shown in Figure 3.10. The quantity of the emitted light is related to the relative contribution of the radiative process.

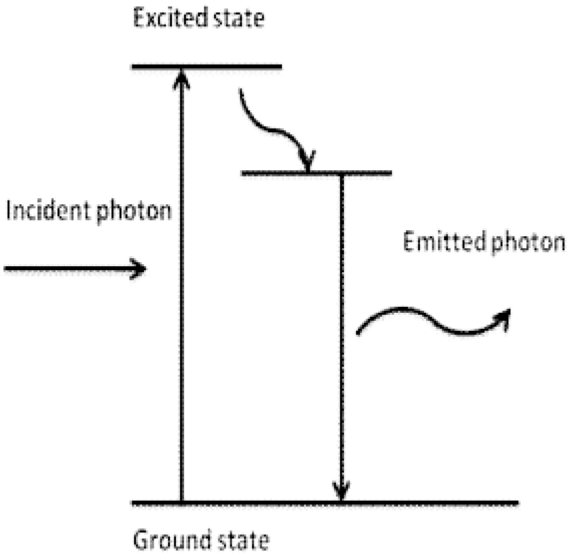


Fig 3.10 : Principle of photoluminescence

The fluorescence instruments contain three basic items: a source of light, a sample holder and a detector. A schematic representation of a fluorimeter is shown in Figure 3.11. The light source produces light photons over a broad energy spectrum, typically ranging from 200 to 900 nm. Photons impinge on the excitation monochromator, which selectively transmits light in a narrow range centered about the specified excitation wavelength. The transmitted light passes through adjustable slits that control magnitude and resolution by further limiting the range of transmitted light. The filtered light passes into the sample cell causing fluorescent emission by fluorphors within the sample. Emitted light enters the emission monochromator, which is positioned at a 90 º angle from the excitation light path to eliminate background signal and minimize noise due to stray light. Again, emitted light is transmitted in a narrow range centered about the specified emission wavelength and exits through adjustable slits, finally entering the photomultiplier tube (PMT). The signal is amplified and creates a voltage that is proportional to the measured emitted intensity. Noise in the counting process arises primarily in the PMT. Therefore, spectral resolution and signal to noise is directly related to the selected slit widths.

Sample preparation process is the same as that of UV - Visible spectroscopy. In both the cases, the sample cell (cuvette) must be free from contaminants.

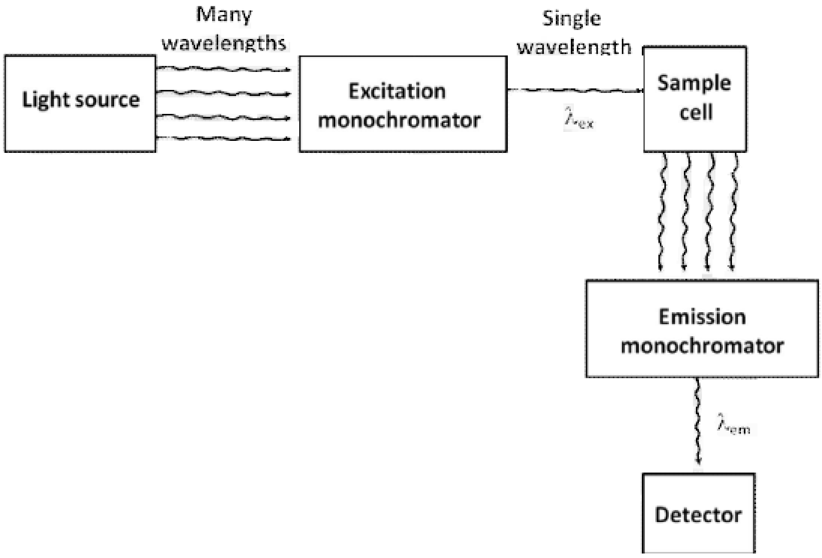


Figure 3.11 Schematic representation of a fluorescence spectrophotometer

MODULE IV

SEMICONDUCTOR QUANTUM NANOSTRUCTURES AND SUPERLATTICES

4.1 INTRODUCTION

Modern microelectronic and optoelectmnic devices can be constructed from a few basic structures, a mong them heterojunctions, or junctions between two semiconductors of different gaps, and metal-oxide-semioonductor (MOS) structures. These two kinds of nanostructures provide electrons with a potential well for electrons of nanometric size at the junction. MOS structure is the basic block of the most important device in microelectronics, the metal-oxide-semiconductor field-effect-transistor (MOSFET. In fact, the discovery of the quantum Hall effect (QHE) by IL von Klitzing in 1980 was based on the study of the transport properties of electrons in the channel of a MOSFET under the influence of simultaneous electric and magnetic fields [1]. However, these new quantum effects (QHE, Aharonov-Bohm effect and Shubnikovde Haas oscillations) are better observed in III-V heterojunctions, since the electron effective mass is much lower in materials such as GaAs than in Si.

In this chapter we shall study first the behaviour of electrons confined in 2D wells at the semiconductor-oxide interface in MOSFET transistors. Next, we shall proceed with the III-V modulation-doped heterojunction used in high-frequency transistors. Next. we will focus on the modulation-doped. square potential, quantum wells. This simple building block is used as a single unit or more often as a multiple quantum well structure in devices. ***When the thickness of the barriers separating the wells is small, tunnelling of elections between neighbouring wells takes place and the resulting device called a superlattice***, displays a band diagram similar to that of electron bands in crystals. However, the allowed energy bands and gaps correspond to much smaller energy intervals, since the SL has a spatial periodicity (equal to the sum of well and barrier thicknesses) which is much larger than the lattice constant. The band structure of SL can be engineered by a proper choice of the well and barrier widths. In this sense, SLs can be considered as artificial solids since their electron energy band structure is similar to the ones in crystals, but they do not exist in nature.

4.2 MOSFET STRUCTURES

The MOSFET is formed by a MOS structure and two p-n"' junctions in which the n material is heavily doped (Figure 4.1(a)), which act as the source and drain of the FET. The gate of the transistor is formed by the MOS structure. The semiconductor is usually p-type silicon over which a thin oxide layer (gate oxide)is grown by thermal oxidation.

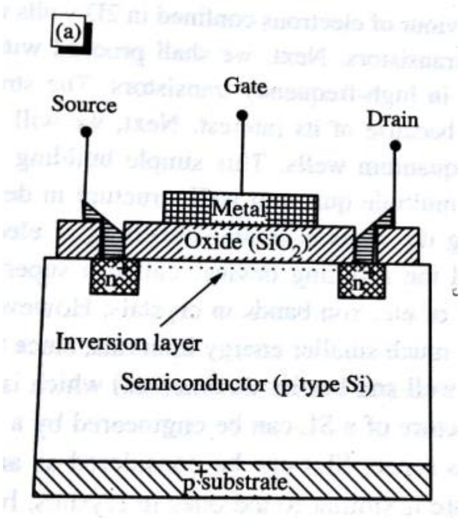


Figure 4.1(a)Structure of MOS Transistor

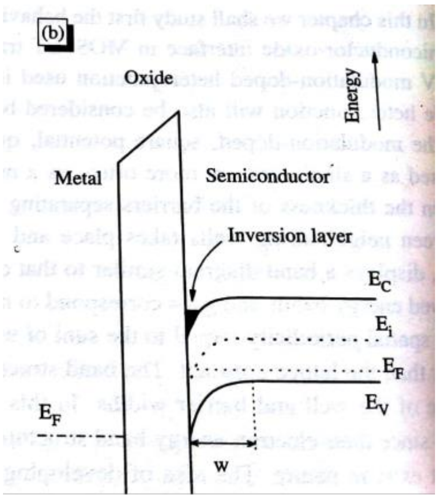


Figure 4.1(b) : Band diagram of MOS structure

Figure 4.1(b) shows the band diagram of the MOS structure for a p-type silicon semiconductor under a fairly strong positive bias. When a positive potential is applied to the gate, the electrons coming from the n+ regions and some from the bulk p-silicon are accumulated at the Si-Si0$_2$ interface. These electrons form the so-called inversion layer or channel and are located in an almost triangular-shaped potential well of nanometric dimensions. The shape of the well is due to the space charge of ionized acceptors in the p-type silicon, whose corresponding holes are repelled by the electric held across the dielectric oxide produced by the positive gate potential. If a positive potential is applied between drain and source, the electrons in the channel will create a current. The current can be modulated by changes at the potential gate, since the amount of electrons in the inversion layer depends on the magnitude of the electric field across the insulator. The name “Field Effect Transistor” is due to the control effect of the electric field applied to the gate.

In order to observe quantum effects at the Si-SiO$_2$ interface, several conditions must be met: (a) the SiO$_2$ insulator, of amorphous nature, should have neither high concentration of impurities (Na+ ions) nor of trapped charge; (b) the smoothness of the Si-Si02 interface should be controlled at the atomic size level, since a rough surface on top of the inversion channel would greatly decrease the electron mobility in the inversion layer.

In order to study the behaviour of electrons in the potential well, it should be recognized that the electron inversion layer can be considered a 2D system of electrons immersed in a triangular-shaped quantum well ,located in the semiconductor, close to the interface with the oxide. In the MOS structure, although the electrons are confined along the perpendicular direction, they are practically free to move in the plane of the interface. Therefore, the quantized values for the energy of confinement should be given by

$$E = En + \frac{h^2}{2m_x{}^*}k_x{}^2 + \frac{h^2}{2m_y{}^*}k_y{}^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Where En corresponds to quantized energy for triangular well.

$$En \approx \left(\frac{3}{2}\pi\left(n-\frac{1}{4}\right)\right)^{\frac{2}{3}}\left(\frac{e^2F^2h^2}{2m_Z{}^*}\right)^{\frac{1}{3}}, \quad n = 1,2,3.. \quad \ldots\ldots(2)$$

Eqn (1) represents parabolas in reciprocal space , the bottoms having values of En.

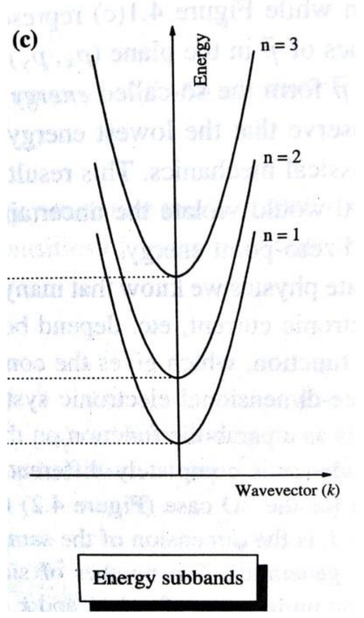


Figure 4.2 : Energy subbands

Similarly the density of states (DOS) function corresponds to the 2D case and is given by

$$g(E) = g_v\frac{m_T{}^*}{\pi h^2} \quad \ldots\ldots\ldots\ldots(3)$$

where we have added the factor g_v which takes into account the conduction band valley degeneracy. This degeneracy arises from the fact that constant energy surfaces of the silicon conduction band are formed by six ellipsoids in the <001> direction of momentum space.

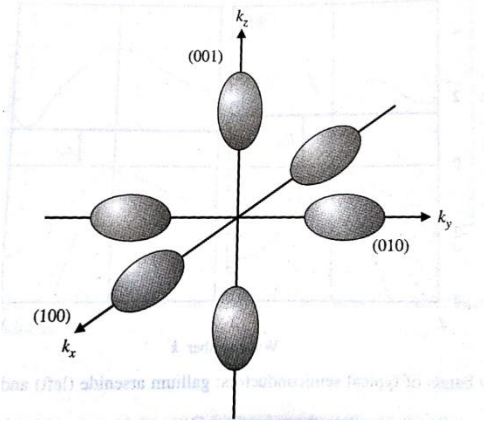


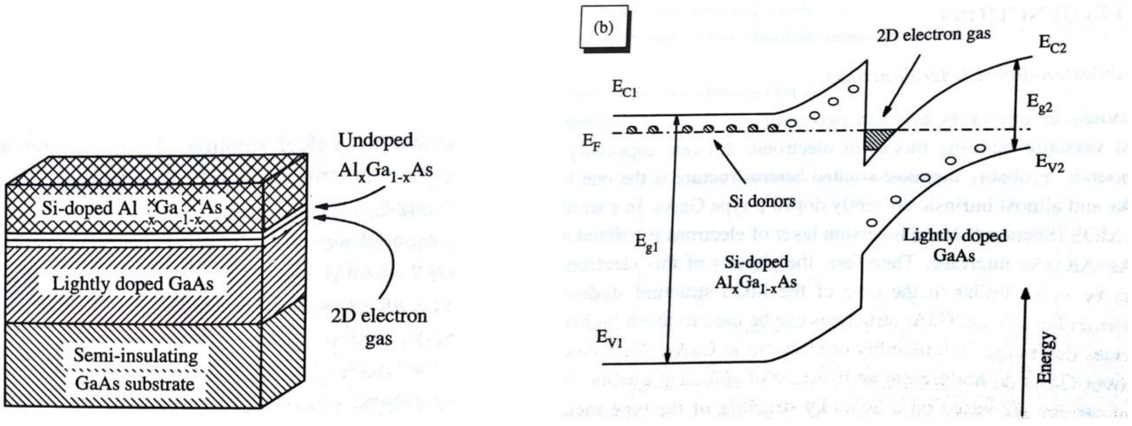
Figure 4.3 : Constant energy surfaces in silicon around the minima of conduction band

The long axis of the ellipsoid corresponds to the longitudinal effective mass $m_T{}^*$=0.91m_o and the two equal short axes to the transversal effective mass $m_L{}^*$ = 0.19mo, which is the one that appears in Eq. (3). Therefore, there are two conduction band minima corresponding to the heavy effective mass and four to the light one. As a consequence, after solving the Schrodinger equation, neglecting coupling of the electrons in the various conduction band minima, one should expect two different subband values or subband ladders for the Si <001>. First consider the electrons in valleys perpendicular to the interfaces. The effective mass which enters Eq. (2) is $m_L{}^*$ Also, $m_x{}^*$ =$m_y{}^*$; = $m_T{}^*$ in Eq. (1) and g_v = 2. This results in subbands of lower energy (higher value of effective mass). Evidently the second subband ladder is originated when the parallel valleys are considered. In this case, m_T is the effective mass for the expression of En in Eq. (2) and gv = 4.

4.2 HETEROJUNCTIONS

4.2.1 MODULATION-DOPED HETEROJUNCTIONS /HIGH ELECTRON MOBILITY TRANSISTOR

Heterojunctions, or interfaces between two semiconductors of different gaps, are one of the most versatile building blocks of electronic devices, especially those based on III-V compounds. Probably the most studied heterostructure is the one formed by n-type Al_x Ga_{1-x} As and almost intrinsic or lightly doped p-type GaAs. In a similar fashion to the case of the MOS, an inversion layer of electrons is formed in the GaAs close to the GaAs-AlGaAs interface. Devices based on AlGaAs structures can be used to much higher frequencies than silicon devices due to the high mobility of electrons in GaAs. Since oxides and insulators deposited over GaAs do not present an interface of sufficient quality, the most important device applications are based on a Schottky structure of the type metal-AlGaAs-GaAs (Figure 4.4 (a)). In this section we will focus on the properties and band diagram of the Alea1_xAs-GaAs heterojunction shown m Figure 4.4(b).



4.4 (a) : Structure of a AlGaAs-GaAs Modulation dopped Hetero junction (b) Band diagram

Let us first consider, from a qualitative point of view, how an electron well of nanometric size is formed at the AlGaAs-GaAs interface.

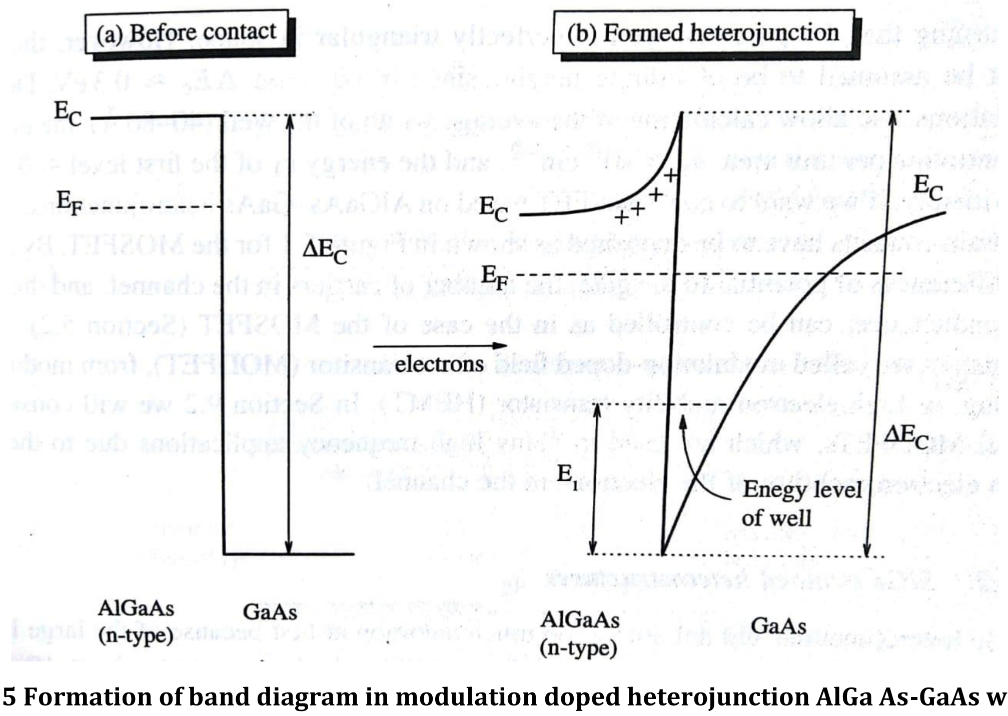


Figure 4.5 Formation of band diagram in modulation doped heterojunction AlGa As-GaAs when both semiconductors enter in contact

Suppose, as in Figure 4.5, that we have an AlGaAs-GaAs heterojunction, where the left material is gallium arsenide doped with aluminium and the right one is near-intrinsic GaAs. This structm‘e is called a modulation-doped heterojunction and the method to produce it is known as modulation doping. First consider the hypothetical situation of Figure 4.5(a), before the two semiconductors enter in contact. In the figure, for simplicity, we only draw the bottom of the conduction bands and the Fermi level, which in the case of n-type AlGaAs is close to the conduction band, and for lightly p-doped GaAs is located close to the middle of the gap. Evidently the bands are flat because the materials are electrically neutral and have uniform doping. The barrier between them in the conduction band, ΔEc, can be approximately found following Anderson's rule. According to this rule, when we join two materials, the vacuum levels should line up. If χ_AA and χ_B are the electron afiinities of the AlGaAs and GaAs, respectively, we should have ΔEc $\equiv \chi_A - \chi_B$, since the electron aflinity of a semiconductor is defined as the energy required for an electron located at the bottom Ec of the conduction band, to get out of the solid, i.e. χ = E_{vac} -EC. According to this rule, one gets a value of ΔEc of 0.35 eV for a doping x in Al_x Ga_{1-x} As around 0.3.

When both materials, AlGaAs and GaAs, enter in contact, some of the electrons from the donors of the n-material will cross the interface reaching the undoped GaAs. Therefore, as in the p-n junction, an internal electric field will be created and directed from the non-neutralized donors in the AlGaAs to the additional electronic charges in the GaAs. This field is the one that causes the band bending shown in Figure 4.4(b). At equilibrium, the two Fermi levels line up, the bands are bent like in the case of the p-n junction, with the only difference that the barrier ΔEc is created. Note also that far from the interface the bottom EC of the conduction bands is flat and at the same distance from the Fermi level E_F as in the case of Figure 4.5(a). Therefore, it is relatively easy to sketch the band diagram of Figure 4.5(b).As it can be appreciated, a quantum well for the electrons has been formed which is limited by a potential well of height ΔEc in the left and a potential energy curve at the right, within the GaAs. The quantum well for the electrons produced at the AlGaAs-GaAs interface has a shape close to a triangle as in the case of the MOS structure. Therefore, if we call Z the direction perpendicular to the interface, the electrons forming the 2D inversion layer are free to move along the (x, y) plane, but their energy for the motion along Z is quantized as in a potential well.

The most important aspect of this heterojunction is that the charge carriers are located in a region (mainly in the GaAs), spatially separated from the AlGaAs semiconductor which originates the free electrons. The electrons in the well should have very high mobility for their motion along the (x, y) plane, since they move within the GaAs which is free of dopent impurities and it is well known that impurity scattering is one of the main factors which limit carrier mobility, especially at low temperatures. Evidently, the electron mobilities are also much higher than in the case of the MOS structure.

4.2.2 SiGe STRAINED HETEROSTRUCTURES

Figure 4.6 shows two typical examples of SiGe heterostructures. In Figure 4.6(a) the substrate is <001> Si (Ex = 1.17 eV) and the strained active layer $Si_{0.7}Ge_{0.3}$ (Eg = 0.78 eV). In this case the conduction band offset is rather small. in contrast to the valence band offset. This situation allows the formation of a 2D hole gas in the SiGe alloy, with electron mobilities around 2 $m^2V^{-1}S^{-1}$, i.e. about half the value found for electrons in a typical MOSFET .In Figure 4.6(b) the situation is reversed and the strained layer is Si. In this case the discontinuity in the conduction band is fairly large and the electrons form a 2D gas, with free motion in the plane of the interface The silicon effective mass corresponding to this motion is the low transversal one ($m_T{}^* \approx 0.19m_0$)therefore yielding a high mobility of around 20 $m^2V^{-1}S^{-1}$,several times higher than the one corresponding to the MOSFET.

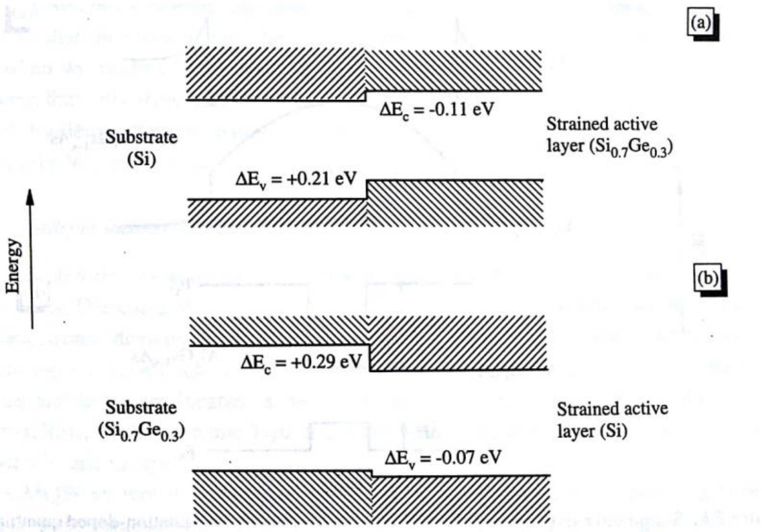


Figure 4.6: Band diagram of two strained SiGe heterostructures
(a). active layer is SiGe (b). Substrate layer is SiGe

SiGe heterostructures have also found an important application in the field of bipolar silicon transistors. One way to improve the efficiency of a bipolar transistors is to use a narrow-bandgap material for the base region, which improves the efficiency of the Si emitter region. In this case the advantage stems from the reduced values of the SiGe bandgap alloys, in comparison to silicon. It is important to mention that the strain that appears in the heterojunction also contributes to the decrease of the bandgap. In addition, the large bandgap offset allows the fabrication of a highly doped, low resistivity, base material, which extends the performance of silicon transistors to much higher frequencies.

4.3 QUANTUM WELLS

4.3.1 MODULATION DOPED QUANTUM WELLS

Modulation doping is a technique for fabricating semiconductors such that the free charge carriers are spatially separated from the donors. Because this eliminates scattering from the donors, modulation-doped semiconductors have very high carrier mobilities.

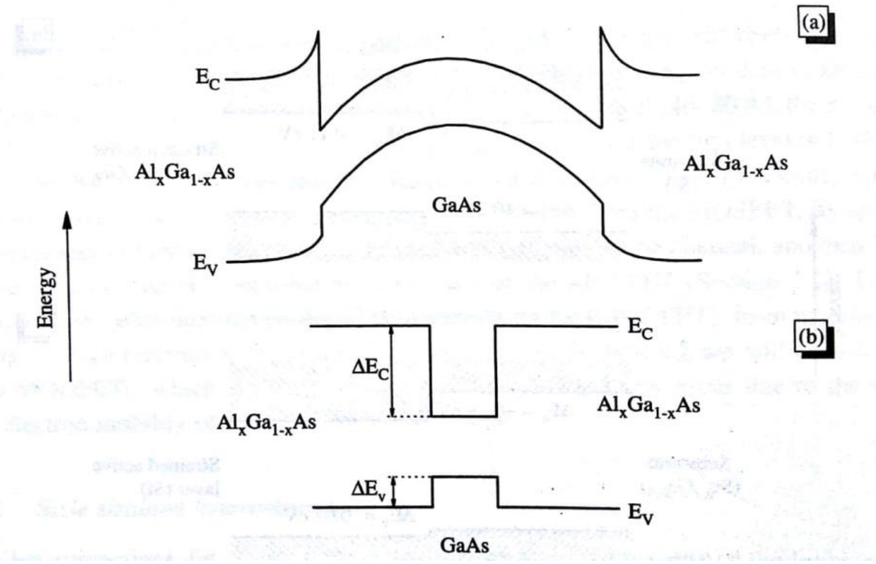


Figure 4.7 Band diagram of AlGaAs-GaAs modulation dopped quantum well (a). wide well (b). Narrow well

To build a symmetric well by facing two AlGaAs-GaAs heterojunctions opposing each other like in Figure 4.7 (a). The wide gap semiconductor material AI_xGa_{1-x} As is located at the ends and the GaAs in the middle. Assuming the distance between the two interfaces is sufficiently small. Then the resulting well (Figure 4.7(b)) for electrons and holes would be almost square with a barrier on each side of the same height as in Figure 4.7(a).

The volume of material inside the well is free of ionized donors (located in the AlGaAs material). Therefore the electrons inside the well, which originated at the neighbouring AlGaAs donor-material, can move into the GaAs region or channel with very high mobility. As in the case of the modulation-doped heterojunctions, MODFET high-frequency transistors can be fabricated if appropriate source and drain contacts are deposited.

Quantum well structures with either high or low mobility for electrons can be fabricated by introducing a controlled amount of impurities. A double quantum well structure with high and low mobilities constitutes the base of the velocity-modulation transistors. In these transistors, the switching from one state to the other is controlled by an electric field transverse to the layers which redistributes the amount of charged electrons and therefore the current in either wells. Velocity-modulation transistors can be operated at very high frequencies.

4.3.2 MULTIPLE QUANTUM WELLS (MQW)

The signal provided by a single quantum well is usually too small to be used in the solid state devices. Therefore, it is often necessary to use an array of quantum wells, especially in optoelectronic devices, such as photodetectors. These structures are called multiple quantum wells (MQW) and are formed by several single quantum wells. If the wells for electrons and holes are located in the same space location, the MQW is called Type I (Figure 4.8(a)), while the name Type II is used when the corresponding wells are located alternatively as in Figure 4.8 (b).

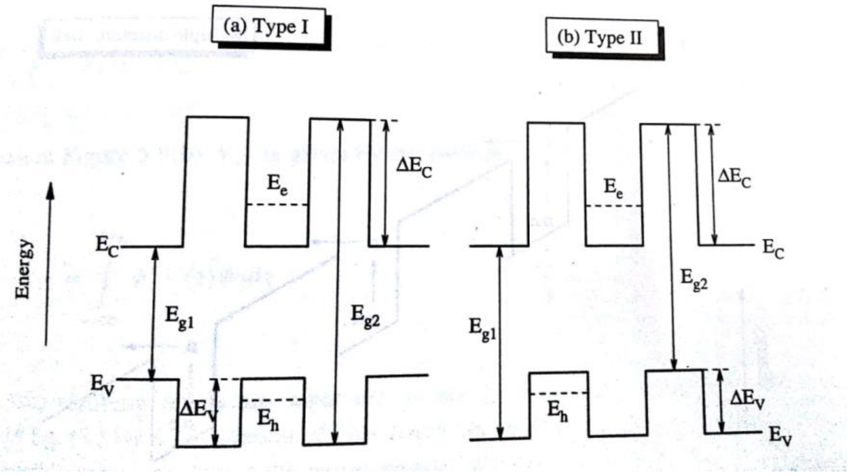


Figure 4.8: (a) Type I MQW (b) Type II MQW

In a MQW system it is assumed that there is no interaction between neighbouring quantum wells, because the barriers separating the wells are thick enough, usually more than about 10 nm.

However, if the energy barriers between consecutive wells are thin enough, the wells will be coupled to each other by tunnelling effects. The discrete energy levels of the quantum wells are then transformed into energy bands. In this case, the system of MQWs is called a ***super lattice*** and the energy spectrum shows very interesting new features.

4.4 THE CONCEPT OF SUPER LATTICES

A superlattice consists of a periodic set of MQW in which the thickness of the energy barriers separating the individual wells is made sufficiently small. As the barriers become thinner, the electron wave functions corresponding to the wells overlap due to the tunnelling effect. As a consequence, the discrete energy levels of the wells broaden and produce energy bands, in a similar way as happens with the states of the individual atoms when they are arranged in a crystal lattice. The most singular aspect of a superlattice consists of introducing at will a new periodicity ***d*** in the material, which is equal to the breadth of the well ***a***, plus the thickness of the barrier ***b.*** Typical thicknesses for a and b could be 4 and 2 nm, respectively. An accurate control over these small thicknesses can only be achieved by techniques for thin film deposition such as molecular beam epitaxy or metal organic chemical vapour deposition.

In order to study the origin of the band structure of superlattices, let us consider first the overlapping between the electron states for a simple two-well system. This is already a familiar problem, because from a quantum mechanical point of view it is formally similar to the case of the diatomic molecule.

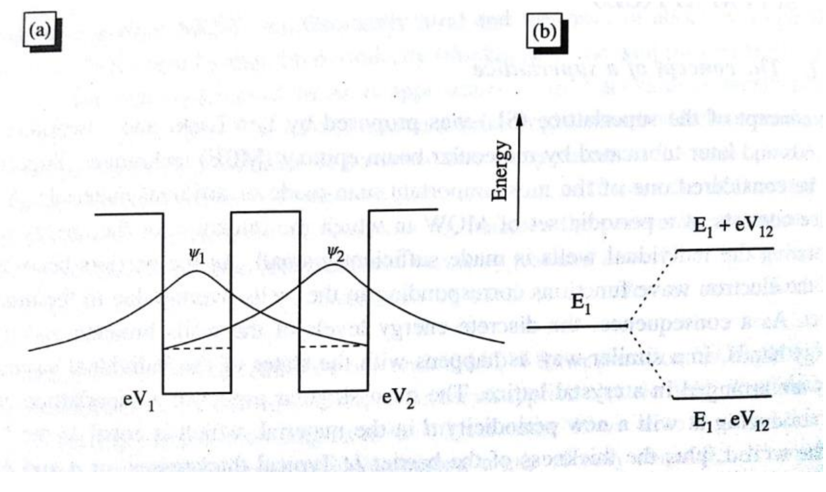


Figure 4.9: (a) Double coupled quantum well (b) Splitting of energy levels

Figure 4.9(a) shows two neighbouring identical quantum wells and corresponding wave functions of what is known as the double coupled quantum well system. The solution for this problem is based on perturbation theory in quantum mechanics. According to it, each original level, say E_1, of the isolated wells splits into two, with energies

$E = E_1 \pm |V_{12}|$ ……………………..(1)

$V_{12} = \int_{-\infty}^{\infty} \psi_1{}^{*} V(z) \psi_2 dz$ ………..(2)

The two resulting levels are separated in energy by $2|V_{12}|$ where the magnitude of V_{12} of Eq. (2) is an indication of how much one well can iniiuence the energy states of the neighbouring one, hence the name overlap integral.

4.5 KRONIG - PENNEY MODEL OF SUPER LATTICE

The fundamental nature of insulators, conductors and semiconductors can be functionally explained based on band theory. The recent development in semiconductor physics, the semiconductor hetero-structures are also analyzed using the concept of band theory. In the Kronig Penney model, instead of experiencing a gradual variation in the strength of the potential electrons experience a maximum potential (potential well) and minimum value (potential barrier) in the presence of the lattice planes. Kronig and Penney established a model for a solid in which the periodic potential seen by the electrons was precisely that of the square type shown in Figure 4.10 for a superlattice potential.

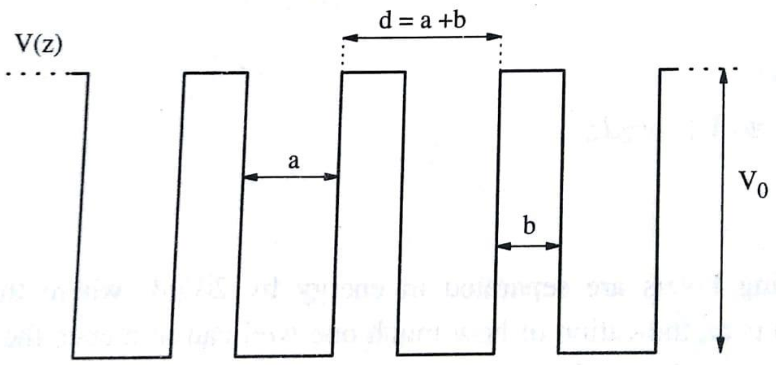


Figure 4.10: Scheme of the periodic potential of a superlattice

This periodic one-dimensional potential is characterized by the following parameters: well thickness a, barrier thickness b, and barrier height V_0. The spatial periodicity is d = a + b. Even if the periodic potential of Figure 4.10 is much simpler than the real one, the Kronig-Penney model yields very interesting results related to the structure of the bands, the forbidden zones, size of the gaps, etc. In the well region (0 < z < a), V = 0, and the wave function is

$\psi(z) = Ae^{ik_0z} + Be^{-ik_0z}$ ....(1)

With

$k_0{}^2 = \frac{2mE}{h^2}$ ....(2)

Due to tunnelling ,the wave function extends inside the energy barrier of height Vo and thickness b.Therefore if -b<z<0;

$\psi(z) = Ce^{qz} + De^{-qz}$ ....(3)

Where the wave vector and the energy are related by

$Vo - E = \frac{h^2q^2}{2m}$ ....(4)

From the condition that both wave functions and their derivatives are continuous at z=0 and z= a

We get after operating

A+B = C+D

$ik_0(A - B) = q(C - D)$ .....(5)

According to the Bloch theorem, we can relate the wave functions at two different locations by

$\psi(a) = \psi(-b)e^{ik(a+b)}$ ....(6)

Where k is the wave vector corresponding to wave functions .Applying this equation to wave functions corresponding to the well and the barrier regions respectively, then

$Ae^{ik_0a} + Be^{-ik_0a} = (Ce^{-qb} + De^{qb})e^{ik_0(a+b)}$ ....(7)

$iko\left(Ae^{ik_0a} - Be^{-ik_0a}\right) = q(Ce^{-qb} - De^{qb})e^{ik_0(a+b)}$ ......(8)

The equations 5,6,7,8 for the amplitudes A,B,C and D have a solution only if the determinant of coefficients equals zero. After some calculations ,we get the relation

$$\frac{q-{k_0}^2}{2qk_0}\sin k_0a \sinh qb + \cos qa \cosh qb = \cos q(a+b) \text{ ………..(9).}$$

Let us assume the simple case of a = b and further that the effective mass of the electron is the same in the well and barrier materials. For the case of GaAs-AlGaAs and E < V_0, solution of Eq. (9) gives the values of allowed and forbidden energies.

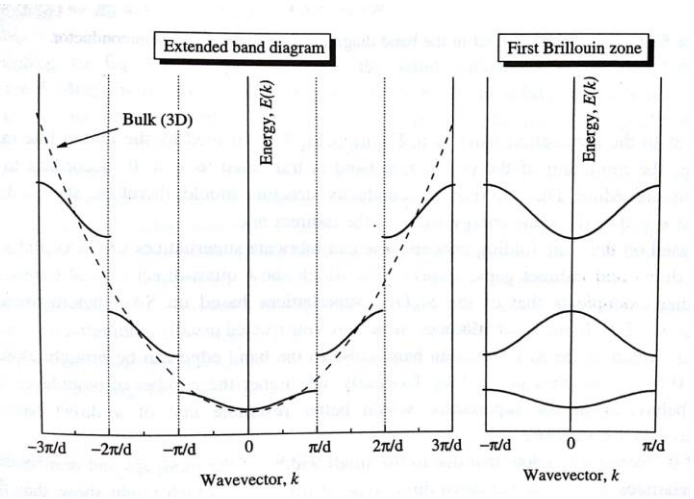


Figure 4.11: Extended band diagram of superlattice and band diagram in reduced first Brillouin zone

It is evident from Figure 4.11 that the general features of the E-k relationship in a superlattice, within the Kronig-Penney model are very similar to the case of electrons in solids if we replace the bulk lattice constant by the larger period of a superlattice. The free electron parabola, therefore, breaks down into several bands and gaps at the edges of the Brillouin zones$k = \pm\frac{n\pi}{d}$, as shown in the extended E-k diagram of Figme 4.11. Next, the portions of the bands can be translated to the reduced zone$(-\frac{\pi}{d} \leq z \leq +\frac{\pi}{d})$. Notice that everything occurs as if the superlattice potential folds the quasi-free energy band of a solid into the centre of the reduced zone, as it can be appreciated if we make the representation in the first Brillouin zone. Since, usually $d \gg a$, the breadth of the bands and gaps in a superlattice are much smaller, often receiving the names of minibands and minigaps. This band folding procedure is typical of superlattices and is called ***zone folding*** since it implies that the pieces of the band in the extended representation are zone-folded into the smaller zone with values of k: smaller than $\frac{2\pi}{d}$

The zone-folding effect has important consequences in the direct or indirect character of semiconductor structures.

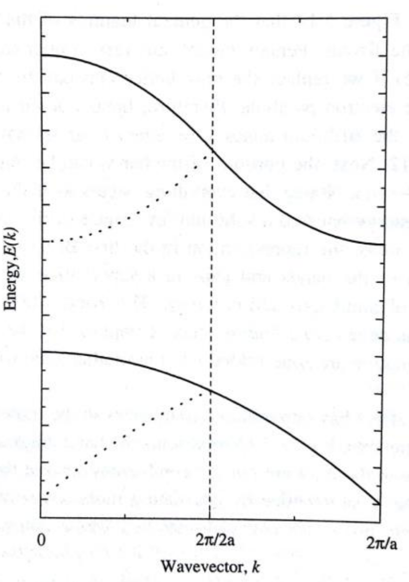


Figure 4.12: Zone folding effect in the band diagram of an indirect gap semiconductor

Figure 4.12 represents the band diagram of a typical indirect gap semiconductor with the minimum of the conduction band at the zone edge. Suppose next that we construct a superlattice by alternating monolayers of two semiconductors with similar electronic properties and well-matched lattice constants, but with one of them with the conduction band as in Figure 4.12. As indicated by the dotted line in the figure, the minimum of the conduction band is translated to k = 0, according to the folding procedure. The resulting semiconductor structure should, therefore, show a direct gap at k = 0 of the same energy value as the indirect one.

Based on the zone folding concept, one can fabricate superlattices of GaAs and AlAs with direct and indirect gaps, respectively, which show quasi-direct optical transitions. Another example is that of the Si_nGe_m superlattices based on SiGe heterostructures .In these superlattices, which are constructed usually symmetrical (n = m), the minimum of the Si conduction band close to the band edge can be brought close to k = 0 by successive zone-folding. Evidently, the higher the number of monolayers (n), the behaviour of the superlattice would better resemble that of a direct bandgap semiconductor structure.

EC 402 NANO ELECTRONICS

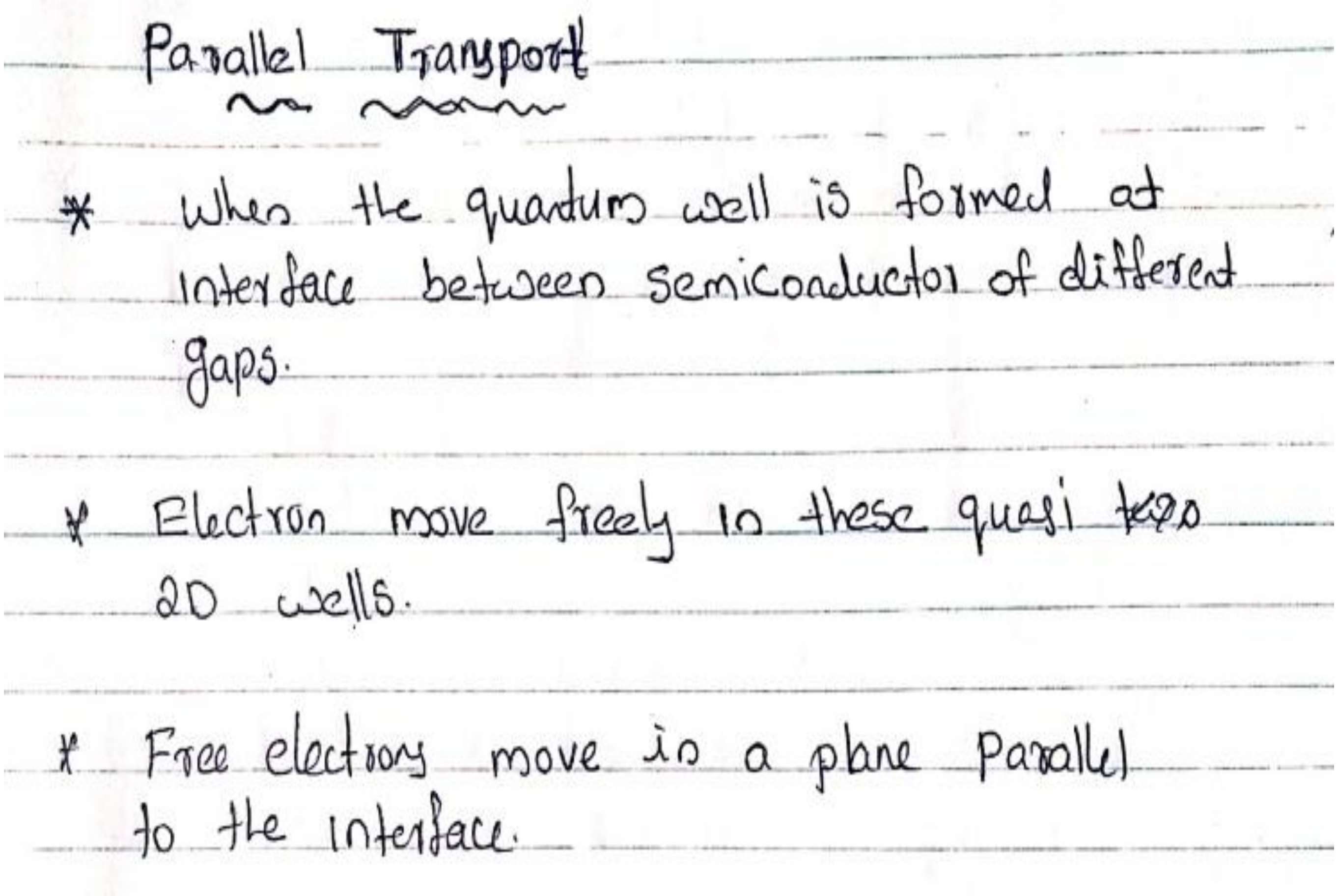
MODULE 5

PREPARED BY

Sajitha A S

Asst.Professor,ECE Dept

NCERC,PAMPADY

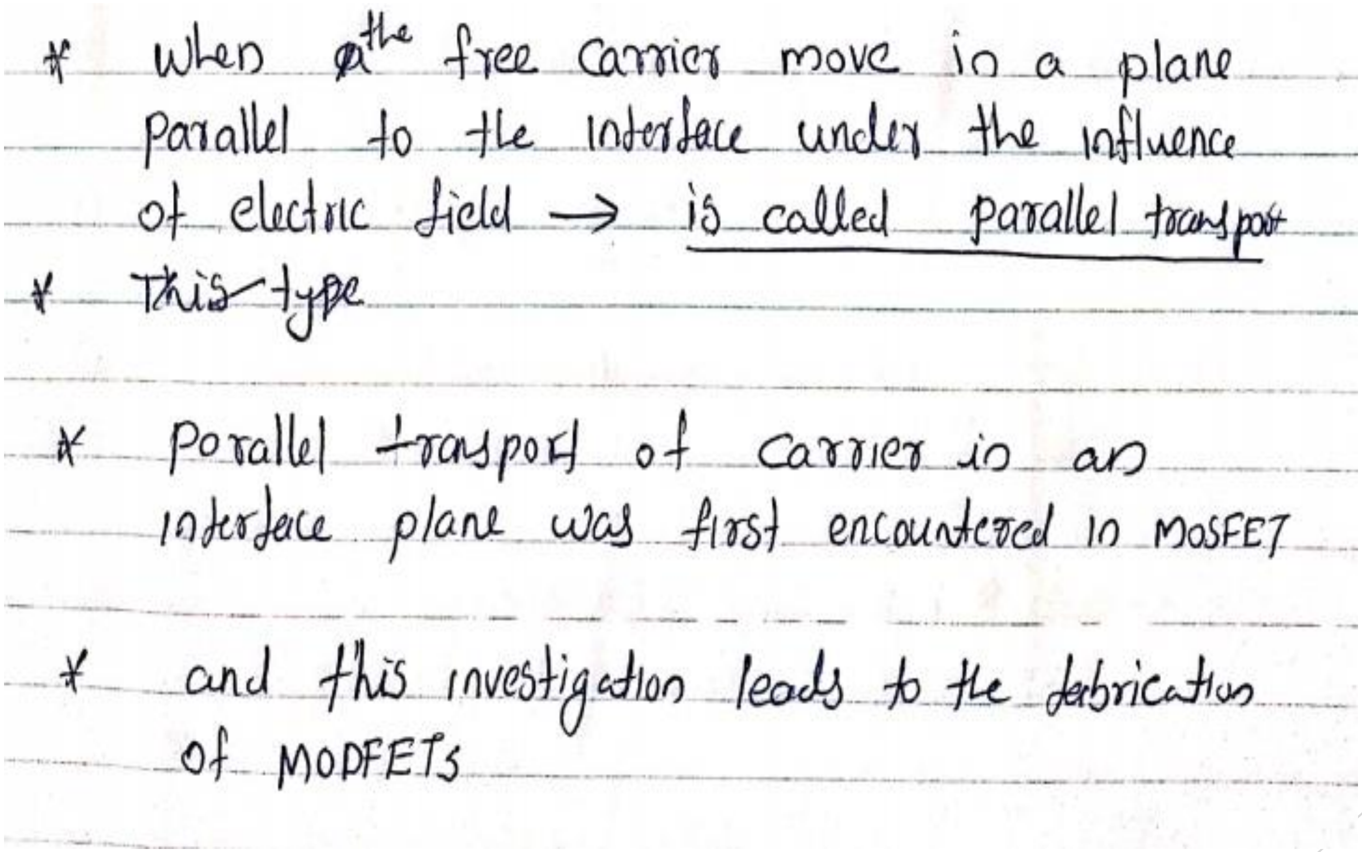


Parallel Transport

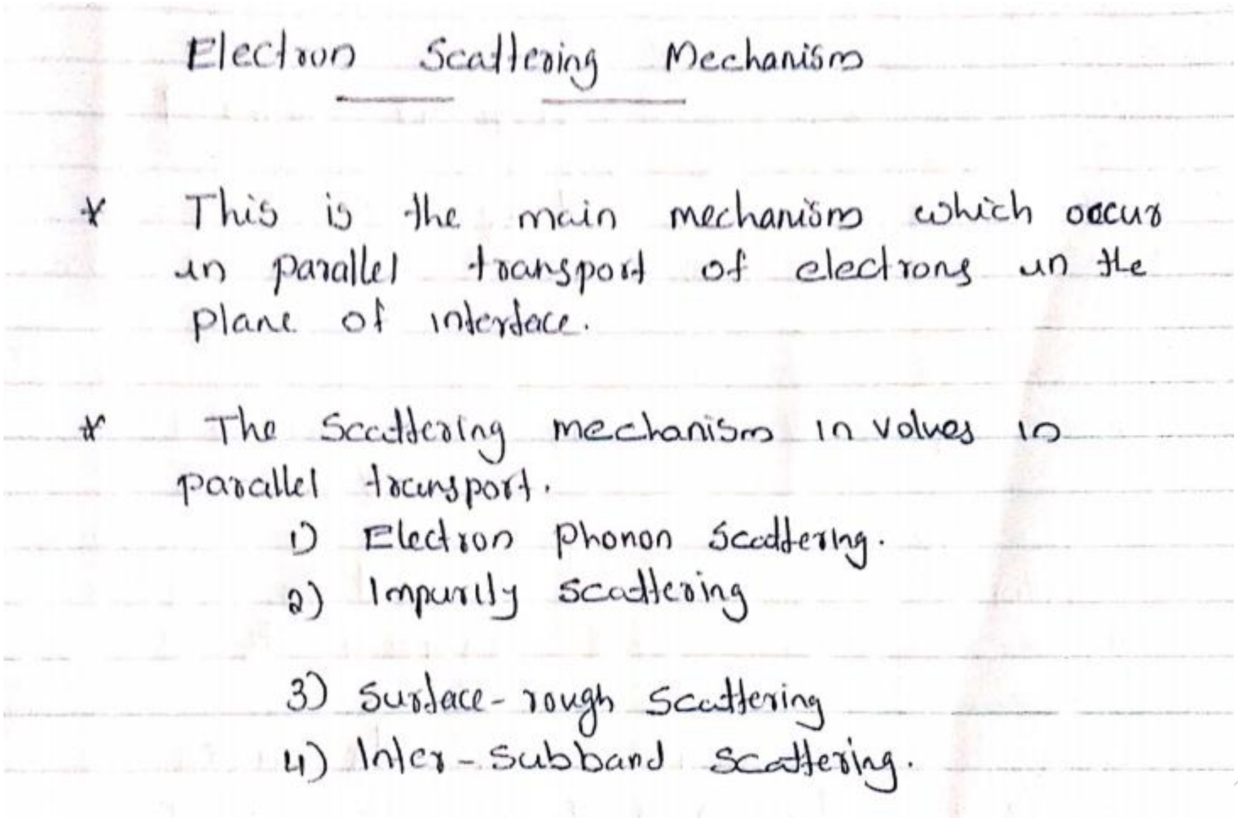
* When the quantum well is formed at interface between semiconductor of different gaps.

* Electron move freely in these quasi two 2D wells.

* Free electrons move in a plane parallel to the interface.



- When ~~a~~ the free carrier move in a plane parallel to the interface under the influence of electric field → <u>is called parallel transport</u>
- This type
- Parallel transport of carrier in an interface plane was first encountered in MOSFET
- and this investigation leads to the fabrication of MODFETs

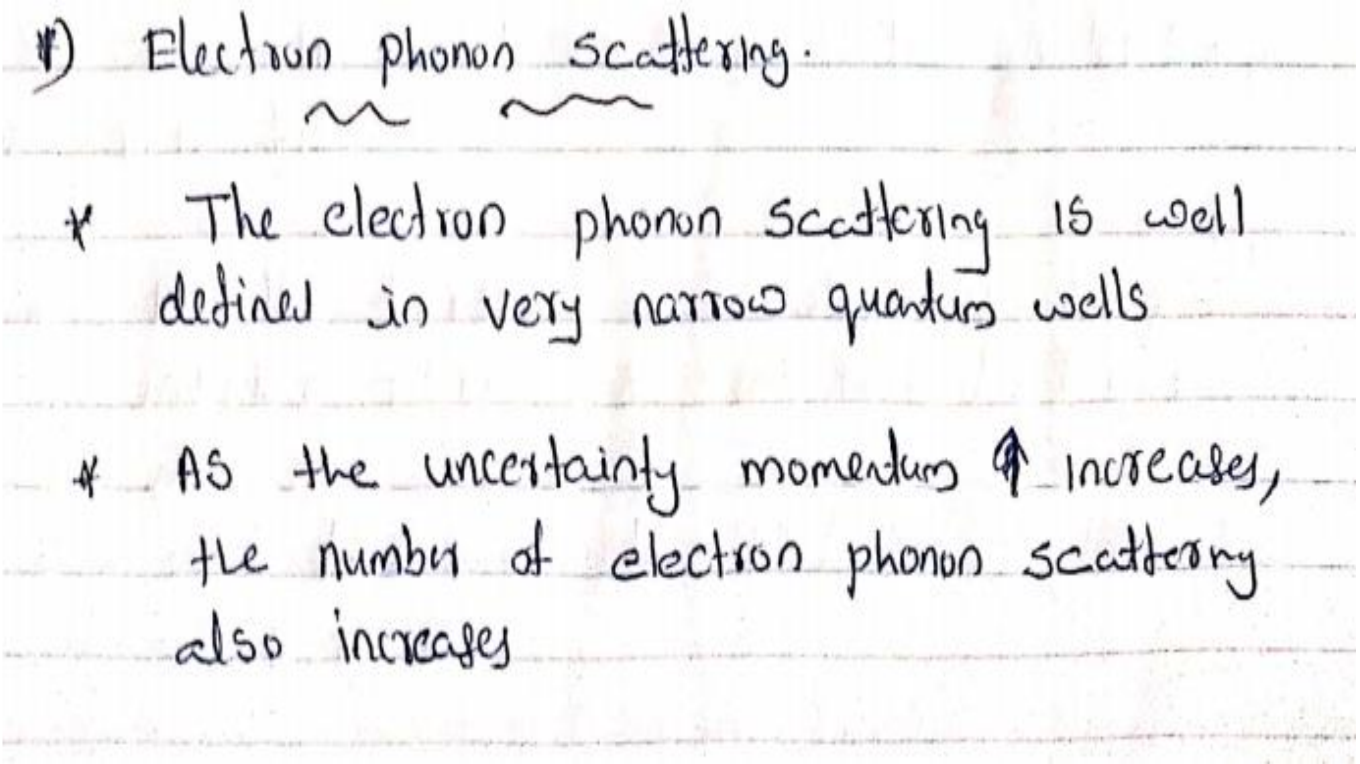


Electron Scattering Mechanism

* This is the main mechanism which occur in parallel transport of electrons in the plane of interface.

* The Scattering mechanism involves in parallel transport.

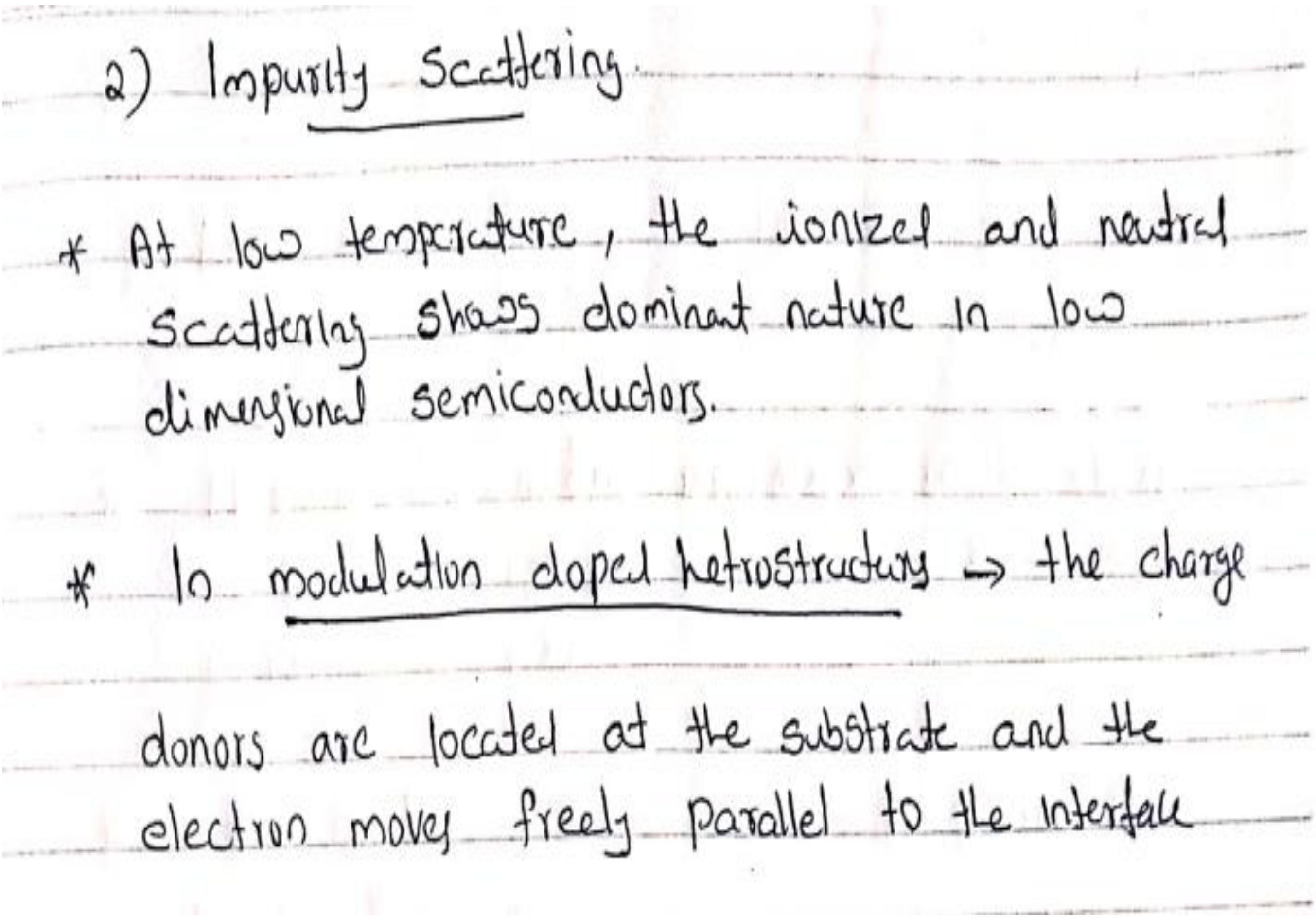
1) Electron Phonon Scattering.
2) Impurity scattering
3) Surface-rough Scattering
4) Inter-subband Scattering.



*) Electron phonon scattering.

* The electron phonon scattering is well defined in very narrow quantum wells

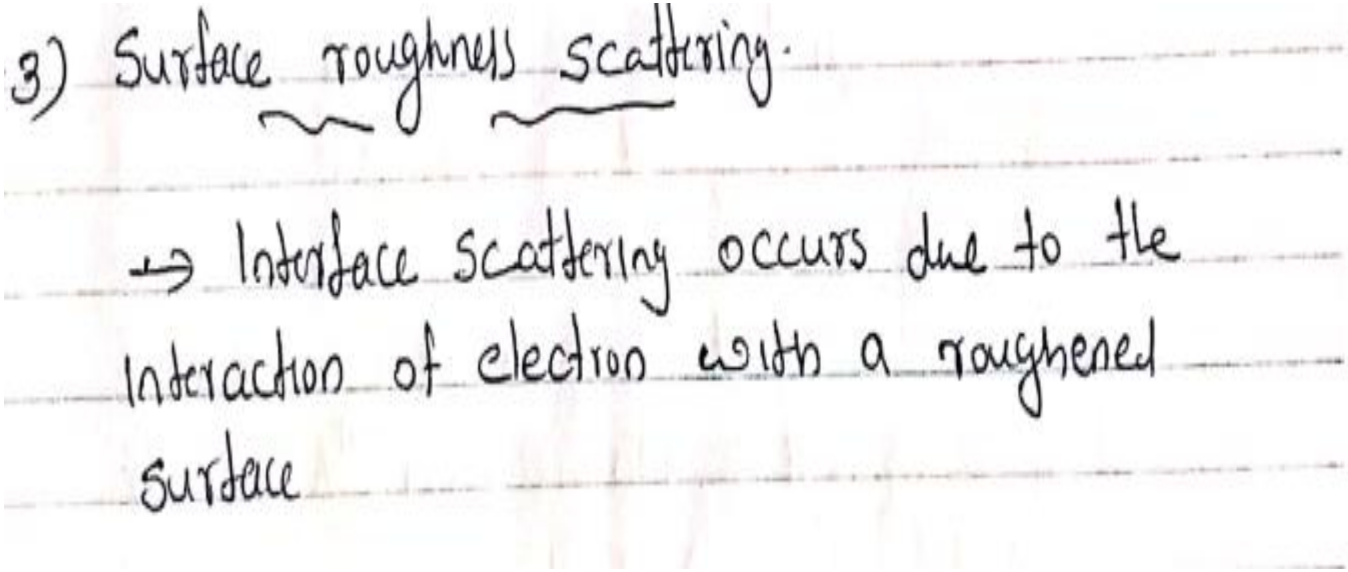
* As the uncertainty momentum ↑ increases, the number of electron phonon scattering also increases



2) Impurity Scattering.

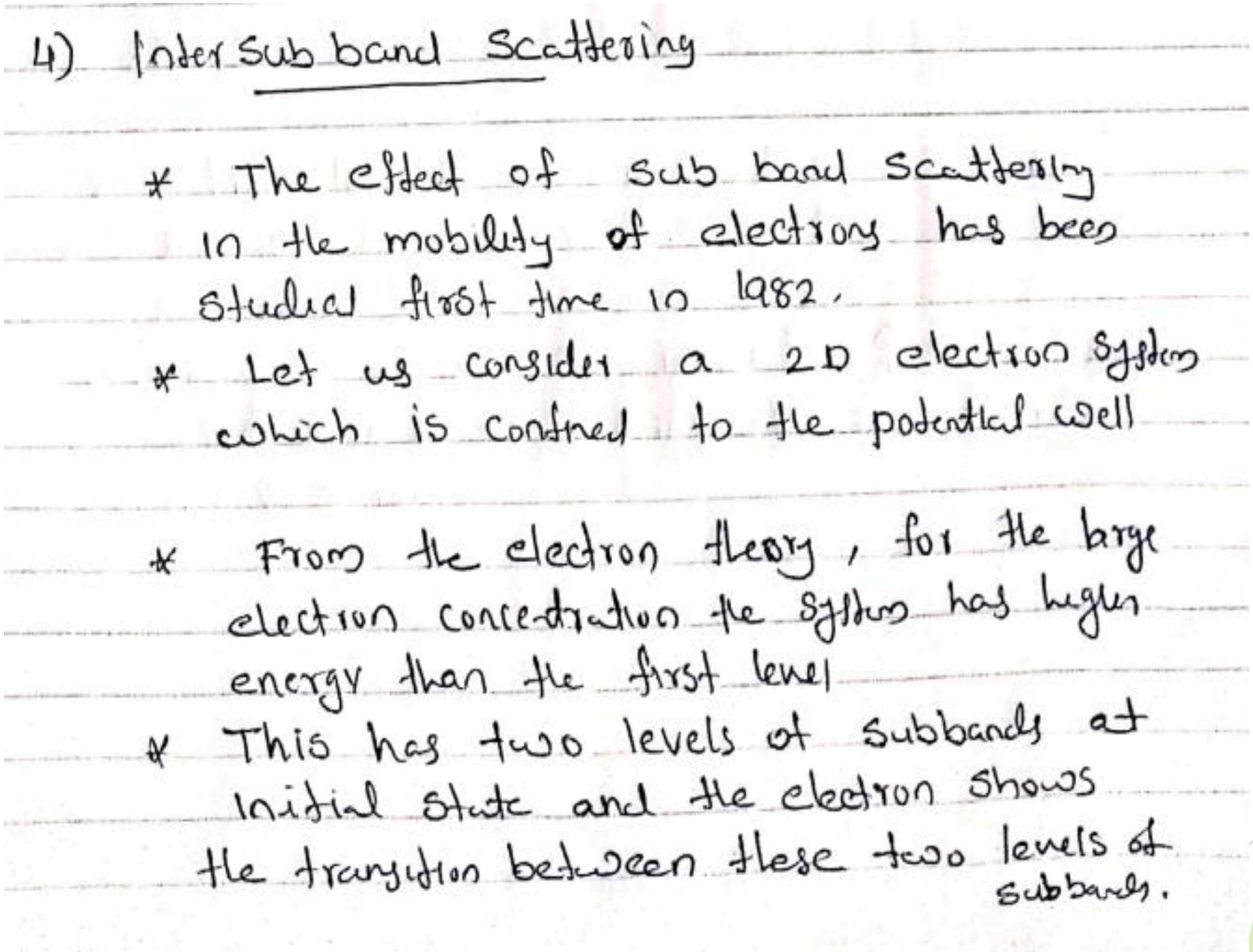
* At low temperature, the ionized and neutral scattering shows dominant nature in low dimensional semiconductors.

* In modulation doped hetrostructures → the charge donors are located at the substrate and the electron moves freely parallel to the interface



3) Surface roughness scattering.

→ Interface scattering occurs due to the interaction of electron with a roughened surface



4) Inter Sub band Scattering

- * The effect of sub band scattering in the mobility of electrons has been studied first time in 1982.
- * Let us consider a 2D electron system which is confined to the potential well
- * From the electron theory, for the large electron concentration the system has higher energy than the first level
- * This has two levels of subbands at initial state and the electron shows the transition between these two levels of subbands.

Hot electrons in parallel Transport

- In the presence of electric field, the electrons in some nano structures and some kind of FET are getting accelerated to a very high velocity.
- The energy of electron in presence of electric field in thermal equilibrium is given by KT
- When the electron energy distribution corresponds to the effective temperature becomes higher than the crystal lattice parameter due to the effect of high electric field then the received electron is termed a Hot electron

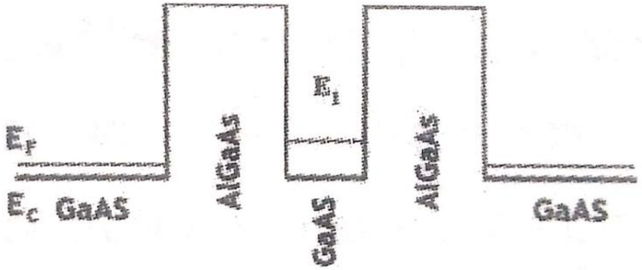
- In the case of parallel transport in AlGaAs/GaAs hetrostructure shows that the velocity under

PERPENDICULAR TRANSPORT

- Perpendicular transport are generally related to tunnelling or quantum transmission.
- In this phenomena the carrier flows perpendicular to the plane of the potential barrier.
- In tunnelling only few carriers can cross the potential barrier as it has insufficient energy and thus cannot cross the barrier height.

RESONANT TUNNELING

- Resonant tunnelling through a potential double barrier plays an important role, when electron diodes and transistor operates in high frequency.
- Let us consider a quantum well with GaAs surrounded by AlGaAs in each side as shown in figure.



- Let E1 the quantized energy level of the well
- Er is ferml energy level of the metallic contact
- Following figures shows the same structure under the application of voltage

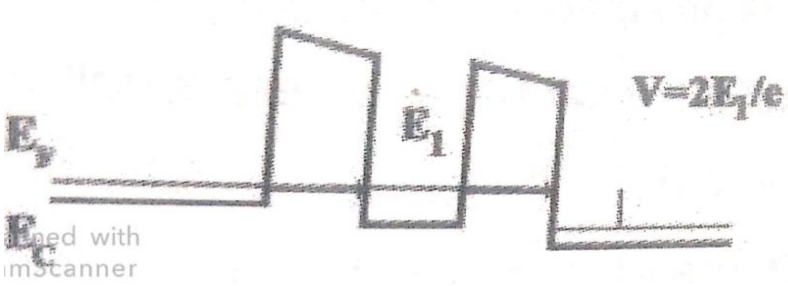


Figure: Under the application of Electric field

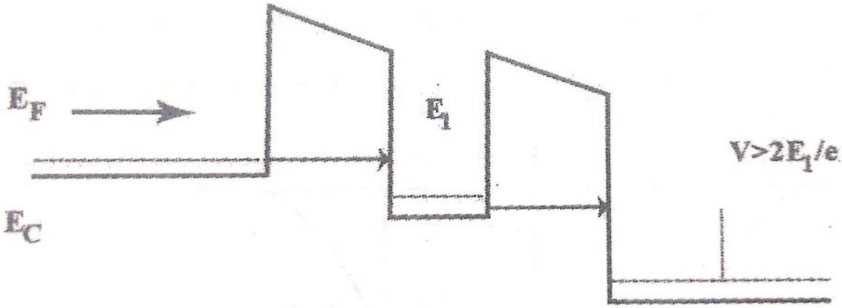


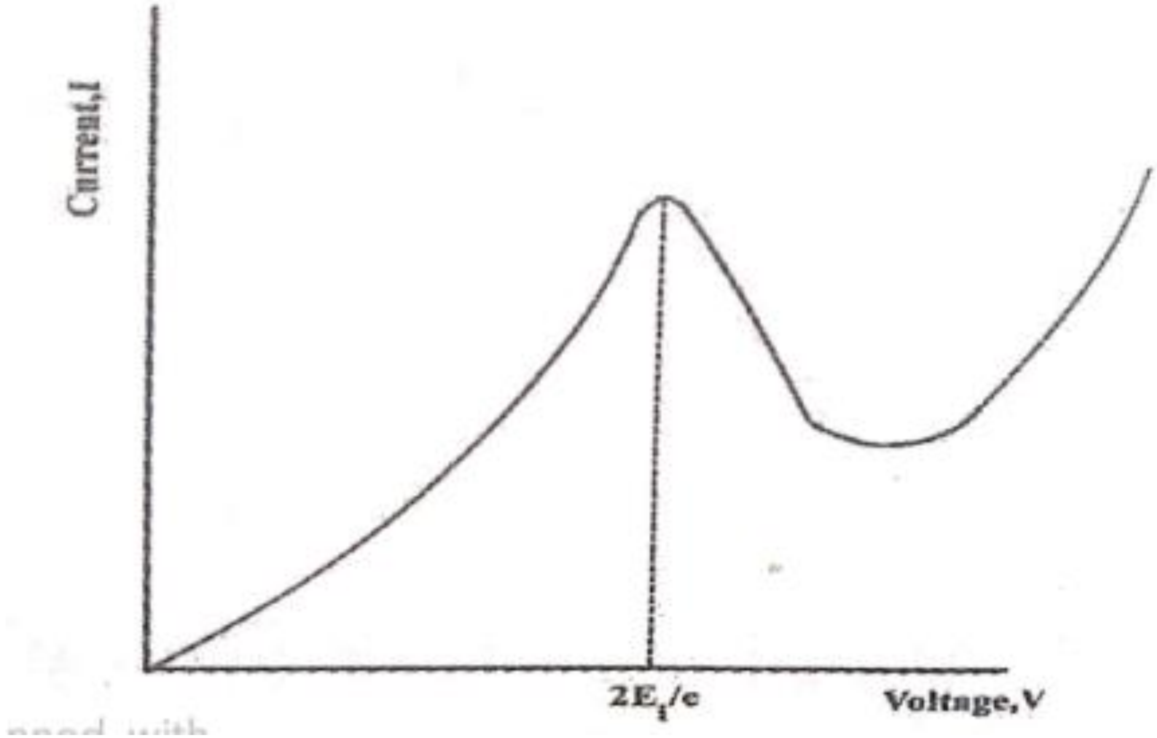
Figure: Under the application of Electric field when V> 2E/e

- The voltage when tunnelling occurs is given by

$$V_1 = \frac{2E}{e}$$

- Where E coincides with the quantized energy level E1.
- Here the Fermi energy level Ef of the metallic contact on the left coincides with n=1 level In the well.
- When the tunnelling transmission coefficient approaches to unity a large current will flow through the structure.
- When the voltage Increases above 2E1/e then the current decreases In the structure due to level mismatch.
- This mismatch shows that the double barrier has negative resistance.

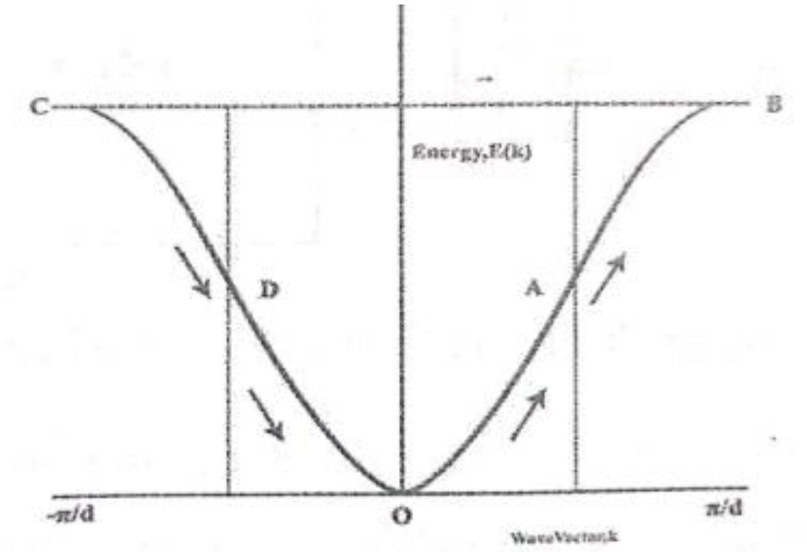
- Figure shows the variation of current as a function of voltage.



- In case of tunnelling, when there Is Increase in voltage, current also increase up to a peak value and then a valley and after that there is a rise.
- Here the tunnelling current increases with the applied voltage and reaches up to a maximum value.
- When the voltage matches the quantized energy level E1 of the well than the current shows its maximum value.
- When the voltage increases above E1 i.e.,V > 2E1/e the current start decreasing and the curve become negative and there exist a negative differential resistance region (NDR).

ELECTRIC FIELD EFFECT IN SUPER LATTICE

- In case of super lattice the electron is grouped in miniband which is very narrow, in comparison with bands in crystal.
- The width of the mini band is very small and it results in the larger dimensions of the super lattice period ‘d’.
- Under the influence of electric field, the electrons in the narrow band of super lattice observe Bloch Oscillation. Figure 5.5 below shows the motion of electrons in the miniband of super lattice which is similar to the motion of electron in K-space.



- Under the influence of electric field the equation of motion of electron 'in this band is given by:

$$h(dK/dt) = -eF$$

- where F is the applied Electric field. Let us assume that the electron is initially aligned at point 0 in the rest position.
- Also, from the figure the direction of the applied electric field is opposite to K. At this condition the electron starts moving from 0 towards A until it reaches to point B.
- At point B the velocity of electron is zero and therefore the electron is transferred to point C.
- From point C it moves towards D by applying electric field and completing one cycle in K-Space when electron reaches to O again.

- The velocity of electron in periodic motion is given by:

$$V = \frac{1}{h\left(\frac{dE}{dK}\right)}$$

- And the period of Bloch oscillation is given by:

$$T_B = \frac{2\pi h}{eFd}$$

where, e is the electronic charge and F is the applied electric field.

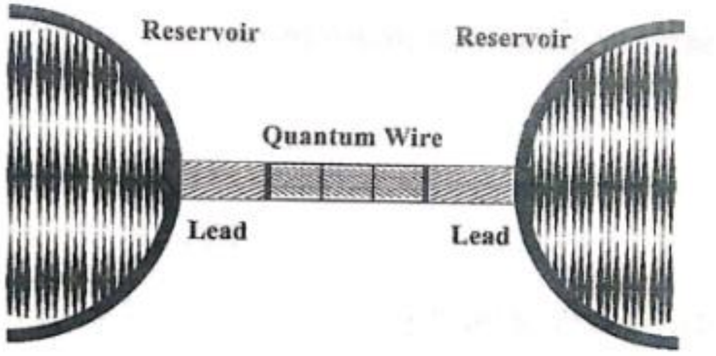
Bloch oscillation does not occur in bulk solid. In super lattice Bloch oscillation is observed only if Tb is shorter than the relaxing time due to scattering. To observe blotch oscillation the miniband should be quite narrow under the effect of electric field.

QUANTUM TRANSPORT IN NANOSTRUCTURE

- When the nanostructure is connected to an external electric field by the means of lead, a phenomena called *Quantum transport* takes place.
- The transport is also called as *Mesoscopic transport* and is referred to systems with a range of size between nano to atom or molecules.
- Quantization of conductance and Coulumb blockade in quantum dots can be observed in very small nanostructures.
- To observe quantum transport effect in semiconductor nanostructure, the effective mass of electron must be small at a given temperature.
- Hence smaller the effective mass, the higher the temperature at which quantum transport can be observed.

QUANTIZED CONDUCTANCE

- Let us consider a 1D mesoscopic semiconductor structure such as such as quantum wire for observing the quantum conductance effect.
- Consider the two reservoirs of electrons which is connected to a short quantum wire as shown in figure



- The length of quantum wire must be shorter than the electron mean path in the materials, so that there would be no scattering and the transport is ballistic.

- Let there will be two Fermi energy level E_{F1} and E_{F2} characterized by the reservoir of the electron.
- A small voltage v is applied in between the reservoir so that the current start flowing through the quantum wires.
- As a result there is a potential energy eV between the two reservoir which is given by,

$$eV = E_{F1} - E_{F2}$$

The current across the wire is given by

$$I = e\, n_{1D}(E)\, V(E) ev$$

Where, $n_{1D}(E)$ is the Density of State (DoS) function,

V(E) is the electron velocity.

By substituting the value of $n_{1D}(E)$ and E as,

$$n_{1D}(E) = \frac{1}{\pi h}\sqrt{\frac{2m_e}{E}} \qquad \& \qquad E = \frac{1}{2} m_e V^2$$

The current equation is given by,

$$I = \frac{2e^2}{h} \times v$$

Now the value of conductance is given by

$$G = \frac{I}{V}$$

$$G = \frac{2e^2}{h}$$

It is interesting to observe that the fundamental conductance of quantum wire is independent of length .The quotient of $G = \frac{e^2}{h}$ is called Quantum unit of conductance and its corresponding to quantum resistance has the value of $R = \frac{h}{e^2} = 25.2\Omega$

Parallel Transport

- when the quantum well is formed at interface between semiconductor of different gaps.

- Electron move freely in these quasi ~~two~~ 2D wells.

- Free electrons move in a plane parallel to the interface.

- when ~~a~~ the free carrier move in a plane parallel to the interface under the influence of electric field → is called parallel transport
- ~~This type~~

- Parallel transport of carrier in an interface plane was first encountered in MOSFET

- and this investigation leads to the fabrication of MODFETS

Electron Scattering Mechanism

* This is the main mechanism which occur in parallel transport of electrons in the plane of interface.

* The Scattering mechanism involves in parallel transport.

1) Electron phonon Scattering.
2) Impurity scattering

3) Surface-rough Scattering
4) Inter-subband Scattering.

1) Electron phonon scattering.

* The electron phonon scattering is well defined in very narrow quantum wells

* As the uncertainty momentum ↑ increases, the number of electron phonon scattering also increases

2) Impurity Scattering.

* At low temperature, the ionized and neutral scattering shows dominant nature in low dimensional semiconductors.

* In modulation doped hetrostructures → the charge donors are located at the substrate and the electron moves freely parallel to the interface

3) Surface roughness scattering.

→ Interface scattering occurs due to the interaction of electron with a roughened surface

4) Inter sub band Scattering

* The effect of sub band scattering in the mobility of electrons has been studied first time in 1982.
* Let us consider a 2D electron system which is confined to the potential well

* From the electron theory, for the large electron concentration the system has higher energy than the first level
* This has two levels of subbands at initial state and the electron shows the transition between these two levels of subbands.

Small sizes that matter:

Opportunities and risks of Nanotechnologies

Report in co-operation with the OECD International Futures Programme

OECD

Allianz

Contents

1. Executive Summary

Nanotechnologies are being spoken of as the driving force behind a new industrial revolution. Both private- and public-sector spending are constantly increasing. Spending on public research has reached levels of well over EUR 3 billion world-wide, but private sector spending is even faster—it is expected to exceed government spending in 2005. Nanotechnologies will be a major technological force for change in shaping Allianz's business environment across all industrial sectors in the foreseeable future and are likely to deliver substantial growth opportunities. The size of the market for nanotechnology products is already comparable to the biotechnology sector, while the expected growth rates over the next few years are far higher. At the same time, scientists have raised concerns that the basic building blocks of nanotechnologies—particles smaller than one billionth of a meter—pose a potential new class of risk to health and the environment. Allianz calls for a precautionary approach based on risk research and good risk management to minimize the likelihood of nanoparticles bringing a new dimension to personal injury and property damage losses or posing third party liability and product-recall risks.

The Allianz Center for Technology and Allianz Global Risks, in co-operation with the OECD International Futures Programme, has reviewed the likely economic impact, investment possibilities, and potential risks of nanotechnologies. This report analyses the opportunities and risks from the perspective of the Allianz Group. The opinions expressed in this report are those of the Allianz Group and do not engage the OECD or its Member governments.

1.1. Nanotechnology and the market place

The term nanotechnology describes a range of technologies performed on a nanometer scale with widespread applications as an enabling technology in various industries. Nanotechnology encompasses the production and application of physical, chemical, and biological systems at scales ranging from individual atoms or molecules to around 100 nanometers, as well as the integration of the resulting nanostructures into larger systems. The area of the dot of this ”i” alone can encompass 1 million nanoparticles.

What is different about materials on a nanoscale compared to the same materials in larger form is that, because of their relatively larger surface-area-to-mass ratio, they can become more chemically reactive and change their strength or other properties. Moreover, below 50 nm, the laws of classical physics give way to quantum effects, provoking different optical, electrical and magnetic behaviours.

Nanoscale materials have been used for decades in applications ranging from window glass and sunglasses to car bumpers and paints. Now, however, the convergence of scientific disciplines (chemistry, biology, electronics, physics, engineering etc.) is leading to a multiplication of applications in materials manufacturing, computer chips, medical diagnosis and health care, energy, biotechnology, space exploration, security and so on. Hence, nanotechnology is expected to have a significant impact on our economy and society within the next 10 to 15 years, growing in importance over the longer term as further scientific and technology breakthroughs are achieved.

It is this convergence of science on the one hand and growing diversity of applications on the other that is driving the potential of nanotechnologies. Indeed, their biggest impacts may arise from unexpected combinations of previously separate aspects, just as the internet and its myriad applications came about through the convergence of telephony and computing.

Sales of emerging nanotechnology products have been estimated by private research to rise from less than 0.1 % of global manufacturing output today to 15 % in 2014. These figures refer however to products ”incorporating nanotechnology” or ”manufactured using nanotechnology”. In many cases nanotechnology might only be a minor – but sometimes decisive -- contribution to the final product.

The first winners in the nanotechnology industry are likely to be the manufacturers of instruments allowing work on a nanoscale. According to market researchers, the nanotechnology tools industry ($245 million in the US alone) will grow by 30 % annually over the next few years.

The following projected three-phase growth path seems credible:

- In the present phase, nanotechnology is incorporated selectively into high-end products, especially in automotive and aerospace applications.
- Through 2009, commercial breakthroughs unlock markets for nanotechnology innovations. Electronics and IT applications dominate as microprocessors and memory chips built using new nanoscale processes come on to the market.

- From 2010 onwards, nanotechnology becomes commonplace in manufactured goods. Health care and life sciences applications finally become significant as nano-enabled pharmaceuticals and medical devices emerge from lengthy human trials.

1.2. Investments in nanotechnology

The financial sector will have a key role in transferring technology knowledge from the research centres to the industry and the markets. For the development of new products and processes and also for the penetration of new markets, sizeable investments are needed, especially in the seed phase. A closer co-operation between the financial community and nanotechnology companies can help to overcome these barriers.

By the end of 2004 venture capitalists had already invested $1 billion in nano companies, nearly half of that alone in 2003 and 2004. It is expected that most of these nanotechnology companies will be sold through trade sales.

For successful investments two aspects will be of critical importance: timing and target selection. Applying the process of ”technical due diligence” will be essential in making acquisitions.

The difficulty and expense involved in building up nanotechnology companies suggests that future winners in the sector will be well-funded companies and institutes that can attract and nurture the scientific and technical expertise needed to understand the problems and challenges. Moreover, the long lead times involved in moving from concept to commercialisation necessitate considerable long-term commitment to projects.

1.3. The environmental, health and safety discussion related to nanoparticles

Along with the discussion of their enormous technological and economic potential, a debate about new and specific risks related to nanotechnologies has started.
The catch-all term ”nanotechnology” is so broad as to be ineffective as a guide to tackling issues of risk management, risk governance and insurance.
A more differentiated approach is needed regarding all the relevant risk management aspects.

With respect to health, environmental and safety risks, almost all concerns that have been raised are related to free, rather than fixed manufactured nanoparticles. The risk and safety discussion related to free nanoparticles will be relevant only for a certain portion of the widespread applications of nanotechnologies.

Epidemiological studies on ambient fine and ultrafine particles incidentally produced in industrial processes and from traffic show a correlation between ambient air concentration and mortality rates. The health effects of ultrafine particles on respiratory and cardiovascular endpoints highlight the need for research also on manufactured nanoparticles that are intentionally produced.

In initial studies, manufactured nanoparticles have shown toxic properties. They can enter the human body in various ways, reach vital organs via the blood stream, and possibly damage tissue. Due to their small size, the properties of nanoparticles not only differ from bulk material of the same composition but also show different interaction patterns with the human body.
A risk assessment for bulk materials is therefore not sufficient to characterise the same material in nanoparticulate form.

The implications of the special properties of nanoparticles with respect to health and safety have not yet been taken into account by regulators. Size effects are not addressed in the framework of the new European chemicals policy REACH. Nanoparticles raise a number of safety and regulatory issues that governments are now starting to tackle. From Allianz’s perspective, a review of current legislation and continuous monitoring by the authorities is needed.

At present, the exposure of the general population to nanoparticles originating from dedicated industrial processes is marginal in relation to those produced and released unintentionally e.g. via combustion processes. The exposure to manufactured nanoparticles is mainly concentrated on workers in nanotechnology research and nanotechnology companies. Over the next few years, more and more consumers will be exposed to manufactured nanoparticles. Labelling requirements for nanoparticles do not exist. It is inevitable that in future manufactured nanoparticles will be released gradually and accidentally into the environment. Studies on biopersistence, bioaccumulation and ecotoxicity have only just started.

From Allianz’s perspective more funding for independent research on risk issues is necessary and we propose a dedicated research center at European level.

1.4. Allianz's position on industrial insurance cover

From an insurance perspective, several basic points define possible risk scenarios from nanoparticles:

- an increasingly high number of persons will be exposed,
- potential harmful effects are expected to evolve over longer periods of many years,
- in individual cases it will be difficult to establish a causal relationship between actions of a company and the resulting injury or damage,
- occupational exposure is a main concern,
- a certain closeness to major liability losses from the past will be evident.

In the absence of more basic evidence, all parties involved should take interim steps to manage risk.

The mechanisms that could lead to liability cases involve not only the development of our scientific understanding of the effects of nanoparticles, but also include legal and socio-economic developments that are difficult to foresee. More and more we realise that long-term illnesses are caused by a complex interaction of different risk factors. It is likely that nanoparticles will be not be so much a single cause or central origin of an illness but more of a contribution to a general health condition. In the traditional regime of liability and compensation, a causal relationship based on a one-to-one assignment of damaging agent and injury needs to be established. In the European legal framework, that causal relationship has to be proven – at least from today's perspective.

For Allianz it seems neither feasible nor appropriate to start a debate about a *general* exclusion of nanotechnologies from the commercial and industrial insurance cover today.

From the available evidence, we believe that the question is not whether or not nanotechnology risks can be controlled – and insured – but rather how they can best be managed and insured in a responsible way.

For a successful risk management of nanotechnologies from our perspective, the following framework is needed:

- sufficient funding of independent research on nanotechnology related risks with active steering by governments,
- transparency about and open access to the results of research activities,
- ongoing dialogue between insurers and commercial and industrial clients,
- international standards and nomenclature,
- adequate regulation of risk issues,
- a global risk governance approach.

Allianz's role is to meet client's demands while, at the same time, prudently protecting its balance sheet. We have started monitoring scientific, legal, social and economic trends in this field. We will constantly adapt our policy towards nanotechnologies as new evidence appears and, possibly, as claims in this field are made.

Given the fact that nanotechnologies have an enabling character and will penetrate almost every industry over the coming years, we expect that nanotechnology risks will be part of the industrial insurance portfolio. However, we will closely watch changes in the field. Where doubts arise, we would, where appropriate, talk with clients. We would also actively steer our portfolio. This might range from assessing the risk appetite for certain classes of business to posing questions about trigger of coverage and making detailed individual risk assessments.

Allianz wants to contribute to a dialogue-oriented approach using sustainability as both a vision and a yardstick of success.

2. What is nanotechnology and what makes it different?

2.1. Introduction

A nanometer (nm) is one thousand millionth of a meter. A single human hair is about 80,000 nm wide, a red blood cell is approximately 7,000 nm wide, a DNA molecule 2 to 2.5 nm, and a water molecule almost 0.3 nm. The term *”nanotechnology”* was created by Norio Taniguchi of Tokyo University in 1974 to describe the precision manufacture of materials with nanometer tolerances[1], but its origins date back to Richard Feynman’s 1959 talk ”There’s Plenty of Room at the Bottom”[2] in which he proposed the direct manipulation of individual atoms as a more powerful form of synthetic chemistry. Eric Drexler of MIT expanded Taniguchi’s definition and popularised nanotechnology in his 1986 book ”Engines of Creation: The Coming Era of Nanotechnology”[3]. On a nanoscale, i.e. from around 100nm down to the size of atoms (approximately 0.2 nm) the properties of materials can be very different from those on a larger scale. Nanoscience is the study of phenomena and manipulation of materials at atomic, molecular and macromolecular scales, in order to understand and exploit properties that differ significantly from those on a larger scale. Nanotechnologies are the design, characterisation, production and application of structures, devices and systems by controlling shape and size on ananometer scale.

Modern industrial nanotechnology had its origins in the 1930s, in processes used to create silver coatings for photographic film; and chemists have been making polymers, which are large molecules made up of nanoscale subunits, for many decades. However, the earliest known use of nanoparticles is in the ninth century during the Abbasid dynasty. Arab potters used nanoparticles in their glazes so that objects would change colour depending on the viewing angle (the so-called polychrome lustre)[4]. Today’s nanotechnology, i.e. the planned manipulation of materials and properties on a nanoscale, exploits the interaction of three technological streams[5]:

1. new and improved control of the size and manipulation of nanoscale building blocks
2. new and improved characterisation of materials on a nanoscale (e.g., spatial resolution, chemical sensitivity)
3. new and improved understanding of the relationships between nanostructure and properties and how these can be engineered.

---

[1] ”Nano-technology' mainly consists of the processing of separation, consolidation, and deformation of materials by one atom or one molecule.
” N. Taniguchi, "On the Basic Concept of 'Nano-Technology'," Proc. Intl. Conf. Prod. Eng. Tokyo, Part II, Japan Society of Precision Engineering, 1974.

[2] A transcript of the classic talk that Richard Feynman gave on December 29th 1959 at the annual meeting of the American Physical Society at the California Institute of Technology (Caltech) was first published in the February 1960 issue of Caltech's Engineering and Science which owns the copyright. It has been made available on the web at http://www.zyvex.com/nanotech/feynman.html with their kind permission.

[3] Engines of Creation was originally published in hardcover by Anchor Books in 1986, and in paperback in 1987. The web version published here http://www.foresight.org/EOC/ was reprinted and adapted by Russell Whitaker with permission of the copyright holder.

[4] ”The oldest known nanotechnology dates back to the ninth century” New Materials International, March 2004 http://www.newmaterials.com/news/680.asp

[5] US National Science and Technology Council, Committee on Technology, Interagency Working Group on NanoScience, Engineering and Technology: ”Nanostructure Science and Technology, A Worldwide Study”. September 1999. http://www.wtec.org/loyola/nano/

The properties of materials can be different on a nanoscale for two main reasons. First, nanomaterials have, relatively, a larger surface area than the same mass of material produced in a larger form. This can make materials more chemically reactive (in some cases materials that are inert in their larger form are reactive when produced in their nanoscale form), and affect their strength or electrical properties. Second, below 50 nm, the laws of classical physics give way to quantum effects, provoking optical, electrical and magnetic behaviours different from those of the same material at a larger scale. These effects can give materials very useful physical properties such as exceptional electrical conduction or resistance, or a high capacity for storing or transferring heat, and can even modify biological properties, with silver for example becoming a bactericide on a nanoscale. These properties, however, can be very difficult to control. For example, if nanoparticles touch each other, they can fuse, losing both their shape and those special properties—such as the magnetism—that scientists hope to exploit for a new generation of microelectronic devices and sensors.

On a nanoscale, chemistry, biology, electronics, physics, materials science, and engineering start to converge and the distinctions as to which property a particular discipline measures no longer apply. All these disciplines contribute to understanding and exploiting the possibilities offered by nanotechnology, but if the basic science is converging, the potential applications are infinitely varied, encompassing everything from tennis rackets to medicines to entirely new energy systems. This twin dynamic of convergent science and multiplying applications means that nanotechnology’s biggest impacts may arise from unexpected combinations of previously separate aspects, just as the internet came about through the convergence of telephony and computing.

2.2. Nanomaterials: basic building blocks

This section outlines the properties of three of the most talked-about nanotechnologies: carbon nanotubes, nanoparticles, and quantum dots[6].

Carbon Nanotubes

Carbon nanotubes, long thin cylinders of atomic layers of graphite, may be the most significant new material since plastics and are the most significant of today's nanomaterials. They come in a range of different structures, allowing a wide variety of properties. They are generally classified as single-walled (SWNT), consisting of a single cylindrical wall, or multiwalled nanotubes (MWNT), which have cylinders within the cylinders. When the press mentions the amazing properties of nanotubes, it is generally SWNT they are referring to. The following table summarises the main properties of SWNT:

Size	0.6 to 1.8 nanometer in diameter	Electron beam lithography can create lines 50 nm wide, a few nm thick
Density	1.33 to 1.40 grams per cubic centimeter	Aluminium has a density of 2.7g/cm^3
Tensile Strength	45 billion pascals	High-strength steel alloys break at about 2 billion Pa
Resilience	Can be bent at large angles and restraightened without damage	Metals and carbon fibers fracture at grain boundaries
Current Carrying Capacity	Estimated at 1 billion amps per square centimeter	Copper wires burn out at about 1 million A/cm^2
Field Emission	Can activate phosphors at 1 to 3 volts if electrodes are spaced 1 micron apart	Molybdenum tips require fields of 50 to 100 V/µm and have very limited lifetimes
Heat Transmission	Predicted to be as high as 6,000 watts per meter per kelvin at room temperature	Nearly pure diamond transmits 3,320 W/m·K
Temperature Stability	Stable up to 2,800 degrees Celsius in vacuum, 750 degrees C in air	Metalwires in microchips melt at 600 to 1,000 degrees C

© Scientific American December 2000[7]

---

[6] Nanotechnology white papers published by Cientifica at http://www.cientifica.com/html/Whitepapers/whitepapers.htm were particularly useful for this section. Registration (free) is required to consult the documents.

[7] Philip G. Collins and Phaedon Avouris "Nanotubes for electronics" Scientific American, December 2000, page 69

However, SWNT are more difficult to make than MWNT, and confusion arises about the quantities of nanotubes actually being manufactured. Carbon Nanotechnologies of Houston, one of the world's leading producers, only makes up to 500g per day. One problem is that economies of scale are practically impossible with today's production technologies – the machines used to manufacture the tubes cannot be scaled up, so producing bigger quantities means using more machines.

Another drawback is that it is difficult to make nanotubes interact with other materials. For example, to fully exploit their strength in composite materials, nanotubes need to be "attached" to a polymer. They are chemically modified to facilitate this (a process known as "functionalization"), but this process reduces the very properties the nanotubes are being used for. In the long-term, the ideal solution would be to use pure nanomaterials, e.g. nanotubes spun into fibers of any desired length, but such a development is unlikely in the next couple of decades unless a radically more efficient production process is developed.

The most promising applications of nanotubes may be in electronics and optoelectronics[8]. Today, the electronics industry is producing the vital components known as MOSFETs (metal oxide semiconductor field-effect transistors) with critical dimensions of just under 100 nm, with half that size projected by 2009 and 22 nm by 2016. However, the industry will then encounter technological barriers and fundamental physical limitations to size reduction. At the same time, there are strong financial incentives to continue the process of scaling, which has been central in the effort to increase the performance of computing systems in the past. A new microchip manufacturing plant costs around $1.5 billion, so extending the technology's life beyond 2010 is important. One approach to overcoming the impending barriers while preserving most of the existing technology, is to use new materials.

With carbon nanotubes, it is possible to get higher performance without having to use ultra thin silicon dioxide gate insulating films. In addition, semiconducting SWNTs, unlike silicon, directly absorb and emit light, thus possibly enabling a future optoelectronics technology. The SWNT devices would still pose manufacturing problems due to quantum effects at the nanoscale, so the most likely advantage in the foreseeable future is that carbon nanotubes will allow a simpler fabrication of devices with superior performance at about the same length as their scaled silicon counterparts.

Other proposed uses for nanotubes:

Chemical and Genetic Probes A nanotube-tipped atomic force microscope can trace a strand of DNA and identify chemical markers that reveal which of several possible variants of a gene is present in the strand. This is the only method yet invented for imaging the chemistry of a surface, but it is not yet used widely. So far it has been used only on relatively short pieces of DNA.

Mechanical memory (nonvolatile RAM). A screen of nanotubes laid on support blocks has been tested as a binary memory device, with voltages forcing some tubes to contact (the "on" state) and others to separate ("off"). The switching speed of the device was not measured, but the speed limit for a mechanical memory is probably around one megahertz, which is much slower than conventional memory chips.

Field Emission Based Devices. Carbon Nanotubes have been demonstrated to be efficient field emitters and are currently being incorporated in several applications including flat-panel display for television sets or computers or any devices requiring an electron producing cathode such as X-ray sources (e.g. for medical applications).

Nanotweezers. Two nanotubes, attached to electrodes on a glass rod, can be opened and closed by changing voltage. Such tweezers have been used to pick up and move objects that are 500 nm in size. Although the tweezers can pick up objects that are large compared with their width, nanotubes are so sticky that most objects can't be released. And there are simpler ways to move such tiny objects.

Supersensitive Sensors. Semiconducting nanotubes change their electrical resistance dramatically when exposed to alkalis, halogens and other gases at room temperature, raising hopes for better chemical sensors. The sensitivity of these devices is 1,000 times that of standard solid state devices.

---

[8] Phaedon Avouris and Joerg Appenzeller: "Electronics and optoelectronics with carbon nanotubes: New discoveries brighten the outlook for innovative technologies" The Industrial Physicist, June/July 2004, American Institute of Physics

Hydrogen and Ion Storage. Nanotubes might store hydrogen in their hollow centers and release it gradually in efficient and inexpensive fuel cells. They can also hold lithium ions, which could lead to longer-lived batteries. So far the best reports indicate 6.5 percent hydrogen uptake, which is not quite dense enough to make fuel cells economical. The work with lithium ions is still preliminary.

Sharper Scanning Microscope. Attached to the tip of a scanning probe microscope, nanotubes can boost the instruments' lateral resolution by a factor of 10 or more, allowing clearer views of proteins and other large molecules. Although commercially available, each tip is still made individually. The nanotube tips don't improve vertical resolution, but they do allow imaging deep pits in nanostructures that were previously hidden.

Superstrong Materials. Embedded into a composite, nanotubes have enormous resilience and tensile strength and could be used to make materials with better safety features, such as cars with panels that absorb significantly more of the force of a collision than traditional materials, or girders that bend rather than rupture in an earthquake. Nanotubes still cost 10 to 1,000 times more than the carbon fibers currently used in composites. And nanotubes are so smooth that they slip out of the matrix, allowing it to fracture easily.

There are still many technical obstacles to overcome before carbon nanotubes can be used on an industrial scale, but their enormous potential in a wide variety of applications has made them the "star" of the nanoworld[9] and encouraged many companies to commit the resources needed to ensure that the problems will be solved. Fujitsu, for example, expects to use carbon nanotubes in 45 nm chips by 2010 and in 32 nm devices in 2013[10].

Nanoparticles

Nanoparticles have been used since antiquity by ceramists in China and the West, while 1.5 million tons of carbon black, the most abundant nanoparticulate material, are produced every year. But, as mentioned earlier, nanotechnology is defined as knowingly exploiting the nanoscale nature of materials, thereby excluding these examples. Although metal oxide ceramic, metal, and silicate nanoparticles constitute the most common of the new generation of nanoparticles, there are others too. A substance called chitosan for example, used in hair conditioners and skin creams, has been made in nanoparticle form to improve absorption.

Moving to nanoscale changes the physical properties of particles, notably by increasing the ratio of surface area to volume, and the emergence of quantum effects. High surface area is a critical factor in the performance of catalysis and structures such as electrodes, allowing improvement in performance of such technologies as fuel cells and batteries. The large surface area also results in useful interactions between the materials in nanocomposites, leading to special properties such as increased strength and/or increased chemical/heat resistance. The fact that nanoparticles have dimensions below the critical wavelength of light renders them transparent, an effect exploited in packaging, cosmetics and coatings.

Quantum dots

Just as carbon nanotubes are often described as the new plastics, so quantum dots are defined as the ball bearings of the nano-age[11]. Quantum dots are like "artificial atoms". They are 1 nm structures made of materials such as silicon, capable of confining a single electron, or a few thousand, whose energy states can be controlled by applying a given voltage. In theory, this could be used to fulfil the alchemist's dream of changing the chemical nature of a material, making lead emulate gold, for example.

One more likely set of possible applications exploits the fact that quantum dots can be made to emit light at different wavelengths, with the smaller the dot the bluer the light. The dots emit over a narrow spectrum making them well suited to imaging, particularly for biological samples. Currently, biological molecules are imaged using naturally fluorescent molecules, such as organic dyes, with a different dye attached to each kind of molecule in a sample. But the dyes emit light over a broad range of wavelengths, which means that their spectra overlap and only about three different

---

[9] "Nanotubes 2003" http://www.researchandmarkets.com/reports/220255/220255.htm

[10] Paul Kallender "Fujitsu touts carbon nanotubes for chip wiring" IDG News Service, March 2005 http://www.infoworld.com/article/05/03/01/HNfujinanotubes_1.html

[11] Paul O'Brien Physics World, December 2003. A summary is available here: http://www.nanotechweb.org/articles/feature/2/12/1/1

dyes emit light over a broad range of wavelengths, which means that their spectra overlap and only about three different dyes can be used at the same time. With quantum dots, full-colour imaging is possible because large numbers of dots of different sizes can be excited by a light source with a single wavelength.

The wide range of colors that can be produced by quantum dots also means they have great potential in security. They could, for example, be hidden in bank notes or credit cards, producing a unique visible image when exposed to ultraviolet light.

It is possible to make light-emitting diodes (LEDs) from quantum dots which could produce white light e.g. for buildings or cars. By controlling the amount of blue in the emission-control the "flavor" or "tone" of the white light can be tuned.

Quantum dots are also possible materials for making ultrafast, all-optical switches and logic gates that work faster than 15 terabits a second. For comparison, the Ethernet generally can handle only 10 megabits per second. Other possible applications are all-optical demultiplexers (for separating various multiplexed signals in an optical fiber), all-optical computing, and encryption, whereby the spin of an electron in a quantum dot represent a quantum bit or qubit of information.

Biologists are experimenting with composites of living cells and quantum dots. These could possibly be used to repair damaged neural pathways or to deliver drugs by activating the dots with light.

Once again, significant advances in manufacturing will be needed to realise the potential of quantum dots. For example, the quantum state needed to make a quantum computer is relatively easily to create, but its behavior is still unpredictable.

2.3. Nano tools and fabrication techniques

Legend has it that the people who profited most from the Klondike gold rush at the end of the 19th century were the sellers of picks and shovels. The same may hold true for nanotechnology, at least in the coming decade before production techniques are improved. According to market researchers Freedonia, the $245 million nanotech tools industry will grow by 30 % annually over the next few years[12]. Microscopes and related tools dominate now, but measurement, fabrication / production and simulation / modelling tools will grow the fastest. Electronics and life sciences markets will emerge first; industrial, construction, energy generation and other applications will arise later.

Microscopy

Nanotechnology uses two main kinds of microscopy. The first involves a stationary sample in line with a high-speed electron gun. Both the scanning electron microscope (SEM) and transmission electron microscope (TEM) are based on this technique. The second class of microscopy involves a stationary scanner and a moving sample. The two microscopes in this class are the atomic force microscope (AFM) and the scanning tunnelling microscope (STM).

Microscopy plays a paradoxical role in nanotechnology because, although it is the key to understanding materials and processes, on a nanoscale samples can be damaged by the high-energy electrons fired at them. This is not a problem with STM, but a further drawback is that most microscopes require very stringent sample preparation. The SEM, TEM, and STM need well prepared samples that are also electrically conductive. There are ways to get around this, but the fact remains that it can take hours to prepare and mount a sample correctly (and hours to actually synthesise the sample).

Top-down and bottom-up synthesis techniques

There are two approaches to building nanostructures, both having their origins in the semiconductor industry[13]. In the traditional ”top-down” approach a

---

[12] ”Nanotech Tools to 2008” August 2004 http://freedonia.ecnext.com/coms2/summary_0285-21108_ITM

[13] Materials Research Society Bulletin, July 2001, special focus on Emerging Methods of Micro- and Nanofabrication

larger material such as a silicon wafer is processed by removing matter until only the nanoscale features remain. Unfortunately, these techniques require the use of lithography, which requires a mask that selectively protects portions of the wafer from light. The distance from the mask to the wafer, and the size of the slit define the minimum feature size possible for a given frequency of light, e.g. extreme ultraviolet light yields feature sizes of 90 nm across, but this scale is near the fundamental limit of lithography. Nonetheless lithography can be used for patterning substrates used to produce nanomaterials, e.g. guiding the growth of quantum dots and nanowires.

The ”bottom-up” approach starts with constituent materials (often gases or liquids) and uses chemical, electrical, or physical forces to build a nanomaterial atom-by-atom or molecule-by-molecule[14]. The simplest bottom up synthesis route is electroplating to create a material layer-by-layer, atom-by-atom. By inducing an electric field with an applied voltage, charged particles are attracted to the surface of a substrate where bonding will occur. Most nanostructured metals with high hardness values are created with this approach. Chemical vapour deposition (CVD) uses a mix of volatile gases and takes advantage of thermodynamic principles to have the source material migrate to the substrate and then bond to the surface. This is the one proven method for creating nanowires and carbon nanotubes, and is a method of choice for creating quantum dots. Molecular self-assembly promises to be a revolutionary new way of creating materials from the bottom up. One way to achieve self-assembly is to use attractive forces like static electricity, Van der Waals forces, and a variety of other short-range forces to orient constituent molecules in a regular array. This has proven very effective in creating large grids of quantum dots.

The bottom up approach promises an unheard-of level of customisability in materials synthesis, but controlling the process is not easy and can only produce simple structures, in time-consuming processes with extremely low yields. It is not yet possible to produce integrated devices from the bottom up, and any overall order aside from repeating grids cannot be done without some sort of top-down influence like lithographic patterning. Nanotechnology synthesis is thus mainly academic, with only a few companies in the world that can claim to be nanotechnology manufacturers. And until understanding of synthesis is complete, it will be impossible to reach a point of mass production.

2.4. Present and future areas of application

What is nanotechnology already used for[15]?

Nanoscale materials, as mentioned above, have been used for many decades in several applications, are already present in a wide range of products, including mass-market consumer products. Among the most well-known are a glass for windows which is coated with titanium oxide nanoparticles that react to sunlight to break down dirt. When water hits the glass, it spreads evenly over the surface, instead of forming droplets, and runs off rapidly, taking the dirt with it. Nanotechnologies are used by the car industry to reinforce certain properties of car bumpers and to improve the adhesive properties of paints. Other uses of nanotechnologies in consumer products include:

Sunglasses using protective and antireflective ultrathin polymer coatings. Nanotechnology also offers scratch-resistant coatings based on nanocomposites that are transparent, ultra-thin, simple to care for, well-suited for daily use and reasonably priced.

Textiles can incorporate nanotechnology to make practical improvements to such properties as windproofing and waterproofing, preventing wrinkling or staining, and guarding against electrostatic discharges. The windproof and waterproof properties of one ski jacket, for example, are obtained not by a surface coating of the jacket but by the use of nanofibers. Given that low-cost countries are capturing an ever-increasing share of clothes manufacturing, high-cost regions are likely to focus on high-tech clothes with the additional benefits for users that nanotech can help implement. Future projects include clothes with additional electronic functionalities, so-called ”smart clothes” or ”wearable electronics”. These could include sensors to monitor body functions or release

---

[14] A layperson’s guide to fabrication techniques can be found here, in the section called ”Synthesis” http://www.ringsurf.com/info/Technology_/Nanotechnology/

[15] ”Current Consumer Products using Nanotechnology” http://www.azonano.com/details.asp?ArticleID=1001

drugs in the required amounts, self-repairing mechanisms or access to the Internet.

Sports equipment manufacturers are also turning to nanotech. A high-performance ski wax, which produces a hard and fast-gliding surface, is already in use. The ultra-thin coating lasts longer than conventional waxing systems. Tennis rackets with carbon nanotubes have increased torsion and flex resistance. The rackets are more rigid than current carbon rackets and pack more power. Long-lasting tennis-balls are made by coating the inner core with clay polymer nanocomposites and have twice the lifetime of conventional balls.

Sunscreens and cosmetics based on nanotech are already widely used. Customers like products that are translucent because they suggest purity and cleanliness, and L'Oréal discovered that when lotions are ground down to 50 or 60 nms, they let light through. For sunscreens, mineral nanoparticles such as titanium dioxide offer several advantages. Traditional chemical UV protection suffers from its poor long-term stability. Titanium dioxide nanoparticles have a comparable UV protection property as the bulk material, but lose the cosmetically undesirable whitening as the particle size is decreased. For anti-wrinkle creams, a polymer capsule is used to transport active agents like vitamins.

Televisions using carbon nanotubes could be in use by late 2006 according to Samsung[16]. Manufacturers expect these "field effect displays," (FED) to consume less energy than plasma or liquid crystal display (LCD) sets and combine the thinness of LCD and the image quality of traditional cathode ray tubes (CRT). The electrons in an FED are fired through a vacuum at a layer of phosphorescent glass covered with pixels. But unlike CRT, the electron source, the carbon, is only 1 to 2 mm from the target glass instead of 60cm with CRT, and, instead of one electron source, the electron gun, there are thousands. FED contain less electronics than LCD and can be produced in a wide range of sizes. Toshiba, for example, will offer screen sizes of at least 50 inches, around 130 cm.

What applications are foreseen in the medium term[17]?

The following list gives a quick overview of the many domains where nanotechnology is expected to fundamentally change products and how they are produced over the next two decades. Some of these uses will be examined in more detail in the section entitled ”Sectoral examples”.

Electronics and communications: recording using nanolayers and dots, flat-panel displays, wireless technology, new devices and processes across the entire range of communication and information technologies, factors of thousands to millions improvements in both data storage capacity and processing speeds and at lower cost and improved power efficiency compared to present electronic circuits

Chemicals and materials: catalysts that increase the energy efficiency of chemical plants and improve the combustion efficiency (thus lowering pollution emission) of motor vehicles, super-hard and tough (i.e., not brittle) drill bits and cutting tools, "smart" magnetic fluids for vacuum seals and lubricants

Pharmaceuticals, healthcare, and life sciences: nanostructured drugs, gene and drug delivery systems targeted to specific sites in the body, bio-compatible replacements for body parts and fluids, self-diagnostics for use in the home, sensors for labs-on-a-chip, material for bone and tissue regeneration

Manufacturing: precision engineering based on new generations of microscopes and measuring techniques, new processes and tools to manipulate matter at an atomic level, nanopowders that are sintered into bulk materials with special properties that may include sensors to detect incipient failures and actuators to repair problems, chemical-mechanical polishing with nanoparticles, self-assembling of structures from molecules, bio-inspired materials and biostructures

Energy technologies: new types of batteries, artificial photosynthesis for clean energy, quantum well

---

[16] Michael Kanellos, ”Carbon TVs to edge out liquid crystal, plasma?” ZDNET News, January 56th, 2005 http://news.zdnet.com/2100-9596_22-5512225.html

[17] « Nanosciences et nanotechnologies » Ministère délégué, recherches et nouvelles technologies, Paris, 2003. English site here : English site here: http://www.nanomicro.recherche.gouv.fr/uk_index.html

solar cells, safe storage of hydrogen for use as a clean fuel, energy savings from using lighter materials and smaller circuits

Space exploration: lightweight space vehicles, economic energy generation and management, ultra-small and capable robotic systems

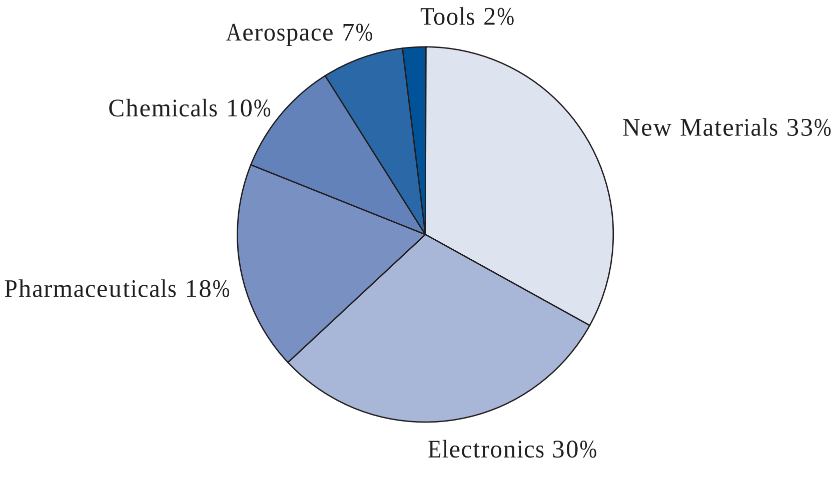
Environment: selective membranes that can filter contaminants or even salt from water, nanostructured traps for removing pollutants from industrial effluents, characterisation of the effects of nanostructures in the environment, maintenance of industrial sustainability by significant reductions in materials and energy use, reduced sources of pollution, increased opportunities for recycling

National security: detectors and detoxifiers of chemical and biological agents, dramatically more capable electronic circuits, hard nanostructured coatings and materials, camouflage materials, light and self-repairing textiles, blood replacement, miniaturised surveillance systems.

3. Market prospects and opportunities

As an OECD survey emphasises[18], most statistical offices do not collect data on nanotechnology R&D, human resources or industrial development, in part because nanotechnology remains a relatively new field of science and technology (and even more so of government policy) and also because of its interdisciplinary and cross-sectoral character. Given this, estimates of potential nanotech markets tend to come from private sources such as specialised consultancy firms who survey a wide number of actors in the field. Lux Research, for example, states that: ”Sales of products incorporating emerging nanotechnology will rise from less than 0.1% of global manufacturing output today to 15% in 2014, totalling $2.6 trillion. This value will approach the size of the information technology and telecom industries combined and will be 10 times larger than biotechnology revenues”[19]. Insurers SwissRe echo this: ”Sales revenues from products manufactured using nanotechnology have already reached eleven-digit figures and are projected to generate twelve-digit sums by 2010, even thirteen-digit sums by 2015”[20]. The chart below shows a projection for the US economy from the National Science Foundation.

Projected contribution of nanotechnology to the US economy, 2015



Source: US National Science Foundation, 2003[21]

---

[18] OECD, Working Party on Innovation and Technology Policy ”Results of OECD Mini-Survey on Nanotechnology R&D Programmes” 7-8 June 2004

[19] Lux Research, October 25, 2004: ”Revenue from nanotechnology-enabled products to equal IT and telecom by 2014, exceed biotech by 10 times.” http://www.luxresearchinc.com/press/RELEASE_SizingReport.pdf

[20] Swiss Re, 2004 ”Nanotechnology : Small matter, many unknowns” http://www.swissre.com/INTERNET/pwswpspr.nsf/fmBookMarkFrameSet?ReadForm&BM=../vwAllbyIDKeyLu/yhan-5yucvt?OpenDocument

[21] OECD Information Technology Outlook 2004, ch. 7, page 264, from M.C. Roco "The Future of National Nanotechnology Initiative" NSF, November 7, 2003

These estimates should be treated with caution, because, as their authors point out, they refer to products ”incorporating nanotechnology” or ”manufactured using nanotechnology”, not to nanotechnology products as such. (It’s as if the value textile industry were calculated by including everything ”incorporating” textiles, be it clothes, aircraft or automobiles.) Lux actually evaluates sales of basic nanomaterials like carbon nanotubes at $13 billion in 2014, a considerable sum, but far from $2.6 trillion. Although the accountancy may be contested, the projected three-phase growth path seems credible:

1. In the present phase, nanotechnology is incorporated selectively into high-end products, especially in automotive and aerospace applications.
2. Through 2009, commercial breakthroughs unlock markets for nanotechnology innovations. Electronics and IT applications dominate as microprocessors and memory chips built using new nanoscale processes come to market.
3. From 2010 onwards, nanotechnology becomes commonplace in manufactured goods. Health- care and life sciences applications finally become significant as nano-enabled pharmaceuticals and medical devices emerge from lengthy human trials.

Following the dotcom fiasco, potential investors are justifiably wary of treating nanotech (or anything else) as ”the next big thing”. While the excitement and hype generated by nanotech’s apostles may be reminiscent of how internet was going to change the world and make everybody rich, there are two crucial differences that counteract the likelihood of a nanobubble: nanotechnology is extremely difficult and extremely expensive, which is why it is concentrated in well-funded companies and institutes that can attract and nurture the scarce scientific and technical expertise needed to understand the problems and challenges. Moreover, the long lead times involved in moving from concept to commercialisation make nanotech particularly unsuitable for making money fast.

3.1. Sectoral example: Medicine[22]

Medical and life-cience applications may prove to be the most lucrative markets for nanotechnologies, with ”lab-on-a-chip” devices already being manufactured and animal testing and early clinical trials starting on nanotechniques for drug delivery. However, the long product approval processes typical of the domain may mean that the health benefits to users and economic benefits to companies will take longer to realise than in other domains. Nanotech’s promise comes from the fact that nanoscale devices are a hundred to ten thousand times smaller than human cells and are similar in size to large biological molecules ("biomolecules") such as enzymes and receptors. For example, haemoglobin, the molecule that carries oxygen in red blood cells, is approximately 5 nm in diameter, DNA 2.5, while a quantum dot is about the same size as a small protein (<10 nm) and some viruses measure less than 100 nm. Devices smaller than 50 nm can easily enter most cells, while those smaller than 20 nm can move out of blood vessels as they circulate through the body.

Because of their small size, nanoscale devices can readily interact with biomolecules on both the surface of cells and inside of cells. By gaining access to so many areas of the body, they have the potential to detect disease and deliver treatment in new ways. Nanotechnology offers the opportunity to study and interact with cells at the molecular and cellular scales in real time, and during the earliest stages of the development of a disease. And since nanocomponents can be made to share some of the same properties as natural nanoscale structures, it is hoped to develop artificial nanostructures that sense and repair damage to the organism, just as naturally-occurring biological nanostructures such as white blood cells do.

Cancer research illustrates many of the medical potentials of nanotechnologies in the longer term. It is hoped that nanoscale devices and processes will help to develop[23]:

- Imaging agents and diagnostics that will allow clinicians to detect cancer in its earliest stages,
- Systems that will provide real-time assessments of therapeutic and surgical efficacy for accelerating

---

[22] Medicine Canadian Institute of Neurosciences, Mental Health and Addiction ”Nanomedicine Taxonomy”, February 2003
http://www.regenerativemedicine.ca/nanomed/Nanomedicine%20Taxonomy%20(Feb%202003).PDF

[23] US Dept. of Health and Human Services: ”Going Small for Big Advances Using Nanotechnology to Advance Cancer Diagnosis, Prevention and Treatment” January 2004
http://nano.cancer.gov/resource_brochure_cancer_nanotechnology.pdf

clinical translation,

- Multifunctional, targeted devices capable of bypassing biological barriers to deliver multiple therapeutic agents directly to cancer cells and those tissues in the microenvironment that play a critical role in the growth and metastasis of cancer,
- Agents that can monitor predictive molecular changes and prevent precancerous cells from becoming malignant,
- Novel methods to manage the symptoms of cancer that adversely impact quality of life,
- Research tools that will enable rapid identification of new targets for clinical development and predict drug resistance.

Drug delivery

This may be the most profitable application of nanotechnology in medicine, and even generally, over the next two decades. Drugs need to be protected during their transit through the body to the target, to maintain their biological and chemicals properties or to stop them damaging the parts of the body they travel through. Once a drug arrives at its destination, it needs to be released at an appropriate rate for it to be effective. This process is called encapsulation, and nanotechnology can improve both the diffusion and degradation characteristics of the encapsulation material, allowing the drug to travel efficiently to the target and be released in an optimal way. Nanoparticle encapsulation is also being investigated for the treatment of neurological disorders to deliver therapeutic molecules directly to the central nervous system beyond the blood-brain barrier, and to the eye beyond the blood-retina barrier. Applications could include Parkinson's, Huntington's chorea, Alzheimer's, ALS and diseases of the eye.

Repair and replacement

Damaged tissues and organs are often replaced by artificial substitutes, and nanotechnology offers a range of new biocompatible coatings for the implants that improves their adhesion, durability and lifespan. New types of nanomaterials are being evaluated as implant coatings to improve interface properties. For example, nanopolymers can be used to coat devices in contact with blood (e.g. artificial hearts, catheters) to disperse clots or prevent their formation. Nanomaterials and nanotechnology fabrication techniques are being investigated as tissue regeneration scaffolds. The ultimate goal is to grow large complex organs. Examples include nanoscale polymers moulded into heart valves, and polymer nanocomposites for bone scaffolds. Commercially viable solutions are thought to be 5 to 10 years away, given the scientific challenges related to a better understanding of molecular / cell biology and fabrication methods for producing large three-dimensional scaffolds.

Nanostructures are promising for temporary implants, e.g. that biodegrade and do not have to be removed in a subsequent operation. Research is also being done on a flexible nanofiber membrane mesh that can be applied to heart tissue in open-heart surgery. The mesh can be infused with antibiotics, painkillers and medicines in small quantities and directly applied to internal tissues.

Subcutaneous chips are already being developed to continuously monitor key body parameters including pulse, temperature and blood glucose. Another application uses optical microsensors implanted into subdermal or deep tissue to monitor tissue circulation after surgery, while a third type of sensor uses MEMS (microelectromechanical system) devices and accelerometers to measure strain, acceleration, angular rate and related parameters for monitoring and treating paralysed limbs, and to improve the design of artificial limbs. Implantable sensors can also work with devices that administer treatment automatically if required, e.g. fluid injection systems to dispense drugs. Initial applications may include chemotherapy that directly targets tumors in the colon and are programmed to dispense precise amounts of medication at convenient times, such as after a patient has fallen asleep. Sensors that monitor the heart's activity level can also work with an implantable defibrillator to regulate heartbeats.

Hearing and vision

Nano and related micro technologies are being used to develop a new generation of smaller and potentially more powerful devices to restore lost vision and hearing. One approach uses a miniature video camera attached to a blind person's glasses to capture visual signals processed by a microcomputer worn on the belt and transmitted to an array of electrodes placed in the eye. Another approach uses of a subretinal implant designed to replace photoreceptors in the retina. The implant uses a microelectrode array powered by up to 3500 microscopic solar cells.

For hearing, an implanted transducer is pressure-fitted onto a bone in the inner ear, causing the bones to vibrate and move the fluid in the inner ear, which stimulates the auditory nerve. An array at the tip of the device uses up to 128 electrodes, five times higher than current devices, to simulate a fuller range of sounds. The implant is connected to a small microprocessor and a microphone in a wearable device that clips behind the ear. This captures and translates sounds into electric pulses transmitted by wire through a tiny hole made in the middle ear.

3.2. Sectoral example: Food and agriculture

Nanotechnology is rapidly converging with biotech and information technology to radically change food and agricultural systems. Over the next two decades, the impacts of nano-scale convergence on farmers and food could even exceed that of farm mechanisation or of the Green Revolution according to some sources such as the ETC group[24]. Food and nutrition products containing nano-scale additives are already commercially available. Likewise, a number of pesticides formulated at the nano-scale are on the market and have been released in the environment. According to Helmut Kaiser Consultancy, some 200 transnational food companies are currently investing in nanotech and are on their way to commercialising products[25]. The US leads, followed by Japan and China. HKC expects the nanofood market to surge from $2.6 billion in 2003 to $7.0 billion in 2006 and to $20.4 billion in 2010.

Companies not associated with food production in the public mind are already supplying nano-enabled ingredients to the industry. BASF, for example, exploits the fact that many vitamins and other substances such as carotinoids are insoluble in water, but can easily be mixed with cold water when formulated as nanoparticles. Many lemonades and fruit juices contain these specially formulated additives, which can also be used to provide an ”attractive” color[26].

Expected breakthroughs in crop DNA decoding and analysis could enable agrifirms to predict, control and improve agricultural production. And with technology for manipulating the molecules and atoms of food, the food industry would have a powerful method to design food with much greater capability and precision, lower costs and improved sustainability. The combination of DNA and nanotechnology research could also generate new nutrition delivery systems, to bring active agents more precisely and efficiently to the desired parts of the human body.

Nanotechnology will not only change how every step of the food chain operates but also who is involved. At stake is the world’s $3 trillion food retail market, agricultural export markets valued at $544 billion, the livelihoods of farmers and the well-being of the rest of us. Converging technologies could reinvigorate the battered agrochemical and agbiotech industries, possibly igniting a still more intense debate – this time over "atomically-modified" foods.

The most cited nano-agricultural developments are:

Nanoseeds: In Thailand, scientists at Chiang Mai University's nuclear physics laboratory have rearranged the DNA of rice by drilling a nano-sized hole through the rice cell's wall and membrane and inserting a nitrogen atom. So far, they've been able to change the colour of the grain, from purple to green.

Nanoparticle pesticides: Monsanto, Syngenta and BASF are developing pesticides enclosed in nanocapsules or made up of nanoparticles. The pesticides can be more easily taken up by plants if they're in nanoparticle form; they can also be programmed to be ”time-released.”

Nanofeed for Chickens: With funding from the US Department of Agriculture (USDA), Clemson University researchers are feeding bioactive polystyrene nanoparticles that bind with bacteria to chickens as an alternative to chemical antibiotics in industrial chicken production.

Nano Ponds: One of the USA’s biggest farmed fish

---

[24] ETC Group ”Down on the Farm: The Impact of Nano-scale Technologies on Food and Agriculture” November 2004 2004 http://www.etcgroup.org/documents/ETC_DOTFarm2004.pdf

[25] HKC ”Nanotechnology in Food and Food Processing Industry Worldwide 2003-2006-2010-2015” 2003 . http://www.hkc22.com/nanofood.html. [The subsequent projections for the world nanofood market may well prove to be underestimates, given the future purchasing power of senior citizens in developed economies and a world-wide functional food market of already $70 billion.]

[26] BASF ”Improved products, more efficient processes, and new properties” http://www.corporate.basf.com/en/innovationen/felder/nanotechnologie/nanotech.htm?printview=on&docid=22321&id=V00-6iy3A6dubbcp-S3

companies, Clear Spring Trout, is adding nanoparticle vaccines to trout ponds, where they are taken up by fish.

”Little Brother”: The USDA is pursuing a project to cover farmers’ fields and herds with small wireless sensors to replace farm labour and expertise with a ubiquitous surveillance system.

Nano foods: Kraft, Nestlé, Unilever and others are employing nanotech to change the structure of food – creating ”interactive” drinks containing nanocapsules that can change colour and flavour (Kraft) and spreads and ice creams with nanoparticle emulsions (Unilever, Nestlé) to improve texture. Others are inventing small nanocapsules that will smuggle nutrients and flavours into the body (what one company calls ”nanoceuticals”).

Nano packaging: BASF, Kraft and others are developing new nanomaterials that extend food shelf life and signal when a food spoils by changing colour.

Food safety: Scientists from the University of Wisconsin have successfully used single bacterial cells to make tiny bio-electronic circuits, which could in the future be used to detect bacteria, toxins and proteins[27].

Nanosensors can work through a variety of methods such as by the use of nanoparticles tailor-made to fluoresce different colors or made from magnetic materials can selectively attach themselves to food pathogens. Handheld sensors employing either infrared light or magnetic materials could then note the presence of even minuscule traces of harmful pathogens. The advantage of such a system is that literally hundreds and potentially thousands of nanoparticles can be placed on a single nanosensor to rapidly, accurately and affordably detect the presence of any number of different bacteria and pathogens. A second advantage of nanosensors is that given their small size they can gain access into the tiny crevices where the pathogens often hide, and nanotechnology may reduce the time it takes to detect the presence of microbial pathogens from two to seven days down to a few hours and, ultimately, minutes or even seconds[28].

3.3. Sectoral example: Semiconductors and computing

The computer industry is already working on a nanoscale. Although the current production range is at 90 nm, 5 nm gates have been proven in labs, although they cannot be manufactured yet. By 2010, world-wide, about $300 billion worth of semiconductor production will be nanotechnology-based (including nanocomponents such as nanolayers, nanoscale treated materials, or other nanostructures) and by 2015, about $500 billion. Because nanotechnology can reduce its basic features, CMOS will continue being used for a decade or more. The intermediate future will have CMOS married to a generation of nanodevices as yet undefined, because there are many alternatives, and it is still too early to tell which will prevail. One solution could be hybrid structures exploiting the advantages of today’s CMOS technology (integration and scaling of transistors and high functionality on a small support) with off-chip optoelectronic interconnects to overcome the throughput bottlenecks[29].

Towards 2015, semiconductor development priorities will change, as the focus shifts from scaling and speed to system architecture and integration, with user-specific applications for bio-nanodevices, the food industry and construction applications. Another trend is the convergence between IT, nanotechnology, biotechnology and cognitive sciences. The higher speeds at which information will be disseminated will change how we work with computers, and also perhaps how we deal with things like damaged nerves, possibly by developing direct interfaces with the nervous system and electronic circuits, so-called neuromorphic engineering, where signals are directly transmitted from a human organism to a machine.

The actual technologies employed are hard to predict. Currently there exist at least four interrelated technical barriers to nanoscale manufacturing:
How to control the assembly of 3-D heterogeneous systems, including alignment, registration and interconnection at 3-D and with multiple functionalities.

---

[27] Robert J. Hamers et al. ”Manipulation and Real-Time Electrical Detection of Individual Bacterial Cells at Electrode Junctions: A Model for Assembly of Nanoscale Biosystems” Nano Letters April 2005. First presented at American Chemical Society, March 2005 http://pubs.acs.org/journals/nalefd/

[28] Manuel Marquez ”Nanotechnology to play important and prominent role in food safety” Advantage Magazine, February 2004 http://www.azonano.com/details.asp?ArticleID=858

[29] Alexander E. Braun ”Nanotechnology : genesis of semiconductors future” Semiconductor International, November 2004 http://www.reed-electronics.com/semiconductor/article/CA476295?pubdate=11%2F1%2F2004

How to handle and process nanoscale structures in a high-rate / high-volume manner without compromising beneficial nanoscale properties.
How to test nanocomponents' long-term reliability, and detect, remove or prevent defects and contamination.
Metrology. At present, using an electron microscope, it is possible to get depth of field, sufficient resolution or low energy (important so as not to damage certain components), but not all three at once. Failure analysis is another metrology issue: how to get a real 3-D view of the structure and defects that may develop during processing or use.

At present, technology front runners include spin electronics, molecular electronics (see below), biocomponents, quantum computing, DNA computing, etc. However, the history of technology teaches that sudden upsets that could change everything are to be expected. As recently as 1998, limited use was predicted for giant magnetoresistance introduced by IBM. But within two years it replaced all equivalent hard disk reading technologies and their extensive production facilities. The technique exploits the electron's spin to produce novel interconnect and device structures, giving rise to the name "spintronics"[30]. Spin is present in all electrons, and manipulating spin would use conventional solid-state semiconductor and metal materials, without the problems associated with nanotubes or molecules. Spin packets have a long lifetime and high mobility in semiconductors, making them attractive for transmitting information in the chip, within the silicon, without using a metal. One major problem with spintronics is that when a magnet heats up, it ceases being ferromagnetic, a condition necessary to exploit the electron spin. It is also difficult to control the ferromagnetic force or direction.

Assuming these problems can be solved, promising applications for spintronics include MRAM (magnetic random access memory), a high-speed non-volatile memory architecture; and logic devices like the spin field effect transistor (spin FET), which consumes less power and operates faster than its conventional counterpart.

Chip makers are already working at around 100 nm, but this is essentially a ”shrinking” of conventional technologies to make them smaller, and this is now reaching its limits. Miniaturisation to much smaller scales will run into problems caused by quantum phenomena, such as electrons tunnelling through the barriers between wires, so an alternative to transistor technology must be found, one whose components will exploit quantum effects rather suffer from them. The first generation of nanocomputers will have components that behave according to quantum mechanics, but their algorithms will probably not involve quantum mechanics. If quantum mechanics could be used in the algorithms as well, the computer would be enormously more powerful than any classical scheme, but such developments are unlikely in the foreseeable future[31].

In the meantime, research is in progress to manipulate molecules to carry out calculations. In "chemical computing", a series of chemical reactions, e.g. of DNA, corresponds to a computation, with the final products of the reactions representing the answer. With this technique, many calculations can be carried out in parallel, but each step requires a long time, and can be very expensive because of the cost of the chemicals used.

A second approach is to use molecules as the ”host” for nuclear spins that form the quantum bits (qubits) in a nuclear magnetic resonance-based computer. However, this approach may not be able to scale up to a computationally useful number of qubits.

The most promising approach is thought to be molecular electronics, using a molecule or group of molecules in a circuit. Bit densities for molecular logic and memory components could be on the order of a terabit / cm^2 (6.5 terabits / in2). Switching speeds could get down into the range of a few picoseconds (1000 times faster than current DRAM)[32].

---

[30] Albert Fert et al ”The new era of spintronics” Europhysics News (2003) Vol. 34 No. 6 http://www.europhysicsnews.com/full/24/article9/article9.html

[31] Simon Benjamin and Artur Ekart ”Towards Quantum Information Technology” CambridgeCentre for Quantum Computation, 2002 2002 http://cam.qubit.org/articles/intros/nano/nano.php

[32] Cientifica ”Molecular electronics” October 2003 http://nanotechweb.org/dl/wp/molecular_electronics_WP.pdf

3.4. Sectoral example: Textiles

The textile industry could be affected quite significantly by nanotechnology, with some estimates talking of a market impact of hundreds of billions of dollars over the next decade. Nanoscience has already produced stain- and wrinkle-resistant clothing, and future developments will focus on upgrading existing functions and performances of textile materials; and developing ”smart” textiles with unprecedented functions such as:

- sensors and information acquisition and transfer,
- multiple and sophisticated protection and detection,
- health-care and wound-healing functions,
- self-cleaning and repair functions.

This last function illustrates how nanotechnology could impact areas outside its immediate application.

US company Nano-Tex is already marketing its NanoCare stain- and wrinkle-resistant technology, and NanoFresh (to freshen sports clothing) is expected soon. Scientists at the Hong Kong Polytechnic University have built a nano layer of particles of titanium dioxide, a substance that reacts with sunlight to break down dirt and other organic material. This layer can be coated on cotton to keep the fabric clean. Clothes simply need to be exposed to natural or ultraviolet light for the cleaning process to begin. Once triggered by sunlight, clothing made out of the fabric will be able to rid itself of dirt, pollutants and micro-organisms. The whole laundry industry would be affected if the technology proves to be economically viable.

Research involving nanotechnology to improve performances or to create new functions is most advanced in nanostructured composite fibers employing nanosize fillers such as nanoparticles (clay, metal oxides, carbon black), graphite nanofibers (GNF) and carbon nanotubes (CNT). The main function of nanosize fillers is to increase mechanical strength and improve physical properties such as conductivity and antistatic behaviours. Being evenly distributed in polymer matrices, nanoparticles can carry load and increase the toughness and abrasion resistance; nanofibers can transfer stress away from polymer matrices and enhance tensile strength of composite fibers. Additional physical and chemical performances imparted to composite fibers vary with specific properties of the nanofillers used. Although some of the filler particles such as clay, metal oxides, and carbon black have previously been used as microfillers in composite materials for decades, reducing their size into nanometer range have resulted in higher performances and generated new market interest.

Carbon Nanofibers and Carbon Nanoparticles

Carbon nanofibers and carbon black nanoparticles are among the most commonly used nanosize filling materials. Carbon nanofibers can effectively increase the tensile strength of composite fibers due to their high aspect ratio, while carbon black nanoparticles can improve abrasion resistance and toughness. Both have high chemical resistance and electric conductivity.

Clay Nanoparticles

Clay nanoparticles or nanoflakes possess electrical, heat and chemical resistance and an ability to block UV light. Composite fibers reinforced with clay nanoparticles exhibit flame retardant, anti-UV and anti-corrosive behaviours.

Metal Oxide Nanoparticles

Certain metal oxide nanoparticles possess photocatalytic ability, electrical conductivity, UV absorption and photo-oxidising capacity against chemical and biological species. Research involving these nanoparticles focuses on antimicrobial, self-decontaminating and UV blocking functions for both military protection gear and civilian health products.

Carbon Nanotubes

Potential applications of CNTs include conductive and high-strength composite fibers, energy storage and energy conversion devices, sensors, and field emission displays. One CNT fiber already exhibits twice the stiffness and strength, and 20 times the toughness of steel wire of the same weight and length. Moreover, toughness can be four times higher than that of spider silk and 17 times greater than Kevlar fibers used in bullet-proof vests, suggesting applications in safety harnesses, explosion-proof blankets, and electromagnetic shielding.

Nanotechnology in Textile Finishing

Nanoscale emulsification, through which finishes can be applied to textile material in a more thorough, even and precise manner provide an unprecedented level of textile performance regarding stain-resistant, hydrophilic, anti-static, wrinkle resistant and shrink-proof properties.

Nanosize metal oxide and ceramic particles have a larger surface area and hence higher efficiency than larger size particles, are transparent, and do not blur the color and brightness of the textile substrates. Fabric treated with nanoparticles $Ti0_2$ and MgO replaces fabrics with active carbon, previously used as chemical and biological protective materials. The photocatalytic activity of $Ti0_2$ and MgO nanoparticles can break down harmful chemicals and biological agents.

Finishing with nanoparticles can convert fabrics into sensor-based materials. If nanocrystalline piezoceramic particles are incorporated into fabrics, the finished fabric can convert exerted mechanical forces into electrical signals enabling the monitoring of bodily functions such as heart rhythm and pulse if they are worn next to skin.

Self assembled Nanolayers

In the longer-term future, self-assembled nanolayer (SAN) coating may challenge traditional textile coating. Research in this area is still in the very early stages, but the idea is to deposit a coating less than one nanometer thick on the textile, and then to vary the number of successive nanolayers to modulate the desired physical properties of the finished article.

3.5. Sectoral example: Energy

Breakthroughs in nanotechnology could provide technologies that would contribute to world-wide energy security and supply. A report published by Rice University (Texas) in February 2005 identified numerous areas in which nanotechnology could contribute to more efficient, inexpensive, and environmentally sound technologies than are readily available[33]. Although the most significant contributions may be to unglamorous applications such as better materials for exploration equipment used in the oil and gas industry or improved catalysis, nanotechnology is being proposed in numerous energy domains, including solar power; wind; clean coal; fusion reactors; new generation fission reactors; fuel cells; batteries; hydrogen production, storage and transportation; and a new electrical grid that ties all the power sources together. The main challenges where nanotechnology could contribute are:

- Lower the costs of photovoltaic solar energy tenfold,
- Achieve commercial photocatalytic reduction of CO_2 to methanol,
- Create a commercial process for direct photo-conversion of light and water to produce hydrogen,
- Lower the costs of fuel cells between tenfold and a hundredfold and create new, sturdier materials,
- Improve the efficiency and storage capacity of batteries and supercapacitors between tenfold and a hundredfold for automotive and distributed generation applications,
- Create new lightweight materials for hydrogen storage for pressure tanks, liquid hydrogen vessels, and an easily reversible hydrogen chemisorption system,
- Develop power cables, superconductors or quantum conductors made of new nanomaterials to rewire the electricity grid and enable long-distance, continental and even international electrical energy transport, also reducing or eliminating thermal sag failures, eddy current losses and resistive losses by replacing copper and aluminium wires,
- Develop thermochemical processes with catalysts to generate hydrogen from water at temperatures lower than 900C at commercial costs,
- Create superstrong, lightweight materials that can be used to improve energy efficiency in cars, planes and in space travel; the latter, if combined with nanoelectronics based robotics, possibly enabling space solar structures on the moon or in space,
- Create efficient lighting to replace incandescent and fluorescent lights,
- Develop nanomaterials and coatings that will enable deep drilling at lower costs to tap energy resources, including geothermal heat, in deep strata,
- Create CO_2 mineralization methods that can work on a vast scale without waste streams.

Solving these challenges will take many years, but commercial and public research institutes are already exploiting nanotechnology for energy applications. Bell Labs, for example, is exploring the possibility of producing a microbattery that would still work 20 years after purchase by postponing the chemical reactions that degrades traditional batteries. The battery is based on a Bell Labs discovery that liquid droplets of electrolyte will stay in a dormant state atop microscopic structures called "nanograss" until stimulated to flow, thereby triggering a reaction producing electricity[34]. Other researchers hope to

---

[33] "Energy and Nanotechnology : Strategy for the Future"
http://www.rice.edu/energy/publications/docs/NanoReport.pdf

[34] "mPhase Technologies and Bell Labs Successfully Demonstrate First Battery Based on 'Nanograss'"
Lucent Technologies, September 28 2004
http://www.lucent.com/press/0904/040928.bla.html

dispense with batteries completely by developing nanotubes-based ”ultra” capacitors powerful enough to propel hybrid-electric cars. Compared with batteries, ultracapacitors can put out much more power for a given weight, can be charged in seconds rather than hours, and can function at more extreme temperatures. They're also more efficient, and they last much longer. The technology is in it earlier stages: the world-wide market was only $38 million in 2002, the most recent year for which figures are available, but researcher firm Frost & Sullivan predicts total 2007 revenues for ultracapacitors of $355 million[35].

Photovoltaics is another area where nanotech is already providing products that could have a significant impact. Three US-based solar cell start-ups (Nanosolar, Nanosys and Konarka Technologies), and corporate players including Matsushita and STMicroelectronics are striving to produce photon-harvesting materials at lower costs and in higher volumes than traditional crystalline silicon photovoltaic cells[36]. Nanosolar has developed a material of metal oxide nanowires that can be sprayed as a liquid onto a plastic substrate where it self-assembles into a photovoltaic film. A roll-to-roll process similar to high-speed printing offers a high-volume approach that does not require high temperatures or vacuum equipment. Nanosys intends its solar coatings to be sprayed onto roofing tiles. And Konarka is developing plastic sheets embedded with titanium dioxide nanocrystals coated with light-absorbing dyes. The company acquired Siemens' organic photovoltaic research activities, and Konarka's recent $18 million third round of funding included the world's first- and fifth-largest energy companies, Electricité de France and ChevronTexaco. If nanotech solar fabrics could be applied to, e.g., buildings and bridges, the energy landscape could change in important ways. Integrated into the roof of a bus or truck, they could split water via electrolysis and generate hydrogen to run a fuel cell. Losers would include current photovoltaic-cell makers and battery manufacturers who failed to react to the new challenge.

Such developments however depend on solving a number of fundamental problems at the nanoscale, but researchers are making fast progress using nanoscale design, include accelerating the kinetics of reactions through catalysis, separating the products at high temperature, and directing products to the next reaction step.

3.6. Nanotechnology and the situation of developing countries

While research and development in nanotechnology is quite limited in most developing countries, there will be increasing opportunities to import nano products and processes. It can be argued of course that nanotechnology could make the situation of developing countries worse by reducing demand for their exports, notably raw materials. Moreover, even in developing countries, few nanotech projects specifically target the needs of the poor, leading to fears of a ”nano divide” similar to the digital divide. The UN’s International Centre for Science and High Technology tackled such issues in February 2005 at a meeting entitled ”North-South dialogue on nanotechnology”[37]. The Centre argued that nanotechnology may offer important benefits to developing countries and it is not correct to assume that it is too difficult or too expensive for them. A similar theme was the subject of a report published in April 2005 by the Canadian Program on Genomics and Global Health (CPGGH) at the University of Toronto Joint Centre for Bioethics (JCB)[38]. The CPGGH asked over 60 international experts to assess the potential impacts of nanotechnologies for developing countries within the framework of the UN Millennium Development Goals. Agreed in 2000 for achievement by 2015, the UN goals are: to halve extreme poverty and hunger; achieve universal primary education; empower women and promote equality between women and men; reduce under-five mortality by two-thirds; reduce maternal mortality by three-quarters; reverse the spread of diseases, especially HIV/AIDS and malaria; ensure environmental sustainability; and create a global partnership for development.

The CPGGH study ranks the 10 nanotechnology applications most likely to have an impact in the areas of water, agriculture, nutrition, health, energy and the environment by 2015. The ranking is markedly different from similar exercises in the more advanced industrial economies, where applications in electronics and computing are generally seen as the most significant,

---

[35] Frost&Sullivan ”World Ultracapacitor Markets” October 2003 http://www.frost.com/prod/servlet/research.pag

[36] ”Forbes Category killers: 5 nanotechnologies that could change the world”, September 2004 http://www.forbesinc.com/newsletters/nanotech/public/samples/Nanotech_5technologies.pdf

[37] http://www.ics.trieste.it/Nanotechnology/

[38] CPGGH ”Nanotechnology and the developing world” April 2005 http://www.utoronto.ca/jcb/home/documents/PLoS_nanotech.pdf

with pharmaceuticals and other health sectors featuring strongly. For developing countries, the experts reckon the top 10 nanotechnology applications are:

1. Energy. There was a high degree of unanimity in ranking this area number 1. Nanomaterials are being used to build a new generation of solar cells, hydrogen fuel cells and novel hydrogen storage systems that could deliver clean energy to countries still reliant on traditional, non-renewable contaminating fuels. Advances in the creation of synthetic nanomembranes embedded with proteins are capable of turning light into chemical energy. If successfully developed on an industrial scale, such technologies could help developing countries avoid recurrent shortages and price fluctuations that come with dependence on fossil fuels, as well as the environmental consequences of mining and burning oil and coal.

2. Agriculture. Researchers are developing a range of inexpensive nanotech applications to increase soil fertility and crop production, and help eliminate malnutrition – a contributor to more than half the deaths of children under five in developing countries. Nanotech materials are in development for the slow release and efficient dosage of fertilisers for plants and of nutrients and medicines for livestock. Other agricultural developments include nanosensors to monitor the health of crops and farm animals and magnetic nanoparticles to remove soil contaminants.

3. Water treatment. Nano-membranes and nano-clays are inexpensive, portable and easily cleaned systems that purify, detoxify and desalinate water more efficiently than conventional bacterial and viral filters. Researchers also have developed a method of large-scale production of carbon nano-tube filters for water quality improvement. Other water applications include systems (based on titanium dioxide and on magnetic nanoparticles) that decompose organic pollutants and remove salts and heavy metals from liquids, enabling the use of heavily contaminated and salt water for irrigation and drinking. Several of the contaminating substances retrieved could then be easily recycled.

4. Disease diagnosis and screening. Technologies include the "lab-on-a-chip", which offers all the diagnostic functions of a medical laboratory, and other biosensors based on nanotubes, wires, magnetic particles and semiconductor crystals (quantum dots). These inexpensive, hand-held diagnostic kits detect the presence of several pathogens at once and could be used for wide-range screening in small peripheral clinics. Other nanotechnology applications are in development that would greatly enhance medical imaging.

5. Drug delivery systems. Nano-capsules, dendrimers (tiny bush-like spheres made of branched polymers), and "buckyballs" (soccerball-shaped structures made of 60 carbon atoms) for slow, sustained drug release systems, characteristics valuable for countries without adequate drug storage capabilities and distribution networks. Nanotechnology could also potentially reduce transportation costs and even required dosages by improving shelf-life, thermo-stability and resistance to changes in humidity of existing medications;

6. Food processing and storage. Improved plastic film coatings for food packaging and storage may enable a wider and more efficient distribution of food products to remote areas in less industrialised countries; antimicrobial emulsions made with nano-materials for the decontamination of food equipment, packaging, or food; and nanotech-based sensors to detect and identify contamination;

7. Air pollution remediation. Nanotech-based innovations that destroy air pollutants with light; make catalytic converters more efficient, cheaper and better controlled; detect toxic materials and leaks; reduce fossil fuel emissions; and separate gases.

8. Construction. Nano-molecular structures to make asphalt and concrete more resistant to water; materials to block ultraviolet and infrared radiation; materials for cheaper and durable housing, surfaces, coatings, glues, concrete, and heat and light exclusion; and self-cleaning for windows, mirrors and toilets.

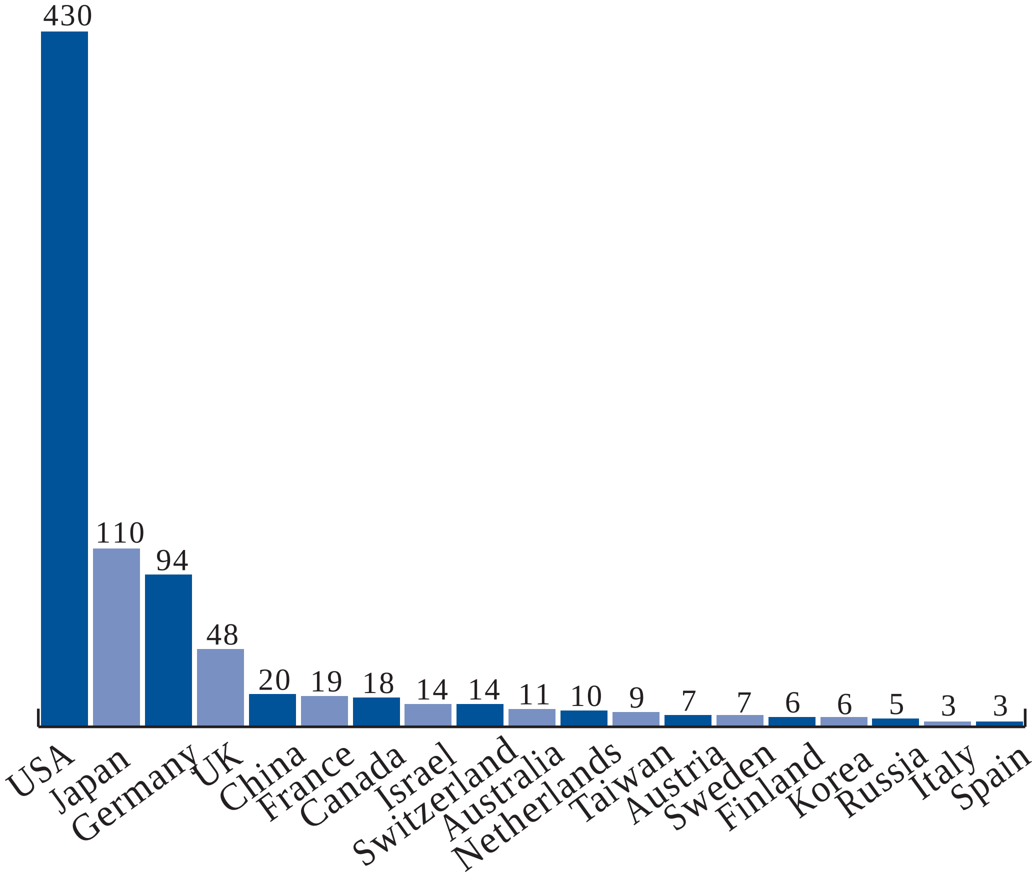
9. Health monitoring. Nano-devices are being developed to keep track of daily changes in physiological variables such as the levels of glucose, of carbon dioxide, and of cholesterol, without the need for drawing blood in a hospital setting. For example, patients suffering from diabetes would know at any given time the concentration of sugar in their blood; similarly, patients with heart diseases would be able to monitor their cholesterol levels constantly.

10. Disease vector and pest detection control. Nanoscale sensors for pest detection, and improved pesticides, insecticides, and insect repellents.

The report also recommends an initiative called Addressing Global Challenges Using Nanotechnology to encourage the development of nanotechnologies targeted at developing nations. It could work along the lines of the Grand Challenges in Global Health initiative launched in 2004 by the Foundation for the National Institutes of Health and the Bill and Melinda Gates Foundation. The authors point out that several developing countries have already launched nanotechnology initiatives. For example, India's Department of Science and Technology will invest $20 million over 2005-2009, and China ranks third in the world behind the United States and Japan in the number of nanotech patent applications. Researchers at China's Tsinghua University have begun clinical tests for a bone scaffold based on nanotechnology which gradually disintegrates as the patient's damaged skeletal tissue heals. This application is especially relevant for developing countries, where the number of skeletal injuries resulting from road traffic accidents is acute. In Brazil, the projected budget for nanoscience over 2004-2007 is about $25 million, and three institutes, four networks, and approximately 300 scientists are working in nanotechnology. Brazilian researchers are investigating the use of modified magnetic nanoparticles to remove oil from oil spills; both the nanoparticles and the oil could potentially be recycled.

4. Players

As the table below shows, most nanotech companies are in the US, mainly because of the more developed venture capital market (over half the venture capital investors in nanotechnology are from the US). Statistics for universities and research institutes also shows a strong, but less marked, US bias.
These figures should be treated with caution, since new companies are created all the time and some do not remain active very long. Moreover, the lack of any clear nanotech strategy even in companies heavily involved in nanotechnology means that tomorrow’s winners will not necessarily be today’s heavy investors. A survey of executives responsible for nanotech at 33 companies with over $5 billion annual revenue found that they have not converged towards any particular model for organising and governing their nanotechnology initiatives. Forty-two percent of represented companies have centralised nanotechnology programs, but an equal share pursue decentralised



© Cientifica and Jaakko Pöyry Consulting 2003[39]

---

[39] Cientifica: "Nanotechnology Opportunity Report, 2nd Edition Executive Summary" June 2003 (unpaginated)

activity with no co-ordination. At 45% of companies, the R&D organisation "owns" nanotech; ownership varies widely across the rest[40].

Barely half of interviewees' firms have a stated nanotech strategy today. When a strategy does exist, it is frequently a platitude like "survey the field and move quickly." Fewer than half of interviewees rate their companies' current approaches to nanotech as "very effective." Pharmaceutical companies are least likely to have an explicit nanotechnology strategy; they also invest the lowest level of people and funding compared with other sectors. (This could however change quite quickly, of course, once critical research thresholds have been crossed.) Asian companies across industries show the highest levels of staffing, funding, and executive sponsorship for nanotech.

These results are all the more surprising given the size of the companies and the sums of money involved. The median corporation had 55 people working on nanotechnology, allocated $33 million in R&D funding in 2004 to nanoscale research, and has partnerships with universities, start-ups, and public sector agencies on multiple nanotech projects. Interviewees expect double-digit increases on each front through 2006.

The roles of the different players in the nanotech sector are no different from those elsewhere[41]:

- large organisations, with the resources to investigate longer-term technologies, seek applications to improve margins, lower costs or increase market share,
- start-ups, seeking to apply technologies in order to capture market share or disrupt existing markets, attract the attention of acquisition-hungry incumbents,
- economic blocs compete for supremacy, mindful of the economic benefits that strength in many of the applications of nanotechnology will bring,
- public agencies attempt to capture the maximum number of links in the value chain.

The long-term outcome is also likely to be similar, with, as happened in microcomputing, one or two new large companies emerging and most of the other viable start-ups being absorbed by large firms. Small companies are likely to target emerging technologies, seeking new generations of products, of which four generations were identified in 2004 by Mihail Roco, Chairman of the Federal Subcommittee on Nanoscale Science Engineering and Technology which oversees the National Nanotechnology Initiative[42]. At the University of Southern California's ”Nano Ethics Conference” in March 2005, Dr Roco revised his estimates for third and fourth generations (making them five years later) and added even a fifth generation[43].

1. 2001. Passive nanostructures. This corresponds to the current state of affairs, with the creation of commercial prototypes and the acquisition of systematic control at the nanoscale for these products using nanostructure polymers, wires, coatings, etc,
2. 2005. "Active" nanostructures, i.e. devices such as actuators that behave like muscles; transistors with active parts created by design; drug delivery within the human body at specific locations and times. Such applications are already in advanced R&D, and some commercial prototypes are expected in the next year or so. Active nanostructure devices will lead to a significant market expansion,
3. 2015-20 (2010 in previous estimate). The third generation will arrive when nanodevices and nanomaterials are integrated into larger nanosystems, and systems of nanosystems with emerging behaviour will be created as commercial prototypes. This includes directed multiscale self-assembling and chemico-mechanical processing. By 2015 nanoscale designed catalysts will expand the use in ”exact” chemical manufacturing to cut and link molecular assemblies, with minimal waste,
4. 2020 (previously 2015). Fourth-generation large nanosystems whose different components will be molecules or macromolecules will emerge over 2005-2020. These will approximate how living systems work (except that living systems are more complex, being integrated on lengthier scales, generally use water, and grow slowly). In detail, this generation includes heterogeneous molecular nanosystems, where each molecule in the nanosystem has a specific structure and plays a different role. Molecules will be used as devices and from their engineered structures and architectures will emerge fundamentally new functions. Research focus will

---

[40] Peter Herbert ”Corporate nanotechnology investments are at risk of being wasted” Lux Research, January 2004 http://www.luxresearchinc.com/press/RELEASE_CEOPlaybook.pdf

[41] Cientifica ”Nanotechnology opportunity report” June 2003 (no longer available online)

[42] Mihail C; Roco ”Nanoscale Science and Engineering: Unifying and Transforming Tools” AIChE Journal Vol. 50, Issue 5, published online in April 2004

[43] USC Nano Ethics Conference, 3-5 March 2005 http://nsts.nano.sc.edu/conferences.html

be on atomic manipulation for design of molecules and supramolecular systems, dynamics of single molecule, molecular machines, design of large heterogeneous molecular systems, controlled interaction between light and matter with relevance to energy conversion among others, exploiting quantum control, emerging behaviour of complex macromolecular assemblies, nanosystem biology for healthcare and agricultural systems, human-machine interface at the tissue and nervous system level, and convergence of nano-bio-info-cognitive domains. Examples are creating multifunctional molecules, catalysts for synthesis and controlling of engineered nanostructures, subcellular interventions, and biomimetics for complex system dynamics and control,

5. Starting around 2030, with nanorobotics, guided assembly, and diverging architectures.

5. Nanotechnology programs of governments

An OECD survey published in 2004 indicates that many countries have developed definitions of nanotechnology as a field of science and technology, and explicitly recognise nanotechnology as a multi- or inter-disciplinary field that draws upon work in the physical science, life sciences and engineering[44]. The survey underlines the lack of progress in developing a statistical definition of nanotechnology that can be used for collecting data on R&D expenditures, human resources, firm creation, etc. This should be borne in mind in the evaluating the usefulness of figures quoted elsewhere in this report.

The organisation of government support for nanotechnology reflects its multi-disciplinary nature, with funding for nanotechnology R&D generally channelled through a number of ministries and research councils with responsibility for different fields of application (e.g., environment, industry, health) or fields of science and technology (e.g., physics, life sciences, engineering). Some countries have begun to take steps to centralise management of their R&D programs, even if program implementation remains distributed. In most cases, co-ordination also entails development of a national strategy for nanotechnology development.

The R&D programs vary considerably in size and scope. In many countries, programs are aimed at developing world-class R&D capability in nanotechnology, recognising its importance in a number of industrial fields and in addressing a number of social needs. Japan, for example, has a range of programs aimed at basic research, nano-materials and nano-electronic devices. Canada has established the National Institute of Nanotechnology as an integrated research institute with 150 researchers from various disciplines and a business incubation facility. Denmark has established three new research centres, two of which focus on interdisciplinary approaches to nanoscience, and one of which will address nano systems engineering.

In addition to these broad-based efforts, some countries (including those mentioned above) have established more focused programs as well, linking R&D to specific national needs. For example, the Netherlands is financing three programs dealing respectively with nanotechnology, microsystems, materials science and microelectronics; electronic devices, circuits and systems; and molecular nanotechnologies. Norway is concentrating on sensors and smart materials, microtechnology, energy (e.g., improving value added from oil and gas resources), environment (e.g., improving process efficiency, recycling and reuse), and new process technology. It also has a program aimed at stimulating innovation and growth in the biomedical industry, based on the production of new materials. Poland's R&D prioritises nanomaterials processing, nano-metals for construction and functional applications, and polymers.

Support for nanotechnology is not limited to R&D. Canada, for example, identifies a number of programs at the federal and provincial levels for major infrastructure support and training. Denmark funds education programs at a number of universities. In Japan, METI supports the Nanotechnology Business Creation Initiative, with 300 corporations, to promote business-matching. MEXT operates a Nanotechnology Support Project which makes specialised facilities available to nanotechnology researchers. Other programs to support international exchanges of young researchers, entrepreneurship, technology licensing and industrial R&D are also expected to support innovation in nanotechnology. Poland has also supported nanotechnology education programs and is creating a virtual institute to strengthen collaboration among nanotechnology researchers in Poland and with

---

[44] OECD: 2004 Results of OECD mini-survey on nanotechnology R&D programmes DSTI/STI/TIP(2004)9

researchers in other countries. The UK's Micro and Nanotechnology Manufacturing Initiative includes funding for a new network of micro and nanotechnology facilities. It is open to public and private companies, research institutes and universities.

Two other programs are worth mentioning in addition to those covered by the OECD survey. In 2003, the United States passed the 21st Century Nanotechnology Research and Development Act, authorising $3.7 billion in federal subsidies for three years beginning in 2005. This is for projects supported by the National Nanotechnology Initiative (NNI), a federal R&D program established in 2001. Government funding for the NNI itself is projected to be $886 million for 2005, roughly 3% of overall US government R&D expenditure[45].

The EU's Sixth Research Framework Programme (FP6) includes nanotechnologies and nano-sciences, knowledge-based multifunctional materials and new production processes and devices among its priorities, with total Community support of around ¤1 billion over 2002-2006. The program's main objectives are the development of a successful European nanotechnology industry, and the application of nanotechnologies in existing industrial sectors. Additional nanotechnology research is supported by other parts of FP6. The follow-up, FP7, calls for almost 5 billion to be spent on nanotechnology over 2007-2013[46].

6. What are the risks of Nanotechnology?

6.1. Broad range of technologies, variety of risks

Keeping in mind the broad range of applications of nanotechnologies outlined in the previous chapters and the variety of industrial sectors that are affected, it is self-evident, that the risks associated with nanotechnologies will also form a complex risk landscape rather than a homogenous set of risks. The emphasis on what kind of risks are key to consider will depend on the perspective of the particular organisation involved in nanotechnologies. To name just a few:

- business risks involved with marketing of nanotechnology enabled products,
- risks related to the protection of intellectual property,
- political risks regarding the impact on the economical development of countries and regions,
- risks regarding privacy when miniature sensors become ubiquitous,
- environmental risks from the release of nanoparticles into the environment,
- safety risks from nanoparticles for workers and consumers,
- futuristic risks like human enhancement and self replications of nano machines.

The catch-all term ”nanotechnology” is not sufficiently precise for risk governance and risk management purposes. From a risk-control point of view it will be necessary to systematically identify those critical issues, which should be looked at in more detail. This risk identification process is a task for all parties involved and it should remain a dynamic process which always takes into account new scientific, technological, societal and legal trends.

This report mainly focuses on potential risks that are relevant to property and casualty insurance, rather than try to cover the broad range of topics indicated above.

Almost all safety concerns that have been raised about nanotechnologies are related to ”free” rather than fixed engineered nanoparticles. The state of the discussion on these risks will be discussed.

The risk and safety discussion related to free nanoparticles will concern only a fraction of the applications of nanotechnologies. In most applications nanoparticles will be embedded in the final product and therefore not come into direct contact with workers, consumers or the environment. They are unlikely to raise concerns because of their immobilisation. Exceptions are possible when the products or materials within which nanoparticles are enclosed are discarded, burned or otherwise destroyed (e.g. in an accident).

Looking at the manufacturing processes involved, they are similar to conventional well-established chemical

---

[45] www.nano.gov

[46] http://www.cordis.lu/nanotechnology/

and engineering processes. Building on established previous experience can be a guideline for risk assessment purposes, but the approach to health and safety issues needs to be modified to address the special characteristics of nanoparticles.

In the early phase of commercialisation most applications of nanotechnology will improve existing products, rather than generating new products. In themselves, most areas of nanotechnology are not likely to present novel safety risks. As with all new technologies new risks might appear that we have not thought about yet, underlining the need for continuous and dynamic risk reviews.

6.2. Positive effects on human health and the environment

A fair assessment of the risks of any new technology must also consider positive contributions to increased safety. The basic innovations that come from nanotechnologies have the potential to contribute to human health and environmental safety in many ways. They have the potential to contribute to solve urgent issues like the provision of clean drinking water or more efficient energy conversion and energy storage. The potential of nanotechnologies regarding economic benefits, the potential to create jobs, wealth and well being is very high. At the moment, public awareness about nanotechnology is limited. What happens over the next few years will determine how the public comes to view it. A transparent discussion of benefits and risks will help people reach a considered, balanced view. This will enable a greater public acceptance, which, in turn, will enable society as a whole to profit from these fundamental technological developments while, at the same time, the risks are kept under control.

Especially in the field of medicine there are quite a few technological developments that promise enhanced diagnostic possibilities, new ways to monitor patients, new ways to treat diseases like cancer and to reduce side effects. To give a few examples:

- Nanoparticles can be used as carriers for targeted drug delivery. Their ability to penetrate certain protective membranes in the body, such as the blood-brain barrier, can be beneficial for many drugs. This could open the way for new drugs from active substances that have not been able to pass clinical trials due to less precise delivery mechanisms,
- Nanosensors and lab-on-a-chip-technologies will foster early recognition and identification of diseases and can be used for continuous monitoring of patients with chronic diseases,
- New therapeutic methods for the treatment of cancer with the help of nanoparticles are investigated.

Ultrasensitive detection of substances will have implications for safety in many other areas such as industrial medicine, environmental medicine and food safety. To give one example: it has recently been shown that bacterial pathogens can be detected in very low concentrations with the help of nanoparticles. Quick and accurate testing is crucial for avoiding potential infections, but in order to be effective many current tests require time-consuming amplification of samples. The new methods are very powerful: new findings indicate that specially treated nanoparticles could allow to detect a single E. coli bacterium in a ground beef sample.

The potential benefits for our environment range from resource-efficient technologies reducing waste to new ways to transform and detoxify a wide variety of environmental contaminants, such as chlorinated organic solvents, organochlorine pesticides, and PCBs.

6.3. Manufactured nanoparticles

The term ”manufactured nanoparticle” is used here to refer to particles that have a physical size of less than 100 nm in at least two dimensions and that are deliberately produced rather than merely emerging as a by-product in activities not targeted for the production of these particles such as combustion processes or welding. While a more rigorous definition would have to take more parameters such as size distribution, diffusion diameter into account, the term nanoparticles shall be used in a rather broad sense here, including agglomerates and aggregates of the primary particles. Nanoparticles exist in various forms such as powders, suspensions or dispersed in a matrix.

In theory manufactured nanoparticles can be produced from nearly any chemical; however, most nanoparticles that are currently in use have been made from transition metals, silicon, carbon (carbon black, carbon nanotubes; fullerenes), and metal oxides. Quite a few of these nanoparticles have been produced for several decades on an industrial scale, but various new materials such as carbon nanotubes, fullerenes or quantum dots that have only been discovered within the last two decades. The development of nanomaterials is rapidly progressing and the variety of ”makes” is increasing with considerable speed.

The following list gives a few examples of manufactured nanoparticles that are of commercial interest[47]:

Type	Examples for use
Metal oxides	
• silica (SiO_2) • titania (TiO_2) • aluminia (AL_2O_3) • iron oxide (Fe_3O_4, Fe_3O_3) • zirconia (ZrO_2) • zinc dioxide (ZNO_2)	• Additives for polymer composites • UV-A protection • Solar cells • Pharmacy / medicine • Additives for scratch resistance coatings
Fullerenes • C_{60}	• mechanical and trbological applications / additives to grease
Carbon Nanotubes • Single-wall carbon nanotubes	• Additives for polymer composites (mechanical performance, conductivity)
• Multiwall carbon nanotubes	• Electronic field emitters • Batteries • Fuel cells
Compound Semiconductors • CdTe • GaAs	• Electronic an optical devices
Organic Nanoparticles	• Micronised drugs and chemicals (vitamins, pigments, pharmaceuticals) • Polymer dispersions
Metals • Au • Ag • Ni	• Catalytic applications • Optoelectronics • Wound dressings

To fully exploit the properties of nanoparticles technologically, not the composition is important, but also parameters like the exact size, morphology

---

[47] see for example: Wolfgang Luther (ed.), Industrial applications of nanomaterials – chances and risks VDI TZ Düsseldorf 2004

(spherical, nanotubes, nanowires, nanocrystals) and surface coatings.

6.4. Nanoparticles and human health

The economic growth in the field of nanotechnologies will lead to an increased variety and increased volumes of engineered nanoparticles that are produced. Even if exposure assessments and data are still lacking it is foreseeable that some degree of exposure to engineered nanoparticles -- for various segments of the population and for the environment – will occur to an increasingly extend over the coming years.

Keeping in mind that these ”free nanoparticles” can enter the human body over various pathways (inhalation, ingestion or via the skin) or disperse into the environment, it is important to understand the implications for human health and the ecosystems.

To assess the risks of nanoparticles, established methods of chemical safety assessments have to be modified to address the special characteristics of nanoparticles[48]. The main difference to the assessment of bulk materials is the fact that additional parameters like size, shape or surface properties will come into play. The same reason that makes nanoparticles technologically interesting leads to the fact that they represent a new category of (potentially) toxic substances. The interaction with the human body and their health effects are expected be different from molecules as well as from bulk materials of the same composition.

It is necessary to understand both, the hazards associated with nanomaterials and the levels of exposure, that are likely to occur. In both areas, the existing knowledge is quite limited and it will be necessary to generate and establish new data in the future.

In the following description the current status of the discussion on hazards and exposure of nanoparticles is summarised as a basis for the subsequent discussion of potential implications for the Allianz Group.

Hazards from engineered nanoparticles

When bulk materials are made into nanoparticles, they tend to become chemically more reactive – this is why they are very interesting as catalysts. Even chemically inert materials like gold or platinum are able to catalyse chemical reactions in nano-powder form.

Many studies indicate that nanoparticles generally are more toxic when incorporated into the human body than larger particles of the same materials. Experts are overwhelmingly of the opinion that the adverse effects of nanoparticles cannot be reliably predicted or derived from the known toxicity of the bulk material[49].

The biggest concern is that free nanoparticles or nanotubes could be inhaled, absorbed through the skin or ingested.

Uptake of nanoparticles via the lung

As pointed out in a recent review article by Hoet in the Journal of Nanobiotechnology[50], most nano-sized spherical solid materials will easily enter the lungs and reach the alveoli.

Inhaled particles can have two major effects on the human body:

- Their primary toxic effect is to induce inflammation in the respiratory tract, causing tissue damage and subsequent systemic effects[51,52]. The property that drives the inflammogenicity of nanoparticles is unknown but is expected relate to particle surface area and number of particles. Nanoparticles can impair the ability of macrophages to phagocytose and clear particles, and this may contribute to inflammatory reactions,

---

[48] NANOTECHNOLOGIES: A PRELIMINARY RISK ANALYSIS ON THE BASIS OF A WORKSHOP ORGANIZED IN BRUSSELS ON 1–2 MARCH 2004 BY THE HEALTH AND CONSUMER PROTECTION DIRECTORATE GENERAL OF THE EUROPEAN COMMISSION.
http://europa.eu.int/comm/health/ph_risk/events_risk_en.htm

[49] Dreher, Kevin L., Health and Environmental Impact of Nanotechnology: Toxicological Assessment of Manufactured Nanoparticles. TOXICOLOGICAL SCIENCES 2004, 77, 3–5

[50] Hoet, Peter HM., Brüske-Hohlfeld, Irene, and Salata Oleg V. Nanoparticles – known and unknown health risks, Journal of Nanobiotechnology 2004, 2:12

[51] Lam C-W. et al. Pulmonary Toxicity of Single-Wall Carbon Nanotubes in Mice 7 and 90 Days After Intratracheal Instillation. TOXICOLOGICAL SCIENCES 2004, 77, 126–134

[52] Warheit D. B. et. al., Comparative Pulmonary Toxicity Assessment of Single-wall Carbon Nanotubes in Rats. TOXICOLOGICAL SCIENCES 2004, 77, 117–125

- Transport through the blood stream to other vital organs or tissues of the body[53]. This may result in cardiovascular and other extrapulmonary effects.

Some scientists have compared nanotubes with asbestos in terms of risks, because they resemble asbestos fiber in their needle like shape. The concern is particularly applicable to fibers of high biopersistence[54]. Although the comparison seems plausible, it has been pointed out, that the nanotube fibers tend to clump together rather than exist as single fibers, thus possibly significantly reducing exposure and their potential to do harm[55].

Uptake via other routes

The scientific literature shows that particles in the nano-size range can also enter the human body via other pathways, i.e. the upper nose and the intestines.

Penetration via the skin seems less evident[56], but research to clarify this is under way[57]. The chances of penetration again depend on the size and surface properties of the particles and also strongly on the point of contact. If nanoparticles penetrate the skin they might facilitate the production of reactive molecules that could lead to cell damage. There is some evidence to show that nanoparticles of titanium dioxide (used in some sun protection products) do not penetrate the skin but it is not clear whether the same conclusion holds for individuals whose skin has been damaged by sun or by common diseases such as eczema. There is insufficient information about whether other nanoparticles used in cosmetics (such as zinc oxide) penetrate the skin and there is a need for more research into this. Much of the information relating to the safety of these ingredients has been carried out by industry and is not published in the open scientific literature.

Distribution in the body

Once in the body, the distribution of the particles in the body is strongly dependent on the nanoparticle in question, e.g. the composition, the size and of the surface characteristics of the particles. It seems that translocation occurs from all uptake routes. Such translocation is facilitated by the propensity to enter cells, to cross cell membranes and to move along axons and dendrites that connect neurons. Surface coatings will have a major effect. The mobility of different types of nanoparticles requires detailed investigations, which have not yet been performed.

To give one example: it is not clear, whether nanoparticles can pass from a pregnant woman's body via the placenta into the unborn child.

It is possible that durable, biopersistent nanoparticles may accumulate in the body, in particular in the lungs, in the brain and in the liver.

For the majority of nanoparticles the toxicological, ecotoxicological data needed to perform a hazard analysis are still lacking. Even if the details are not yet clear, it is evident that the interaction with the human body will depend on various parameters such as chemical composition, particle size, surface area, biopersistence and surface coatings among others. Therefore, until a theory of the impact of nanoparticles on human health has been established, each nanomaterial should be treated individually when health hazards are evaluated. A systematic risk screening will be helpful to establish the basic know-how to understand the interaction with the human body and the environment and to establish the theoretical framework needed.

Besides toxic effects, the interaction of nanoparticles that have entered cells opens a wide field of potential effects resulting from the interaction with cell structures such as ribosomes and DNA.

Exposure to engineered nanoparticles

Injury can be caused by chemicals only if they reach sensitive parts of a person at a sufficiently high

---

[53] Nemmar et al, "Passage of inhaled particles into the blood circulation in humans", Circulation 2002;105(4):411-41

[54] Hoet, Peter HM., Health Impact of nanomaterials? Nature Biotechnology 2004, Volume 22, Number 1

[55] Maynard, A. D. Exposure to Carbon Nanotube Material: Aerosol Release During the Handling of Unrefined Single-Walled Carbon Nanotube Material. Journal of Toxicology and Environmental Health Part A Volume 67, Number 1 / January 9, 2004

[56] Pflücker F., The human stratum corneum layer: an effective barrier against dermal uptake of different forms of topically applied micronised titanium dioxide. Skin Pharmacol Appl Skin Physiol. 2001;14 Suppl 1:92-7

[57] see for example the NANODERM research project funded by the European Commission http://www.uni-leipzig.de/~nanoderm/Brochure_NANODERM_WWW.pdf

concentration and for a sufficient length of time. Besides the physicochemical properties of the nanoparticles, the nature of exposure circumstances and the health state of the persons at risk have to be considered.

Nanoparticles exist in nature (e.g. from volcanic eruptions or forest fires) or can be produced by human activities. Intentional nanoparticles are manufactured under (normally strict) control while unintentional ones can come from high temperature combustion, explosions, mechanical abrasion or other industrial processes. The main source of unintentional nanoparticles (in this context also called ultra-fine particles) is automobile traffic.

We live surrounded by nanoparticles. To provide an example: a normal room can contain 10,000 to 20,000 nanoparticles per cm^3, whilst these figures can reach 50,000 nanoparticles per cm^3 in a forest and 100,000 nanoparticles per cm^3 in urban streets.

At present, the nanoparticles originating from dedicated industrial production are marginal in relation to those produced and released unintentionally, such as through combustion processes.

However, higher levels of exposure are expected with the industrial processes in which nanoparticles are intentionally produced or used. Furthermore, as manufactured nanomaterials become more widespread in use, the range of scenarios in which exposure becomes possible will increase.

Exposure assessments for engineered nanoparticles should be able to cover all identified uses for the entire life-cycle of the individual nanoparticle. This could include:

- processes involved in the production,
- processes involved in the identified use,
- activities of workers related to the processes and the duration and frequency of their exposure,
- risk management measures to reduce or avoid exposure of humans,
- waste management measures,
- activities of consumers and the duration and frequency of their exposure.

Nanoparticles are currently being produced in low volumes and, aside from their use in cosmetics, involve as yet little or no exposure to populations outside the workplace. However, keeping in mind the predicted growth in production volumes and the expected expansion of the product range, this situation is subject to change over the next few years.

With an increasing number of companies involved in nanotechnologies, the quality of risk management encountered in practice will be ranging from highly sophisticated to poor. The level of risk management often depends on the branch of industry, but mainly on individual management practises and management attention in the companies. Accordingly, the assessment of possible exposures must take into account poor risk management practises, especially in the absence of specific regulations.

With respect to hazards, there is enough evidence to suggest that exposure to nanoparticles, particularly to those insoluble in water, should be minimised as a precaution.

Occupational hazards

The US national nanotechnology initiative has estimated that around 20,000 researchers are working in the field of nanotechnology. For the UK, the Institute of Occupational Medicine has estimated that approximately 2,000 people are employed in new nanotechnology companies and universities where they may be potentially exposed to nanoparticles[58].

The primary production of nanoparticles takes places with the help of chemical and engineering processes that are well established. These are considered by various organisations to be relatively safe, except for accidental releases via leakages. During normal production, subsequent steps like product recovery and powder handling may result in respirable concentrations of agglomerated nanoparticles, to dermal exposure or to ingestion (mainly via hand-to-mouth contact).

So far only one study has addressed the exposure of workers to engineered nanoparticles. In the study into potential airborne and dermal exposure to carbon nanotubes samples were taken during typical activities in the production process like material removal, filling, pouring and clean up (techniques: laser ablation and high pressure carbon monoxide). The study has shown that the material only becomes airborne with a sufficient

---

[58] Institute of Occupational Medicine. Nanoparticles: An occupational hygiene review Research Report 274. Edinburgh 2004

level of agitation. So-called Van der Waals forces make the nanoparticles ”sticky” and imparts a strong tendency to form agglomerates. Airborne concentrations found while handling unrefined material were considered to be ”very low” (between 0.7 and 53 $\mu g/m^3$). The released material seems to form larger agglomerates in the size range of 1 µm rather than leading to a high number concentration of fine particles. Up to 6 mg of nanotubes were found on individual gloves. While the gloves can minimise dermal exposure, airborne clumps of material can lead to exposure of less well protected parts of the skin[59].

Contrary to intuition, air filtration systems such as respiratory protective equipment should be effective when used correctly. Below a particle size of 100 nm filtration efficiency for High Efficiency Particulate Arrestor (HEPA) filters even increases with decreasing particle size. The reason for this is strong Brownian motion that leads to increased probability to contact the filter elements. Once a filter element is hit, strong Van der Waals forces keep the nanoparticle stuck to the surface.

It has been inferred by the UK Institute of Occupational Medicine that the primary route of exposure for many nanoparticles could be dermal exposure and subsequent ingestion exposure.

Generally accepted, realistic methods for exposure assessments are still lacking for workplaces. These methods must be biologically relevant—that is, they should be able to measure the most appropriate metric that characterises the exposure. That can be surface area in the case of airborne particles, but also mass or number of particles in the case of dermal exposure or ingestion. Size selective sampling methods will be needed to ensure that only the relevant size range is sampled. It remains a technical challenge to develop effective methods and standards for controls. For some nanoparticles it may be necessary to measure and control to very low levels on the order of ng/m^3.

As with hazard aspects of nanoparticles, also in the area of exposure there are many unknowns today and also technical challenges lie ahead. As an intermediate state between bulk material and individual molecules the characteristics of nanoparticles add some complexity to the exposure assessment (such as surface area, agglomeration, size distributions) which can be reduced once a better understanding, better measurement methods and standards for these particles will exist. It seems that existing precautionary measures are also effective for the control of nanoparticles, however in the absence of more evidence this is difficult to demonstrate.

To quote the report of the Institute of Occupational Medicine to the UK’s Health and Safety Executive: ”In summary, we conclude that there is little evidence to suggest that the exposure of workers arising from the production of nanoparticles has been adequately assessed.”

Exposure of consumers to products

There are a number of ways in which engineered nanoparticles from products can come into direct contact with consumers during their use. In some applications these nanoparticles are active parts of the product, in some applications only an accidental exposure is possible.

Due to the fact that nanoparticles in powder form are only used as intermediate products during stages of the production process, it seems likely that the exposure from consumer products will mainly occur via ingestion and via the skin. It is unlikely that engineered nanoparticles that are bound in a matrix or somehow fixed in a product are released. Manipulations like grinding or cutting do not necessarily release nanoparticles, but rather more likely particles of larger size in which the nanoparticles are still bound.

The following table gives some examples of existing end products and end products that will be marketed in the foreseeable future, and which can lead to an exposure to certain nanoparticles:

---

[59] Maynard, A. D. Exposure to Carbon Nanotube Material: Aerosol Release During the Handling of Unrefined Single-Walled Carbon Nanotube Material. Journal of Toxicology and Environmental Health Part A Volume 67, Number 1 / January 9, 2004

Application	Functionality of nanoparticle	Example
Cosmetic products	• active ingredient	• suntan lotions (transparent UV protection) • lotions • tooth paste
Medical products / devices		• wound dressings
Food products	• improved production processes • pesticide • enhanced food properties • extended shelf life • transport of nutrients and flavours into the body	• drinks • ice cream • fish • …
Automotive	• explosive	• airbag gas generators
Sports articles	• enhanced gliding	• ski wax

Whether or not there are health risks involved from particular applications will probably be subject to considerable debate in the near future. It is a challenge to create an approach to exposure assessments that is fit for such diverse applications with differing nanomaterials and different potential exposure paths. Depending on the type of product different regulatory regimes apply and have to be accounted for. Comparative risk classification schemes will have to be developed that can serve as a guideline / road map in the overall risk identification process. Generally, to demonstrate the control of the exposure a life cycle assessment will be needed.

While the overall approach will probably not differ substantially from existing product safety assessments, it will take some time until the basis to follow this path has been established and answers related to the special characteristics of nanoparticles have been found.

Open points linked with the exposure to nanoparticles include:

- which parameters characterising the particles should be measured?
- the development of efficient measurement methods for these parameters,
- testing methods for various applications and situations during use of the products,
- fate of various types of nanoparticles in the environment
- fate of the nanoparticles when a product is destroyed, burnt or discarded,
- extent of uptake in the human body.

Conclusions about the safety of engineered nanoparticles

With the production of engineered nanoparticles we are confronted with a new class of materials that have novel properties compared to bulk material. Information describing the health risk of engineered nanoparticles is only evolving and many questions are still open.

The uncertainties involved have to be seen against the background of the ever more demanding public views regarding the safety of products from new technologies and an increasing potential for exposure as the quantity and types of nanoparticles used in society grow.

From animal experiments and analogies to studies on incidentally produced ultra fine particles (such as . from burning of fuels) it is possible that at least some nanoparticles are hazardous for the human body and that the exposure to these nanoparticles should be avoided or at least minimised.

One promising path to prevent potential health hazards proposed by a number of scientists is to make the particles biodegradable. Particles that are degradable either by water or by lysis with enzymes will greatly reduce the risks involved because they do not persist in the body. It is self-evident that the slower the particles are cleared (high persistence), the higher the tissue burden can be. With a short bio-durability, long-term effects can be minimised or even excluded. Biocompatability can serve as one major engineering parameter for nanomaterials in the future.

There is a broad variety of nanoparticles being investigated and a high number of parameters that influence the functionality of these nanoparticles as well as their interaction with the human body. It is expected that it will take several years until some major critical risk assessment issues regarding hazards and exposure can be answered. Accordingly, in a workshop in January 2005, the European Commission has identified a variety of research needs related to nanoparticles. The recommendations for research include:

- ecotoxicity of nanoparticles for cases of inadvented contact by children, adults and susceptible individuals,
- biomonitoring studies of specific biological impact of nanoparticles, according to their most likely routes of exposure,
- standards development and (certified) reference materials,
- sensors for detecting nanoparticles and assessing exposure, both stationary and portable with a focus on low-cost, highly specific sensors giving a real time response to environment an health relevant properties,
- a strategic concept for a comprehensive risk assessment.

The regulatory risk assessment of chemicals will in the future be carried out under the New European Union Chemicals Policy (Registration, Evaluation and Authorisation of Chemicals). In its current state REACH will be concerned only with the safe use of bulk materials, and will not take account of size. Additionally, under REACH production mass limits are planned (limited evaluation procedure for production of less than 100 tons, no registration for production volume below 1 ton). Many nanoparticles that are produced in low volume will therefore generally not be accounted for in the REACH framework. Further examinations of this regulatory situation might be required on the basis of new scientific evidence related to manufactured nanoparticles.

It will be a challenge for industry, legislators and risk assessors to fill all relevant knowledge gaps and to construct a set of high throughput and low cost tests for nanoparticles. In that way a risk assessment of each nanoparticle should be established before the larger quantities are manufactured. This process is unlikely to be fast enough without active steering and support by governments and the EU.

The uncertainties involved – especially long term -- will have to be addressed by all organisations involved in the process of the introduction of nanotechnologies. It will be a necessary for the various industries involved to perform life-cycle assessments and demonstrate the safety aspects to a broader public. Successful communication will depend on a general feeling of trust towards nanotechnologies. This in turn needs an open dialogue and interaction involving all different stakeholders, including a high degree of transparency regarding scientific results.

6.5. Nanoparticles and the environment

As nanotechnologies move into large-scale production in many industries, it is a just a matter of time before gradual as well as accidental releases of engineered nanoparticles into the environment occur. The possible routes for an exposure of the environment range over the whole lifecycle of products and applications that contain engineered nanoparticles:

- Discharge / leakage during production / transport and storage of intermediate and finished products,
- Discharge / leakage from waste,
- Release of particles during use of the products,
- Diffusion , transport and transformation in air, soil and water.

Some applications like cosmetic products or food ingredients will be diffuse sources of nanoparticles.

In addition, certain applications such as environmental remediation with the help of nanoparticles could lead to the deliberate introduction of nanoparticles into the environment. This is an area which will probably lead to the most significant releases in terms of quantity of nanoparticles in the coming years.

The main criteria used to assess the risks of chemicals for the environment and indirectly for human health are toxicity, persistence and bioaccumulation. Substances that can cause direct damage to organisms (high toxicity), that decay very slowly in the environment (high persistence) and that can concentrate in fatty tissues (high potential for bioaccumulation) are of particular concern. For engineered nanoparticles the particular characteristics of nanomaterials will have to be taken into account for a specific risk assessment. The existing information about properties of the bulk material will not be sufficient to classify the environmental risk of the same material in the form of nanoparticles. The possible environmental effect will therefore have to be assessed specifically for each type / class of nanomaterial.

Only few studies on this very complex subject exist. From a scientific point of view, the results should be seen as indications rather than a sound basis for decision making.

In the first study on the toxic effects of manufactured nanoparticles on aquatic organisms, fish (largemouth bass) were exposed to uncoated fullerene carbon 60 (C_{60}) nanoparticles[60]. The fullerenes are one type of manufactured nanoparticle that is being produced by tons each year. Significant lipid peroxidation (oxidation of fats) was found in the brain of the animals after exposure to 0.5 ppm uncoated nC_{60}. The study demonstrates that manufactured nanomaterials can have adverse effects on aquatic and possibly other organisms.

Nanoscale iron particles have been investigated as a new generation of environmental remediation technology[61]. Due to their high surface reactivity and large surface area they can be used to transform and detoxify environmental contaminants like PCBs. Field tests in the US have shown that the nanoparticles remain reactive in soil and water for several weeks and that they can travel in groundwater as far as 20 meters. The risks associated with free nanoparticles on ecosystems was not discussed in the original publication, but should be looked at in sufficient detail before environmental applications are brought to the market. The Royal Society has called for the prohibition of the use of free nanoparticles in environmental applications until appropriate research has been undertaken.

A very specific environmental issue in the case of nanoparticles is their propensity to bind with other substances, possibly toxins in the environment such as Cadmium. Their high surface area can lead to adsorption of molecular contaminants. Colloids (natural micro- and nanoparticles) are known for their transport and holdings capacity of pollutants. The adsorbed pollutants could possibly be transported over longer distances / periods of time by nanoparticles.

On the other hand nanoparticles are less mobile than we intuitively might think. It seems that their movement is very case specific and that that are generally less mobile than larger particles. Here again their large surface area and their maximised chemical interaction comes into play. Their sticky nature considerably slows their transport through porous media like soil.

In summary, the information about nanoparticles and the environment is only at an early stage. Among the research needs, there are topics like the effect of nanoparticles on species other than humans, about how they behave in the air, water or soil, or about their ability to accumulate in food chains. Taking into account the high number of parameters that characterise nanoparticles (like size, shape, specific surface treatment, chemical composition) as well as the variety of nanoparticles, it will need considerable research efforts to close the knowledge gaps. For cost efficient and quicker results, harmonisation of research is required that focuses on the most important materials and parameters and to concentrate on nanoparticles that are more likely to be produced.

6.6. Explosion hazards of nanoparticles

For many industries, the explosion of dust clouds is a potential hazard in the production process. A dust explosion occurs when a combustible material is dispersed in the air, forming a flammable cloud which is hit by a flame. The concentrations needed for a dust explosion are rarely seen outside of process vessels, hence most severe dust explosions start within a piece of equipment (such as mills, mixers, filters, silos).

Various materials that are not stable oxides can be involved in dust explosions, e.g. natural organic materials (grain, sugar, etc); synthetic organic materials (organic pigments, pesticides, etc) coal and peat metals (aluminium, zinc, iron, etc). There is a clear dependence on size and surface area of dust particles, it does however not vary linearly with the explosibility of the powders. Dust with a particle size above 0.4 mm is normally not explosive.

---

[60] Oberdorster E. Manufactured nanomaterials (fullerenes, C60) induce oxidative stress in the brain of juvenile largemouth bass. Environ Health Perspect. 2004 July; 112(10): 1058-62

[61] Zhang W-X. Nanoscale iron particles for environmental remediation: An overview. Journal of Nanoparticle Research 5: 323–332, 2003

In certain applications, the stronger reaction of smaller particles can be used to pack more explosive power in a given volume. Today, microsize aluminium particles already gets incorporated into rocket fuels and bombs. Nanoaluminium is being investigated as advanced technology in these military applications. Aluminium nanopowders could also find their way into airbag gas generators. The toxicity is being investigated and has to be compared to the properties of existing propellants and their combustion products (particles and gases).

The explosibility of nanopowders has so far not received much attention in the public debate on health and safety risks of nanotechnologies.

A review of literature available on this topic has been performed by the Health and Safety Laboratory in the United Kingdom[62].The review could not find data for nanopowders with particle sizes in the range of 1 to 100 nm. It states that the extrapolation of existing data for larger particles to the nano-size range is not possible because of the changed physical and chemical characteristics on the nano-scale.

The report recommends that the explosion characteristics of a representative range of nanopowders be determined since an increasing range of materials that are capable of producing explosive dust clouds are being produced as nanopowders. It is foreseen that the production of nanopowders is likely to increase significantly over the next few years.

The report of the Royal Society recommends that the explosion risks be managed by ensuring that large quantities of combustible nanoparticles do not become airborne until the explosion hazard has been properly evaluated. As long as nanopowders are produced in small quantities of grams, the explosion hazard will be negligible.

Besides property damage, workers safety and business continuity issues, one major concern about the explosion of nanopowders is the release of larger quantities of nanomaterial into the environment and the resultant potential pollution problem.

6.7. Self replication of miniature machines

In 1986 Eric Drexler published an influential book called ”Engines of Creation: The Coming Era of Nanotechnology,” in which he imagined the fabrication of molecular machines[62]. These machines would be able to produce any (macroscopic) item from molecular building blocks. For this plan to work, these machines would have to be able to produce machines of their own kind, a process called self replication. His most compelling argument about the feasibility of these machines is the observation that biology gives us many examples of nanoscale machines that function on this scale.

The idea that engineering a synthetic form of life with self-replicating machines has, in turn, created the fear that once designed, these nano-robots could spread across the biosphere. Drexler called this scenario the ”gray goo”.

This view has been challenged by many scientists. Most scientists have dismissed the ”gray goo” scenario as ”science fiction”. There are fundamental questions that have not been resolved so far, like the energy management, the strong surface forces on the nanoscale or Brownian motion. After all, biological systems in our environment have been optimised over billions of years of evolution and make extensive use of the particular physical characteristics that govern the nanoscale world, which we are just beginning to understand. Our knowledge even of the processes in a simple cell is limited.

It seems therefore safe to say that the construction of self-replicating nano-robots will remain beyond our capabilities for the foreseeable future. From an insurance perspective the risks associated with self-replication of machines will therefore remain futuristic and will -- in all probability -- have no relevance over the next decade.

---

[62] D K Pritchard, Health and Safety Laboratory. Literature review – explosion hazards associated with nanopowders HSL/2004/12

[62] K. Eric Drexler Engines of Creation: The Coming Era of Nanotechnology, Anchor Books 1986 ISBN 0-385-19973-2

6.8. Regulatory considerations of authorities and other stakeholders

In September 2004, a paper on possible emerging safety issues of nanotechnologies was submitted to the Joint Meeting of the OECD's Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology[65]. The paper points out that nanotechnology exploits properties not generally seen in large-scale solids of the same chemical composition, and that these same properties have led to health and safety concerns. For example, the high surface reactivity of nanoparticles and the ability to cross cell membranes might have negative human health impacts. On the other hand, not all nanoparticles, and not all uses of nanotechnology, will necessarily lead to new human health or environmental hazards. A distinction can be already be made, for example, between free and fixed nanoparticles. Fixed nanoparticles are less likely to pose a problem because they are immobilised within a matrix and cannot freely move or disperse within the human body or the environment. In other words, there is likely to be low human exposure.

It may be possible to assess many products involving nanoparticles from the chemicals industry using existing mechanisms or techniques for risk assessment, but it may often be necessary to adapt existing techniques or devise new risk assessment methods. In any case, the OECD paper argues, it will be important to address safety issues in a proactive way in order to avoid public uneasiness about the new technology as a result of a lack of information and attention to safety aspects which could lead to limitations on the safe development of a promising technology. There will be a Special Session of the Joint Meeting in June 2005 to investigate in more detail the potential safety implications of manufactured nanomaterials for human health and environmental safety.

The International Risk Governance Council (IRGC) discussed nanotechnology at a meeting of its Scientific and Technical Council in February 2005[65], agreeing that it is a broad technology which raises important long-term issues (including societal issues) but arguing that grouping all nanoscale R&D and potential products within the single term ”nanotechnology” makes it difficult to distinguish between technologies with different risk governance issues. The situation is complicated by the fact that some possibly harmful commercial applications, such as in cosmetics, already exist and are subject to little new control, while potentially beneficial technologies may be retarded. Drug delivery, which may turn out to be the most important area economically, will take some time to come to market because of the time needed to complete approval processes. The drug example raises the possibility that nanotechnology may suffer from a new problem, termed ”toxicology bottleneck” as the science advances too fast for the risk assessment processes of organisations such as the US FDA. This is further complicated by the difficulty in both measuring and modelling nanoscale particle behaviours in both air and fluids.

The report by the UK's Royal Society and Royal Academy of Engineering on nanotechnologies expects the likelihood of nanoparticles or nanotubes being released from products in which they have been fixed or embedded (such as composites) to be low, but recommends that manufacturers assess this potential exposure risk for the lifecycle of the product and make their findings available to the relevant regulatory bodies[66]. The report highlights the lack of knowledge about the behaviour of nanoparticles — for example, the effects of inhaling free manufactured nanoparticles have not been studied extensively. Analogies with results from studies on exposure to other small particles such as the pollutant nanoparticles in urban air and mineral dusts in some workplaces suggest that at least some manufactured nanoparticles will be more toxic per unit of mass than larger particles of the same chemical. It also seems likely that nanoparticles will penetrate cells more readily than larger particles. The report also emphasises the diversity of technologies lumped together under the term ‘nanotechnology’ and the implications of that diversity for the approach to public dialogue, research and regulation.

---

[64] OECD: 2004 ”Nanotechnology: Emerging safety issues?” ENV/JM (2004)32

[65] Unpublished. The IRGC's June 2004 factsheet on nanotechnology can be found here http://www.irgc.org/_cgidata/mhscms/_images/12384-3-1.pdf

[66] The Royal Society & The Royal Academy of Engineering: 2004, ”Nanoscience and nanotechnologies: opportunities and uncertainties” www.nanotec.org.uk/finalReport.htm

In its response to the Royal Society and Royal Academy of Engineering Report from February 2005 the Minister of Science and Innovation sets out the UK's government's agenda on nanotechnologies[67]. Planned steps to ensure a safe and ethical development of nanotechnologies include:

- setting up a research co-ordination group to investigate risks from nanoparticles,
- Initiatives for public dialogue to help the scientific community and the public to explore issues relating to the regulation of nanotechnologies,

It is also planned to review the adequacy of the current regulatory frameworks. Specifically, the government wants to work with the EU regulatory authorities with respect to:

- the assessment of risks associated with medicines and medical devices,
- the safety of unbound nanoparticles in cosmetics and other consumer products,
- disclosure of testing methodologies used by industry,
- labelling requirements on consumer products,
- sector specific regulations for products of nanotechnologies in addition to REACH at a European level.

At present no specific regulations exist in Europe which refer specifically to the production and use of nanoparticles either for workers or consumers' safety or for environmental protection. Current rules and operational practices are applied.

There are several research projects funded by the European Commission which deal with nanoparticles (eco)toxicity and risk. The NanoDerm project investigates the quality of skin as a barrier to ultra-fine particles. NanoSafe and NanoSafe2 (in negotiation) study risk in production and use of nanoparticles. NanoPathology investigates the role of micro- and nanoparticles in inducing biomaterial disease. Other projects are currently being negotiated such as the specific support action NanoTox that aims to provide support for the elucidation of the toxicological impact of nanoparticles on human health and the environment, and the co-ordinating action Impart that aims to improve understanding of the impact of nanoparticles on human health and the environment. Moreover, the German project NanoCare will develop an inhalation toxicity model for testing of nanomaterials.

In the proceedings of a workshop from January 2005, the EU Commission states that further examinations of the regulation may be needed on the basis of new scientific evidence.

6.9. Position of the industry

Over the last few years national nanotechnology associations have been established and / or industry associations have formed subsections that deal with nanotechnologies in various countries. A number of theses associations have initiatives that deal with risks and risk management related to nanotechnologies.

As one example, the German DECHEMA (Society for Chemical Engineering and Biotechnology) and the VCI (German chemical industry association) have founded a working group on ”Responsible Production and Use of Nanomaterials”. The two organisations represent the interests of 1,600 German chemical companies as well as 5,000 private members. The working group has actively promoted a risk management approach to nanotechnologies. The group aims at the successful realisation of the economical and technological chances by initiation of suitable measures, taking into account ethical, ecological, social and health care aspects. The tasks of the working group are:

- To identify and prioritise research topics which have to be addressed in order to assess possible risks of nanomaterials,
- To prepare project proposals on the basis of identified research topics and to support their realisation,
- International co-operation to gain synergies and to widen the database,
- Dialogue with stakeholders,
- Promoting related communication.

On the risk management agenda, the emphasis of the working group lies on identifying possible risks of chemical nanotechnology focusing on nanoparticles and nanotubes in both their free and bounded state

The working group favours a structured approach by assessing hazards and exposure and implementing risk management measures such as:

- occupational protection measures (organisational, technical, and individual) for workers,

---

[67] HM GOVERNMENT UK. RESPONSE TO THE ROYAL SOCIETY AND ROYAL ACADEMY OF ENGINEERING REPORT:'Nanoscience and nanotechnologies:opportunities and uncertainties', February 2005

- use of embedded nanomaterials, use of nanomaterials chemical fixed / covered at surfaces for consumer products as well as individual toxicological testing and governmental authorisation.

This approach reflects the attitude of industry representatives, who generally call to explore the potential risks case by case, application by application and material by material. One basis for this approach is establishing databases for the scientific evaluation of risks. It is not clear whether public access to this kind of databases is intended.

Industry associations on an international basis are still in their very early stages. A new European association to represent the nanotechnology business, the Nanotechnology Trade Association (ENTA), will be set up in June 2005. ENTA is an association formed for the nanotechnology industry in Europe and aims to liaise between governments, science and industry. It is supported by several major multinational companies such as British Petroleum, and Proctor and Gamble, and it will share information explore, networking opportunities, disseminate news and also perform risk evaluations.

6.10. Position of pressure groups

One small non-government organisation called the ETC group has published several reports raising concerns about nanotechnology that have been regularly quoted in the international press[69,70]. The head of ETC, Pat Mooney, is best known for his role in the fight against Monsanto's genetically modified seeds.

In 2002, the ETC Group called for a moratorium on the commercialisation of new nano-scale materials until laboratory protocols and regulatory regimes are in place that take into account the special characteristics of these materials, and until they are shown to be safe.

In a recent publication called ”Down on the Farm”, the ETC Group recommends specifically that all food, feed and beverage products incorporating manufactured nanoparticles be removed from the shelves and new ones be prohibited from commercialisation until companies and regulators have shown that they have taken nano-scale property changes into account. They also call for the prohibition of nano-scale formulations of agricultural products such as pesticides and fertilisers from environmental release until a regulatory regime specifically designed to examine these nano-scale products finds them safe.

A report for the Greenpeace Environmental Trust called for a commitment by industry to environmentally sound practises and to fund relevant research on human health on a far greater scale than currently witnessed. Greenpeace does not, however, echo the call for a nanotech moratorium, because that ”seems both unpractical and probably damaging”[71].

6.11. Position of reinsurers and insurers

The topic of nanotechnologies has been taken up by the major re-insurers, mainly Swiss Re[72], Munich Re[73] and GenRe[74].

They all agree that the insurance industry is going to have to live with the uncertainties of nanotechnology related risks for a longer period of time and that it will not be able to quantify the probability of potential losses occurring and their possible extent.

In principle, many lines of business are considered to be potentially affected, including:

- Workers’ compensation,
- General and products liability
- Products recall,
- Environmental liability,
- Property (dust cloud explosion).

It is assumed that if health effects from certain engineered nanoparticles become ever manifest and causation can be established, then series of claims will almost certainly follow.

Currently, there are no specific policy exclusions or terms in regular use that are tailored to address risks from nanotechnologies.

---

[69] ETC Group. Nanotech: Unpredictable and Un-Regulated 2004

[70] ETC Group. DOWN ON THE FARM: The Impact of Nano-Scale Technologies on Food and Agriculture. November 2004

[71] Greenpeace Environmental Trust / Imperial College London. Future Technologies, Today's Choices. Nanotechnology, Artificial Intelligence and Robotics. A technical, political and institutional map of emerging technologies July 2003

[72] Swiss Reinsurance Company. Nanotechnology Small Matter - Many Unknowns 2003

[73] Munich Re Group Nanotechnology – What’s in store for us? Topics 2003/1

[74] GenRe, Nanotechnology: Will Minute Items have a Huge Impact on the P/C Industry? Hazardous Times 2004

One major issue for the re-insurers is the fact that the extent of nanotechnology risks will become apparent only in the long term. Dangers will be chronic rather than acute. The fears here are that ”unforeseeable, ruinous loss accumulation unleashed by a flood of late claims” might build up. As nanotechnology becomes pervasive and certain nanomaterials are found to cause illness, many insureds (policy holders) are likely to be affected.

From the fact that the risk is considered to be incalculable, Swiss Re concludes that the insurance industry will have to work with loss scenarios and certain loss-limiting measures to protect the balance sheet.

A major issue is the ”stacking of limits” problem: insurers using traditional occurrence policies in some jurisdictions may face trigger obligations under multiple policies over several years. Both Swiss Re and GenRe call for an avoidance of this problem by ”claims made” covers and exact descriptions under which a loss may be said to have occurred. GenRe states that today most nanotechnology risks are written on occurrence policies.

7. Chances and risks for the Allianz Group

7.1. Nanotechnologies and investments

The predicted strong overall economic impact of nanotechnologies means that probable commercial success stories will run parallel to a demand for financing solutions. On the other hand, as with other new technologies, business risks depend not only on selecting promising business concepts but also on the timing of investments.

Private equity and venture capital

A survey by the VDI (the Society of German Engineers) shows that one of the major barriers for innovation by small- and medium-sized enterprises in Germany in the area of nanotechnology is a lack of capital. To develop new products and processes and also to penetrate new markets, sizeable investments are needed, especially in the seed phase. The VDI calls for closer co-operation between the financial community and nanotechnology companies[75].

Venture capital firms in nanotechnology will have a key role in transferring technology knowledge from the research centers to the industry and the markets. According to Business Week, at the end of 2004 venture capitalists had already invested $1 billion in nano companies, nearly half of it alone in 2003 and 2004.. It is expected that the majority of nanotech exits will be through trade sales.

For successful investments, two factors will be of critical importance: timing and target selection.

Timing

The overwhelming majority of observers agree that, looking at the broad picture, the question is not whether nanotechnologies will reach the market, but how long they will take to make a broad impact on real products and, in turn, create sizeable turnover and profits

This time scale is likely to vary widely for individual products and companies simply because of the vast number of activities bundled together under the term ”nanotechnologies.”

This can be illustrated by three examples:

- ➞ Short term: equipment manufacturers already have a broad customer base as their tools are used in nanotech laboratories all over the world. They are creating the basis for progress in nanotechnologies.
- ➞ Mid term: several major electronics companies plan to produce TV and computer displays featuring carbon nanotubes. These screens will be lighter, cheaper, brighter, and more energy-efficient than today's models and the market associated with this application is huge. These screens will be in stores around 2006.
- ➞ Long term: even after successful tests on animals, it will take years until a general medical clearance is given for nanobiotechnological cancer therapy. On average, its takes more than 10 years for a drug to reach the market.

---

[75] W. Luther, N. Malanowski ”Nanotechnologie als wirtschaftlicher Wachstumsmarkt, Innovations- und Technikanalyse”, VDI, Düsseldorf September 2004 (in German)

Selection

In contrast to the insurance sector, with its high number of insureds (policy holders), private equity and venture capital give detailed consideration to a smaller number of ”targets.”

Careful target selection will be necessary to minimize early stage investment risks. This should involve in-depth analyses not only of market, legal, and financial aspects, but also of technical and environmental factors in a process of ”acquisition due diligence”
In many cases, a sound scientific understanding combined with know-how about commercialisation will be key factors.

The health and environmental risks outlined above can create secondary risks for companies that produce and market nanotech products. The discussion on nanotechnology will strongly influence the marketability – and thereby the commercial success -- of certain products. The likelihood of bad investments should be minimized by carefully reviewing the evidence about risk.

Public companies

Because nanotechnology is an enabling technology, it offers new possibilities for thousands of materials and products that already exist. This necessarily involves large industrial companies such as Degussa, Bayer, DuPont, and General Electric. Already, 19 of the 30 companies in the Dow Jones industrial index have launched nanotechnology initiatives. But, while some of these companies are enthusiastic about the opportunities, they cannot be said to be nanotechnology driven.

Investment company Nanostart AG has identified more than 180 small and medium-size listed companies whose operations are mainly in the nanotechnology field, while Merrill Lynch has created a Nanotech Index (NNZ) to help it keep track of the industry. The index includes firms which state in public documents that nanotechnology initiatives form a significant component of business strategy. Most firms in this category are in the fields of semiconductors, biotechnology, instrumentation, sensors, diagnostics, drug delivery, drug development, genomics, and materials.

The large performance spread of firms on the index underlines the importance of selection and timing.

The influence of nanotechnologies on existing industries

In the medium to long term, new nanotechnology enabled products will strongly influence existing industries. This influence is likely to include the deployment of enhanced products that compete with existing ones as well as, in the case of some firms, disruptive changes. The impact of this influence could well extend beyond the immediate area of application.

If economically viable, the following examples of nanotech applications (which have been mentioned earlier) are likely to exert an influence on established industries:

- Clothing that is able to rid itself of dirt, pollutants and micro-organisms is likely to affect the entire laundry industry.
- Low-cost photovoltaic films and sprays for buildings, cars, bridges etc. will have a possible disruptive influence on silicon photovoltaic cell manufacturers and battery makers among others.

It will be necessary constantly to exercise technical due diligence to isolate the opportunities, risks and possible impact of nanotechnologies. This process also holds true for ”traditional” investments.

7.2. Nanotechnology and industrial insurance: Managing Chances and risks

What are the implications of nanotechnology for commercial and industrial insurance? The commercialisation of products that either contain nanoparticles or use nanotechnologies is an ongoing process. Today, Allianz is insuring many industrial and commercial clients active in this field. The activities of our insureds(policy holders) range across all field of nanotechnologies from chemical companies that produce nanoparticles to manufacturers of consumer products. Over time, the insurance portfolio will contain an increasing number of insureds (policy holders) with an increasing proportion of their commercial activities in nanotechnologies.

By carrying these risks, the insurance industry is already contributing to the early commercial phase of nanotechnologies. Especially for small- and medium-sized enterprises. adequate insurance cover is an important prerequisite for entrepreneurial activity.

This approach is straightforward insofar as the overall risk that is carried is only a very small fraction of the overall risk portfolio. The ”innocent” acceptance of risks is typical of the insurance industry in the early phase of many new technologies. With an increasing proportion of nanotechnology risks in the portfolio, questions about adequate risk assessment procedures, control of accumulation risks and of insurability become more and more relevant over time. A balance needs to be kept between managing a sustainable insurance portfolio with adequate returns and a maintaining a responsible approach towards economic development. This is because the insurance industry, in its risk carrying capacity, is always an enabler of new technologies.

Looking at possible risk scenarios related to nanotechnologies, the following basic features can be distinguished:

- an increasingly high number of persons will be exposed to engineered nanoparticles,
- potential harmful effects will evolve over longer periods of time,
- a high number of companies from various branches of industry could be involved,
- in individual cases it will be difficult to establish a causal relationship between action or omission of a company and the resulting damage, injury or financial loss,
- potential loss scenarios resemble major product liability cases from the past
- occupational exposure is a main concern.

In the context of insurance, the primary area of discussion will therefore be industrial liability insurance with its different types of cover.

There is a class of risks related to new and sometimes revolutionary technologies that is typical of problematic fields in industrial liability insurance. Typically, this class of risk has an inherent mass tort potential which is difficult to specify. The time scale potential here is simply incalculable.

Scenarios like the asbestos claims – especially in the USA – come into mind. The three waves of asbestos claims have cost US insurers and re-insurers approximately $135 billion. The estimates for the fourth wave of claims are as high as additional 200 to $275 billion.

However, to include nanotechnology risks in this context does not seem helpful as the sheer size of the sums tend to mitigate against any sound risk analysis and constructive dialogue between the different stakeholders.
As outlined in the chapter on risks, the – limited – evidence base does not imply that such a direct and broad comparison is valid. The public dialogue and political solutions will gain from differentiation rather than from sweeping arguments. This will be a difficult task for all parties involved – including the media, the industry and the NGOs – but a lot can be gained in terms of economic, social and ecological sustainability along this path.

One of the major legal issues will be causality. Epidemiological studies on ambient fine and ultra fine particles reveal a correlation between ambient air concentration and mortality counts. There does not seem to be a monocausal relationship. Instead, the issue is rather the effects of fine particles on the health of susceptible persons exposed to many other factors such as cigarette smoke, genetic disposition, or lifestyle. It is becoming increasingly clear that long-term illnesses are caused by a complex interplay of different risk factors. The traditional laws and rules that govern liability and compensation are based on a one-to-one assignment of injury and damaging agent. They clearly reach their limits here. This also means limits for traditional liability insurance schemes.

It is likely that any health effects from engineered nanoparticles will be subject to this changed perception. In many cases of illness it might be extremely difficult to clarify whether the exposure to nanoparticles is just one of many contributing factors or a more central factor. In the European legal framework the latter would have to be proven – at least as the situation is today.

How far the burden of proof will be changed remains open. In 1948 the World Health Organisation, in its constitution, defined health as a state of ”complete physical, mental and social well-being,” and not consisting only of the absence of disease or infirmity. Monitoring changes—whether they are scientific, technical, social or legal—is one of the major tasks of risk management. It is a task well suited for insurers.

What risk management measures will be taken by the Allianz Group in relation to new technologies in general and nanotechnologies in particular?

Risk management approach to nanotechnology from an insurer's perspective

Risk awareness

There is a much uncertainty about emerging risks associated with nanotechnologies. It will take years for studies about exposure routes, the effects on human health and the environment to reach conclusive results. While it is still too early to make conclusive statements, our own risk management will need constantly to ”put its feelers out.”

The first step in our risk management toolbox is to create an awareness of the risks and an understanding of the hazards. The first step is to determine how underwriters and risk engineers should deal with critical issues such as direct exposure to nanoparticles or their release into the environment.

Risk identification and risk evaluation

The next step is to identify and evaluate the risks in a continuing process that considers scientific, technical, legal and regulatory factors. Public opinion is also important because that influences many political, industrial, and legal decisions.

Yet close scrutiny of the individual risk, often with support from risk-consulting units, is a prerequisite to taking economically sustainable decisions. Often more important than the underlying basic risk, the risk management practices in these companies will range from very basic to highly sophisticated. The last point is of key importance: even if a certain technology creates new risks, the quality of risk management practices will determine whether these hazards will actually materialise. This is an area where the insurer can contribute through his experience with a multitude of clients from various industries.

Risk handling and accumulation control

Given the knowledge gap about risks outlined above, it is a matter of debate whether one main criterion of insurability – namely the assessability of risks with respect to probability and severity – is or can be fulfilled. The combination of limited evidence about the hazards and potential latency claims again warrants close monitoring of the risk.

On the other hand, it seems neither feasible nor appropriate to start a debate about a general exclusion of nanotechnologies from the insurance coverage today. There are several arguments for this:

- nanotechnologies cover a very broad field with far from uniform risk characteristics,
- the terminology used in nanotechnologies is very broad. No uniform language or set of definitions exist,
- the sheer variety of nanotechnologies and their applications across broad sections of industrial segments mean a positive diversification effect for insurance portfolios,
- the exposure of the general population to nanotechnologies is still comparatively low.

But all this does not rule out specific applications such as the use of nanoparticles in environmental remediation from being subject to more intense risk analysis—and possibly even being excluded from cover.

Furthermore, Allianz has a wide range of measures at hand to optimally manage the risk and ”sculpt” a portfolio. The appetite for certain classes of risk can be defined, and details of coverage can be tailored, both to meet client demands and protect Allianz's assets.

A major challenge with nanotechnologies as well as with other emerging risks will not only be selection and pricing, but also adequate risk capital allocation.

A ”wait-and-see” attitude is not appropriate, especially because of the existence of latent risks, which can take many years to show.

7.3. Conclusions for industrial and commercial insurance

Several steps should get high priority for all stakeholders involved in nanotechnologies in the next few years. These include:

➡ Independent research into the risks of nanoparticles, exposure routes and the effects on humans and the environment. Strengthening the evidence base and allowing public access to the results. Transparency will be a key factor for adequate risk management and public trust,

Developing comparative risk classification schemes and databases, possibly for cross-cutting use by different organisations. Focusing underwriters' and risk engineers' efforts on critical issues such as direct exposure to nanoparticles or their release into the environment.

- Bringing the discussion about nanotechnology to the front line of insurance, that is, to meetings between clients and underwriters and risk engineers,
- Encouraging a willingness to discuss the subject in a way that is not dominated by ideology, and making good use of reviews by independent organisations. Using sustainability as a vision and success criterion.

Editor:
Dr. Christoph Lauterwasser
Allianz Center for Technology

Allianz AG
Group Communication
Königinstr. 28
D-80802 München
Germany
www.allianzgroup.com
Contact:
Michael Anthony
Email: Michael.Anthony@allianz.com

OECD
The OECD International Futures Programme
2, rue André Pascal
F-75775 Paris Cedex 16
France

Contact:
Barrie Stevens
Email: Barrie.Stevens@oecd.org