Pollution of Coastal Seas

M J Varkey

M J Varkey is a scientist at the Physical Oceanography Division, National Institute of Oceanography, Goa.

Pollution of various environments is a consequence of population growth and industrialisation. Coastal seas form part of marine environment and are very rich in minerals, crude oil, fishes etc. They are also being used for disposal of wastes from cities. Various types of wastes, if not properly treated, would cause serious pollution of these shallow seas endangering marine life and spoiling recreational facilities. Different polluting agents like sewage, chemicals, industrial coolants etc. and their effects in the coastal seas are discussed with a historical background.

Introduction

The world's coastal lands are, perhaps, the best places for both habitation and industries. Many of the big metropolises are situated at river mouths e.g. Calcutta, Shanghai, Cairo and Rome. The large rivers with their tributaries provide abundant fresh water supply, cheap transport, fertile land and good sea foods. These are the things ideally required for locating industries also. The mega-cities with their large populations and industries generate millions of tons of domestic and other wastes every year, safe disposal of which is a big problem. During the sixties and seventies many coastal seas were so polluted that some of them became lifeless. Fishes caught from some were poisonous and gases emanating from some others made the surrounding air impure. Now, the situations have improved considerably all over the world. Many dead waters are brought back to life; many beaches are bleached of their stink. But a relapse or failure of the cleaning and monitoring processes might happen, if the administering agencies are not always aware of the seriousness of the problem (see the section on Thames). Almost all marine pollution problems are present in the coastal seas and brackish waters except those caused by deep sea mining and deep sea dumping of wastes. Some aspects of coastal pollution are described in the following sections.

What is Pollution?

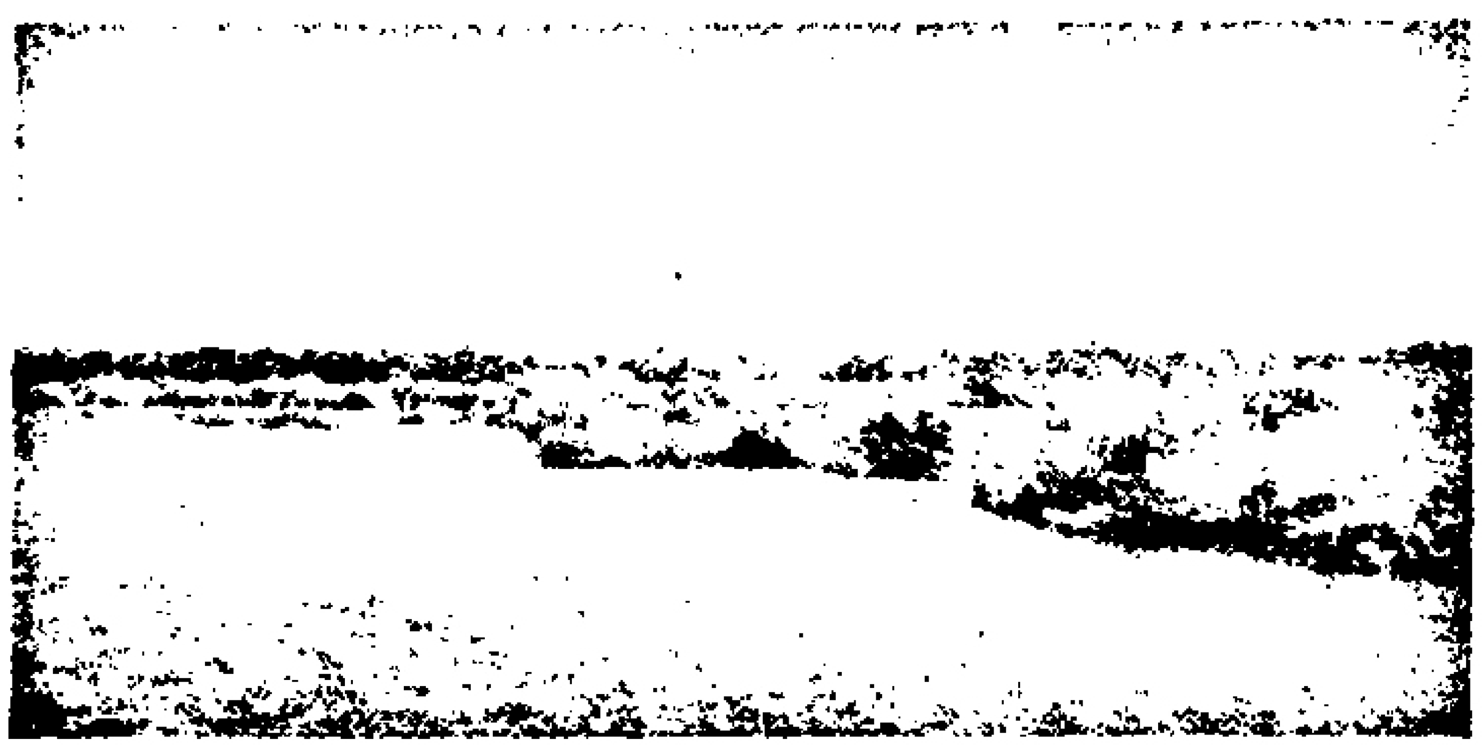
The Inter-Governmental Oceanographic Commission, an international agency (within UNESCO) for ocean research and related matters, defines: 'marine pollution is the introduction by man, directly or indirectly, of substances or energy into the marine environment (including estuaries), resulting in such deleterious effects as: harm to living resources; hazards to human health; hindrance to marine activities including fishing; impairing the quality for use of sea water and reduction of amenities'. Almost all human activities in coastal waters can cause changes or damage, minor or major, short term or long term to the environment, if not properly planned and monitored. Even seemingly harmless (?) activities like constructions, shipping, fishing and boating can dirty or damage the water bodies, especially the shallow waters. (Here it is important to state that while many activities have potentials for serious pollution, pollution can be prevented or mitigated with proper planning, treatment of wastes and monitoring).

Marine pollution is the introduction by man, directly or indirectly, of substances or energy into the marine environment (including estuaries), resulting in such deleterious effects as; harm to living resources; hazards to human health; hindrance to marine activities including fishing; impairing the quality for use of sea water and reduction of amenities.

Now, it seems that the word 'deleterious' in the above old definition if resolved further can lead to a better definition of pollution. If the damage to the water or environment is of minor nature or easily and naturally recoupable, then this can be considered as unavoidable and insignificant. But, if the destruction is of a serious nature, whose recouping might take many years of effort involving research and technology, then this is the situation where the word 'deleterious' rightly fits in. Many sewage discharging problems are results of carelessness and lack of civic values and can be solved by community based efforts.

Coastal Areas and Pollution Sources

India has over 6000 kms of coastline. Its territorial seas alone come to about 0.13 million square kms (with 12 nautical miles as limit). The exclusive economic zone (EEZ) covers an area of



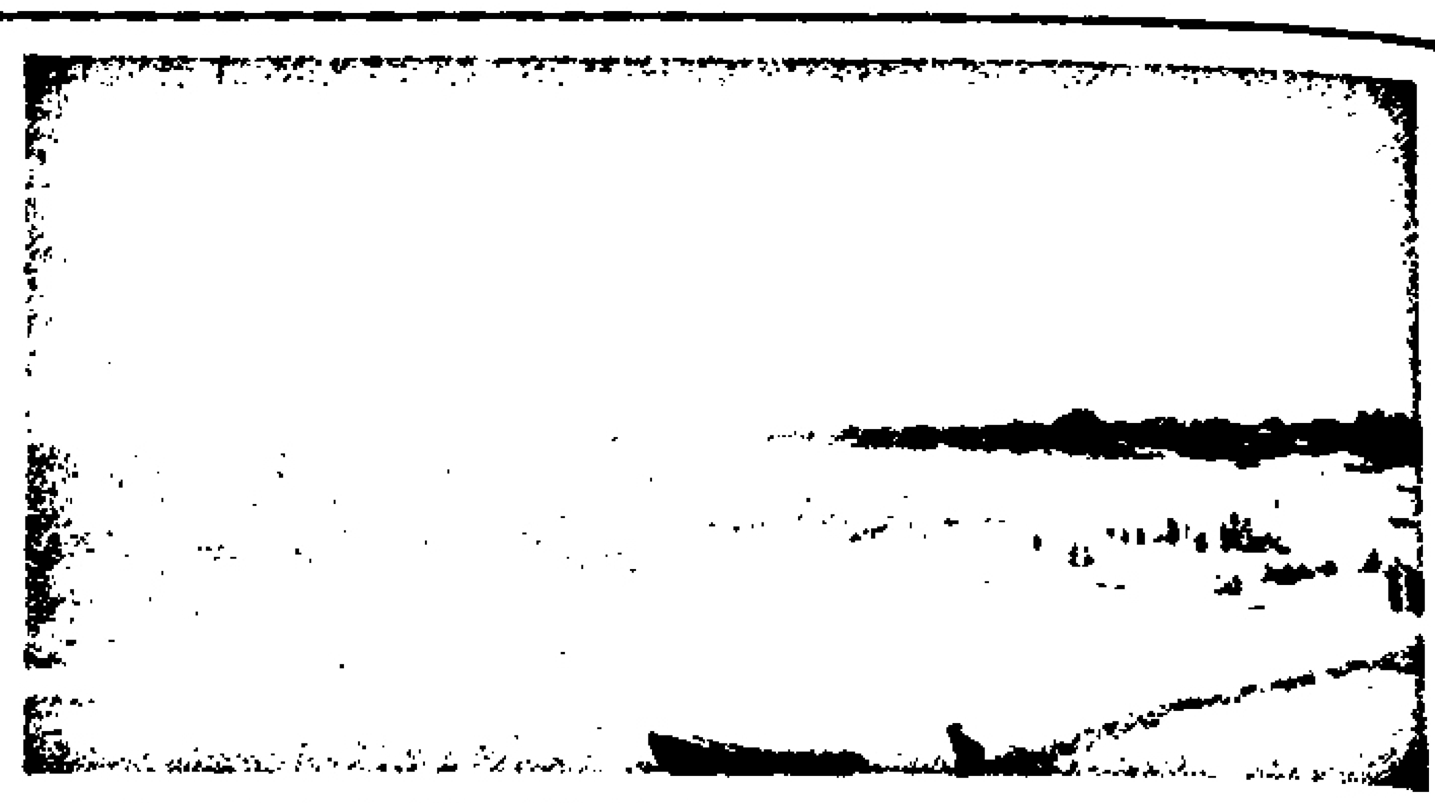


Figure 1 (left). Photograph of a marshy area.
Figure 2 (right). Photograph of a wide beach used for recreation.

about 2.3 million square kms including those around the islands. As far as coastal development and pollution are concerned only the territorial seas and brackish waters are of relevance. This long stretch of coastline consists of various types of lands like marshes (*Figure 1*), mangrove swamps, hilly or raised areas, rocky coasts, sandy coasts, cities and ports. These varied coastal environments are used for ports and shipping, industries, waste dumping, fishing, salt pans, aquaculture, and recreation (*Figure 2*) and water sports. The Central Pollution Control Board (Government of India) has classified the entire coastal areas into five types based on evaluations of 'best-use-of-coastal-segment' from traditional and organised uses and activities. They are: (1) SW1 – salt pans, mariculture, contact water sports and ecologically sensitive areas, (2) SW2 – fishing and noncontact recreation, (3) SW3 – industrial cooling and aesthetics, (4) SW4 – harbours and (5) SW5 – navigation and coastal waste disposal. According to this classification the central west coast of India including northern Karnataka, Goa and southern Maharashtra is grouped under SW1. The main sources of pollution of the Indian coasts are urban sewage, drilling and shipping of crude oil, industrial effluents like chemicals, heat etc., and radioactive wastes,

Sewage and Coastal Dumping

Many of the world's big cities were letting out their domestic waste directly into nearby rivers or seas during the 19th and early 20th centuries. This polluted coastal seas and rivers causing dangerous contagious diseases and eutrophication[1]. City administrators were later required to install treatment plants for

[1] The enrichment of natural waters with inorganic materials, especially nitrogen and phosphorous compounds, that support excessive growth of water plants, choking out other forms of life.

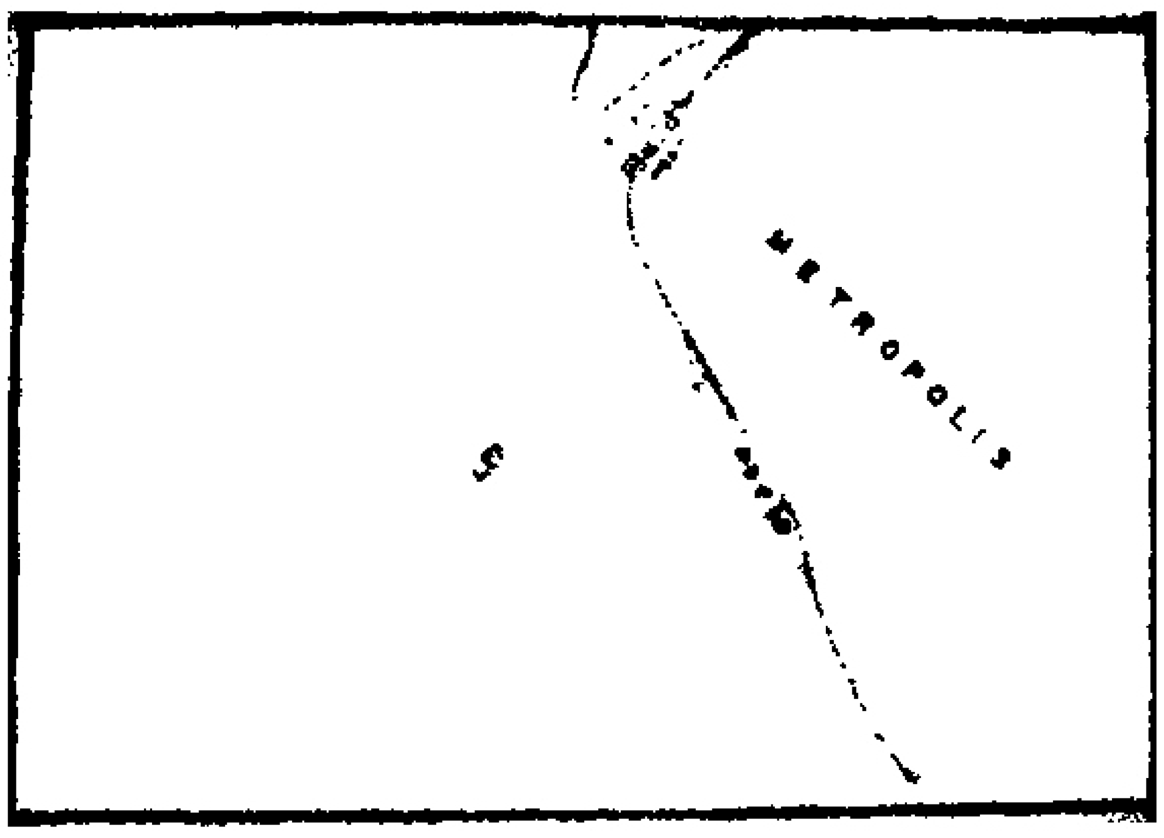


Figure 3. Schematic diagram of oxygen depletion patterns around a sewage sludge dumping site (hatched area) in coastal water.

making the domestic wastes harmless for dispersal in the seas. The sludge coming out of the treatment plants was dumped in safe (well circulated) marine areas. *Figure 3* (Even this sludge is a potential source of pollution if not well dispersed and unmonitored). This improved the sewerage systems and city environments.

But for the financial constraints of the poor nations, sewage disposal with or without treatment, is not a big technological problem today. Sometimes, it would be better to let out the untreated sewage many kilometres away into the open seas where sufficient oxygen supply and mixing exist for degradation and dispersal of the wastes (*Figure 4*). Calculations show that expenses for a large treatment plant are comparable to that of a long pipeline of about 10 kms length for dumping untreated waste in well circulated deep areas of coastal seas.

Box 1.

The ambitious marine outfall projects at Worli and Bandra are on their way to becoming the first of its kind in India to discharge sewage 3kms away from the coastline, (through a tunnel below sea bottom at about 50m. *The Times of India*, May 17, 1997).

Chemical Wastes

A large variety of chemical wastes are formed by various industrial

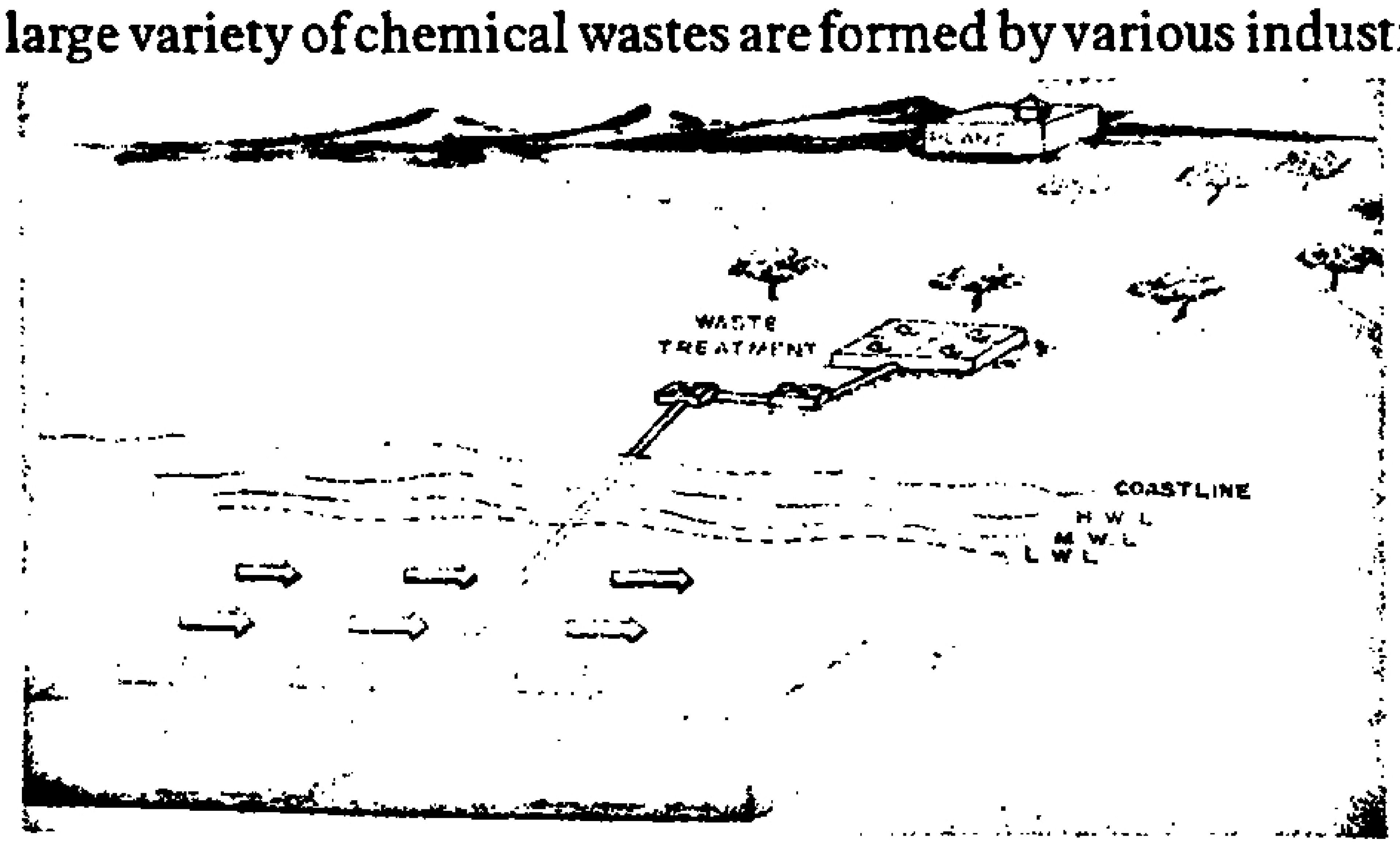


Figure 4. Schematic diagram of dispersal of treated factory waste in deep waters of a coastal sea.

Box 2.

Showa Denko, the company blamed for the mercury poisoning in Niigata prefecture, known as Minimata disease, has 'sincerely apologised' to the victims of the illness and agreed to pay those not already compensated by the government. This came 30 years after the disease was first officially recognised. (news, *Marine Pollution Bulletin.* March 1996).

activities. Many of them are of minor significance as polluting materials. But mercury, cadmium, lead, vanadium, nickel, chromium and pesticides like telodrin, diedrin and endrin are some of the common dangerous waste chemicals. Mercury, cadmium and lead are extremely toxic to organisms. Many of the chemical wastes may be treated before dumping, to remove the poisonous contents and to degrade to easily decomposable or absorbable materials in sea or land. Sewage sludge is found to have higher concentrations (10–200 times) of heavy metals than that of normal Earth's crust material. These concentrations can further increase if industrial effluents form part of sewage.

Many of the dispersed chemicals can accumulate in marine organisms depending on ambient concentrations, tissue peculiarities etc. and their concentrations would go on increasing in other organisms who feed on them. Humans and other organisms when fed on food items contaminated with pollutants are affected with various types of physiological problems or diseases. The Minimata disaster in Japan is a glaring example in which dozens died and hundreds became sick due to mercury poisoning. DDT poisoning was found to kill thousands of pairs of breeder birds (terns) off the coasts of the Netherlands in 1965. A drastic reduction in number of healthy pelicans was noted in California where DDT pollution existed. Studies revealed heavy contamination of pelican eggs by DDT. Efforts to reduce the DDT in the environment resulted in lowering the concentration levels of DDT in pelican eggs from 907 to 97 (mg/kg of fat) during 1969–1974. Effects of chemicals on life are of a very different type compared to eutrophication effects.

Box 3.

A Russian Milke class nuclear powered submarine sank in the Norwegian Sea on 7 April after an explosion. Norwegian and Russian scientists have been on the site since the accident, analysing air, surface and deep water samples for radioactive contamination. (News, *Mar. Poll. Bull.*, June, 1989).

Thermal and Radioactive Wastes

Thermal pollution of rivers and coastal seas by heat released from coolants from the factories is a serious environmental problem causing destruction and imbalance of aquatic life. Many aquatic lives are very temperature sensitive. The warm waste water released into coastal seas raises the ambient temperature causing physiological and other problems to aquatic organisms. Standard guidelines exist for levels of permitted *warmings* to avoid harm to aquatic life. In Europe (temperate regions) 2° C is the upper limit of differential warming with a maximum of 26°C. In tropical regions 2°C rise can be harmful. In all thermal pollution problems some basic points are to be properly weighed before taking remedial measures. These are (1) warming can favour some life forms at the disadvantage of others; (2) indirect effects of favoured growth can be bad to aquatic ecology and human food items and (3) if the whole range of aquatic life (other than favoured) is found not affected over long monitoring periods warming consequences can be considered nondeleterious and dealt with calmly.

Radioactive waste water released from nuclear plants is of a very different nature due to their long half life periods and damaging effects on cells compared to waste heat. In the oceans the amounts of natural and induced radioactivities are low. The main sources of induced radioactivity are (1) nuclear weapons testing, (2) nuclear reactors and processing plants, and (3) nuclear powered ships and submarines. One important point to be noted is that the half life periods of almost all radiation products (induced) are very low (< 30 years) compared to periods of natural radioactivity (of the order of 1000 yrs). Radioactive wastes affect surface planktonic organisms and fishes living near the bottom to various degrees. Even during the sixties and seventies, the radioactive wastes in the seas were at tolerant levels. Now, after instituting better 'test bans', decommissioning bad reactors and with international inspection mechanisms, radiation pollution is of no alarming nature. This is not a plea for complacency. Presently, deep waters of international seas and national territories of safe geological formations are used to deposit nuclear wastes safely for decades.

Petroleum Products and Shipping

Large cities flush out lakhs of tons of waste hydrocarbons into their sewage. Drilling, pumping and marine transport of crude oil are major sources of oil spills in the seas (about 2 million

Box 4.
Another lawsuit has added to the roster of those seeking compensation from Exxon Corporation .... the suit seeks restitution for the thousands of birds, otters and seals killed, a claim that could cost the company billions of dollars .... The suit is just one of almost 200 pending against the company. (*Mar. Poll. Bull.*, 1989, 20 (10): 485.)

tons). Polluted rivers and port operations and clean-ups in harbours add thousands of tons of oil wastes into coastal seas. Oil pollution caused by marine transport (largest from the Arabian oil fields) can reach the coasts of Africa, Asia and the Americas.

Effects of crude oil and its remnants in the sea are of many types. Oil on sea surface is subject to evaporation, dissolution, emulsification and biodegradation. After these processes, the residue would have solid shapes like peas and are called tar clumps. They spread on the water column and may sink as sediments. Sometimes they drift to the beaches. Large scale oil spills from accident sites of ships and harbours had caused widespread damage to marine organisms like planktons, fishes and birds. The cases of Torrey Canyon (1967), Amoco Cadiz (1978), Exxon Valdez (1989) and the biggest Gulf War oil slick (1991) are well-known examples.

The effects of landing of oil on different types of coastlines are different. Wave activity is found to have considerable effect on the dissipation of oil on coasts. On rocky coasts with high wave activity, oil clean-up is not generally necessary. But salt marshes, mangrove areas and sheltered (from waves) tidal flats would have serious long-term after-effects. As regards to oil pollution on coasts, the following conclusion of the Warren Spring Laboratory of UK is noteworthy: *"from the point of view of pollution of the environment the best thing to do with oil pollution of the beach is to do nothing. With the climatic conditions and the types of sea around the British Islands, oil left alone will fairly rapidly become innocuous or disappear altogether."* Even though this is specific to British coasts, the conclusion is applicable for other coasts also depending on their nature and wave activity.

Box 5.
'*Saving Oiled Seabirds*' a booklet prepared by the International Bird Reserve Research Centre, Berkeley California and distributed by the American Petroleum Institute, summarises the best established techniques for rehabilitating oiled birds (News, *Mar. Poll. Bull.*, May, 1978).

Story of Thames

The brackish waters of river Thames provide a glaring, long story of pollution, cleaning and restoration, during two centuries (see for details the book *The Restoration of the Tidal Thames* by

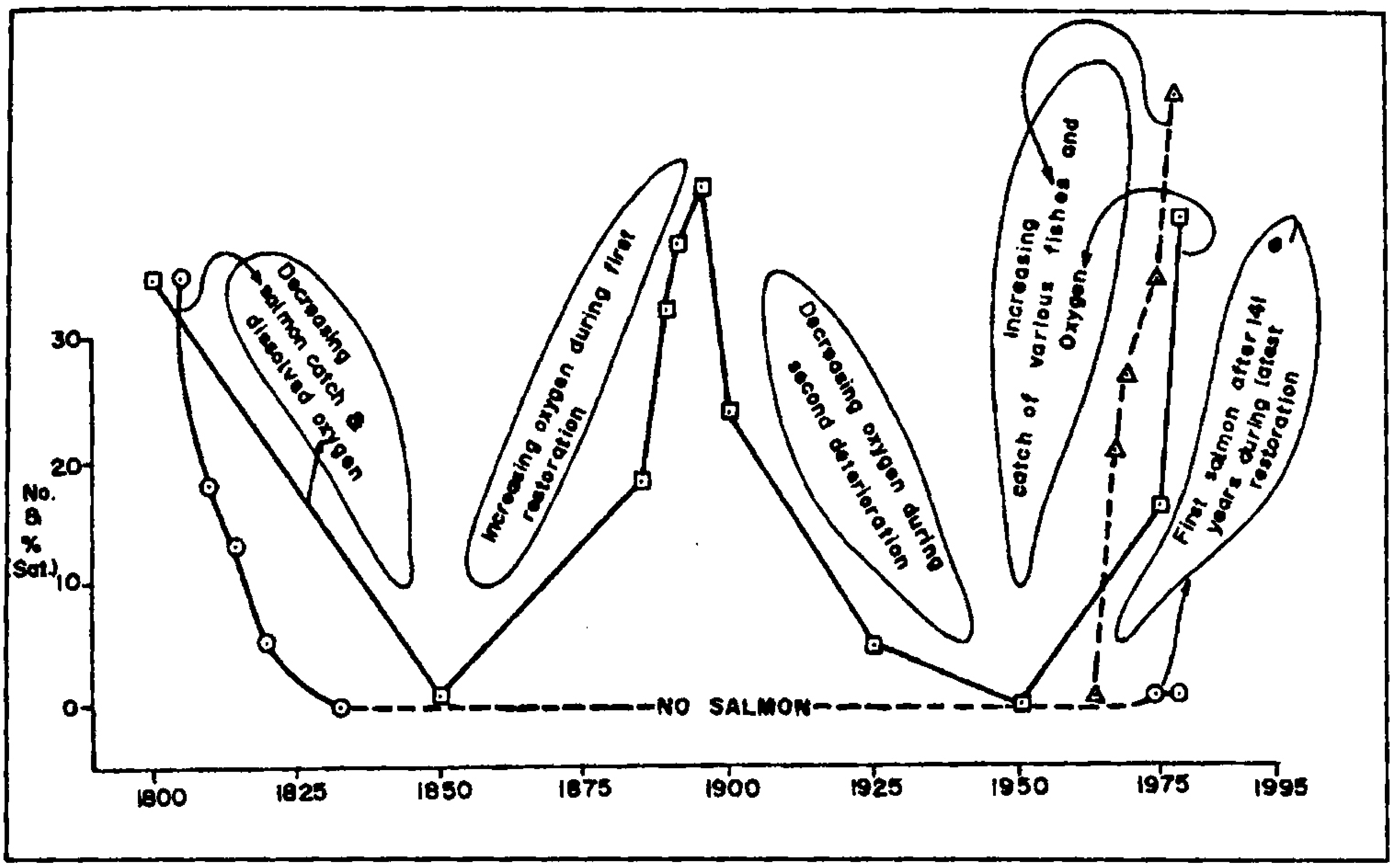


Figure 5. A composite of oxygen saturation levels and fish catch in tidal Thames for about two centuries.

L B Wood 1982). This book gives data and information on many pollutants, treatment methods, oxygen depletion, fish catch etc. during different periods at different locations along the tidal Thames from 1800 to 1980. The adjoining figure on oxygen (percentage of saturation) and fish catch is composited from these data (*Figure 5*). The tidal Thames passed through two cycles of 'pollution-restoration' during 180 years each with a period of about 90 years. It would be very interesting to search out the reasons for the second phase of deterioration during 1895–1950, even though the city experienced one cycle of 'deterioration-restoration' earlier. Was it due to lack of civic values or callousness of ordinary citizens or administrators? Is it likely that such cycles can happen in other large cities also? What were the exact socio-economic reasons? The answers would be definitely useful for other cities and city administrators.

A Message

Many case histories of pollutions of various types in various

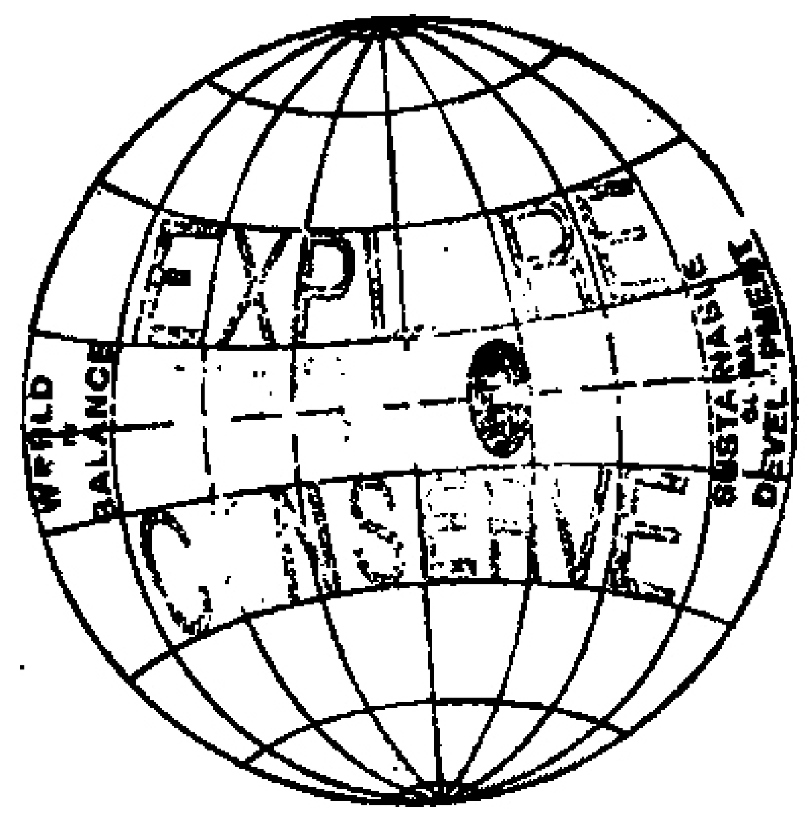


Figure 6. A visualisation of global human endeavours and sustainable development.

places make it clear that pollution is a result of ignorance or lack of values in ordinary people, local administrators and factory managers. Problems of ozone holes and greenhouse gases were, perhaps, beyond anybody's forethinking and are of global nature. These required long term international programmes for remedies and are being carried out. As far as the future of our Earth is concerned 'sustainable development' of all seems to be the only solution (*Figure 6*).

The search (explore and exploit) for new resources and techniques have to continue. But in the case of planned, new industrial processes, projects and sewage systems the following points (for conservation) are of high importance:

(1) When a factory is planned with new manufacturing processes or for new products the pathways of degradation and end-products of wastes and products are also to be studied and necessary treatment methods are to be implemented.

(2) Efficient and harmless (over long periods) waste disposal systems are to be installed.

(3) Continuous monitoring of the environment for any accumulated after-effects caused by various uses of products and waste disposal (in sea or land) are to be carried out. This can prevent unpredictable consequences (e.g. ozone hole).

Address for Correspondence
M J Varkey
Scientist E-2
Physical Oceanography Division
National Institute of Oceanography, Dona Paula
Goa 403 004, India

Suggested Reading

[1] S A Gerlach. *Marine Pollution*. Springer-Verlag. Berlin, 1981.

[2] L B Wood. *The Restoration of the Tidal Thames*. Adam Hilger Ltd. Bristol. United Kingdom, 1982.

Systems and Control Engineering

1. Notions of Control

A Rama Kalyan and J R Vengateswaran

To control an object means to influence its behaviour so as to achieve a desired goal. To implement this influence, engineers build various devices that incorporate several mathematical techniques. The study of these devices and their interaction with the object being controlled is the subject of this article.

Rama Kalyan is currently with the Department of Instrumentation and Control Engineering, Regional Engineering College, Tiruchirapalli. He is deeply interested in the concepts of control wherever they exist, be they physical, biological, or abstract systems. The mathematics behind control theory always fascinates him. His other interests are Carnatic music and philosophy.

J R Vengateswaran is presently a third year student of the same department. He would like to work in control engineering and its applications in robotics and bio-medical engineering.

Introduction

Let's begin with a simple question. When did you last use the word *control*? Perhaps, one may have to think for a while. But the paradox is that one invariably uses this almost in every walk of life but fails to take notice.

- He has no *control* over his expenditure
- The driver lost *control* and collided head on
- The law and order situation in the city is out of *control*
- Pest *control* in orchards

These are a few common phrases we come across all the time. There are many new products and services being introduced everyday that depend on control systems. Yet, they are never identified as control systems. The user of the system does not focus on the control system, but rather on the desired result. Thus, control systems are invisible.

Should we lament this apparent invisibility of control systems? Perhaps, we do as control engineers. This invisibility, however reflects the fact that control systems is a mature technology. It is the very nature of control systems, that the performance of an object or process which is controlled is improved.

The label of mature technology both exalts and haunts control

Controls discipline has played critical roles in technological advances over the last few decades.

science and engineering. There is no question that the controls discipline has played critical roles in technological advances over the last few decades. If not for modern control theory, aircraft would not be flying as far or as fast; manufactured goods, from paper to steel to gasoline to Mars Bars, would not be as readily available; Rover would still be just a dog's name; ... the list is endless. Control has been a linchpin of the modern age.

We avoid a formal definition of the word *system*. Instead, we shall rely on the popular or intuitive concept of a system that one has. That we are able to do so illustrates one of the facets of system theory which makes it so important, namely, we can find the ideas and techniques of system theory in such seemingly unrelated areas as computer design, investigation of the central nervous system, guided missiles, etc. Whenever a system is designed such that the dynamic behaviour of an initial process is intentionally altered to achieve some desired performance, then we have a control system.

An Illustration

Let us imagine we are driving an automobile. We start from a point A and our destination is B, say, 60 km away. Obviously, depending upon the time constraint, we choose to drive at a pre-calculated speed. Let us say, we need to reach our destination exactly in 60 minutes. Hence, the pre-calculated speed is 60 kmph. We would like to express this in the following sentence. 'If we drive the automobile at 60 kmph, then we reach our destination in exactly an hour'. The antecedent clause in this sentence refers to the *cause* and the consequent refers to the *effect*. This may be depicted as shown in *Figure 1*.

The cause may be physical, for example, the pressure we apply on the accelerator pedal so that the speed of the vehicle is

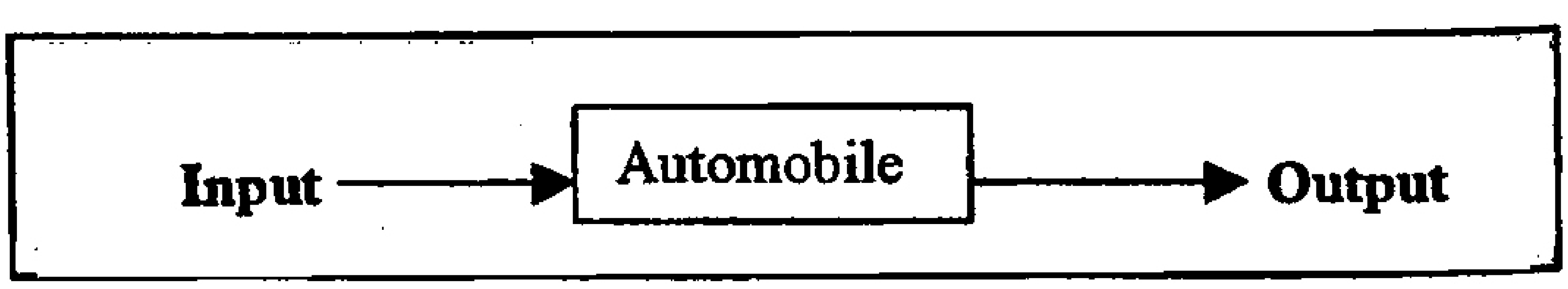


Figure 1.

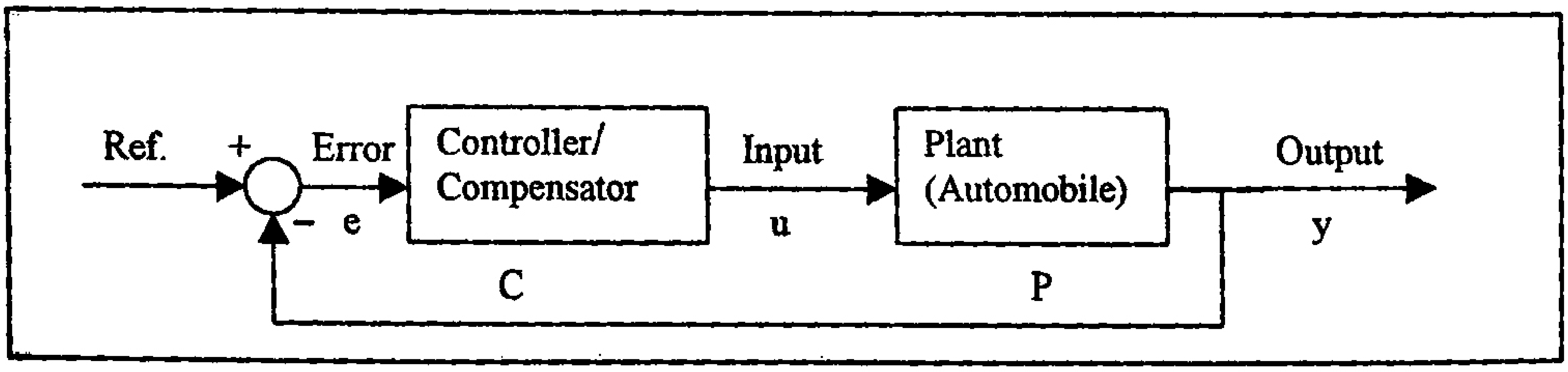
constant at 60 kmph. We also use the word *input* for convenience. Similarly, the effect may be the rate of displacement or the speed of the vehicle; or equivalently, the distance covered by the vehicle. Sometimes, we use the word *output* for the effect. Thus, we assume that if an input is applied there will be a unique response. This relationship between the input and the output is essential in defining a system.

Whenever a system is designed such that the dynamic behaviour of an initial process is intentionally altered to achieve some desired performance, then we have a control system.

Evidently, so far, we have considered an ideal case where it is assumed that there is absolutely no traffic, no speed-breakers etc. Let us now include some of these practical factors. Suppose there is some obstacle, say a railway crossing, after travelling 10 km, when we need to apply the brakes (another control!) and wait for 10 minutes. We have to increase the speed to 75 kmph so that the time lost may be 'compensated', and we reach the destination on time. What exactly is the process that we accomplished just now? We *measured* the variables – the distance and the time – and accordingly changed or modified the speed. In other words, we controlled the automobile by using the difference between the *desired* distance to be covered and the *actual* distance covered so far. This may be depicted as shown in *Figure 2.*

For obvious reasons, we call the control system in *Figure 1* an *open-loop* system and that in *Figure 2* a *closed-loop* system. In this illustration, the system we attempted to control is an automobile. In general, any system that is required to be controlled is called a *plant.* The input applied to the plant is called the *control signal* or *control law* which is provided by a compensator. In the closed-loop control system, the plant output is compared with a reference input and an error is computed. Accordingly, we also call this

Figure 2.



system, a *negative unity feedback* control configuration. Other configurations also exist but are beyond the scope of the present article. This negative unity feedback configuration is the most used one as it is conceptually quite simple.

Success Stories of Control Systems

We often hear about new control applications in flight control, process control, manufacturing, environmental control, and many other areas. In *Box 1* and *Box 2*, we describe the application of feedback control concepts to two novel domains.

Box 1. Application of Control Systems to Multi-Storeyed Buildings

Since the collapse of the Tacoma Narrows bridge in 1940, the relevance of control engineering to the design of civil engineering structures has been noted. Protecting civil structures is of great economic and social importance. Buildings and other physical structures have traditionally relied on their strength and ability to dissipate energy to survive under severe dynamic loading. In recent years, however, world wide attention has been directed towards the use of control and automation to mitigate the effects of those dynamic loads on the structures. The basic task is to determine a control strategy that uses the measured structural responses to calculate an appropriate control signal that will enhance structural safety and serviceability. The first full-scale application was accomplished by the Kajima Corporation, Japan in 1989. The Kyobashi Seiwa building is a eleven story (33.1 m) building in Tokyo, having a total floor area of 423 sq.m. A control system was installed, consisting of two active mass damper systems (AMDs) – the primary AMD is used for transverse motion and has a mass of 4 tons, while the secondary AMD has a mass of one ton and is employed to reduce torsional motion. The role of the active (or feedback) system is to reduce building vibration under strong winds and moderate earthquake excitations and consequently to increase the comfort of occupants of the building. To date, there have been more than 20 buildings and 10 bridges that have employed feedback control strategies in full-scale implementation. A few of them are listed below:

Full Scale Structure	Location	Year	Scale of building
Sendagaya INTES	Tokyo, Japan	1992	58m, 3280 ton, 11 stories
Osaka Resort City 2000	Osaka, Japan	1992	200m, 56980 ton, 50 stories
Yokohama Landmark Tower	Yokohama Kanagawa, Japan	1993	296m, 260610 ton, 70 stories
Hiroshima Richga Royal Hotel	Hiroshima, Japan	1994	150m, 83000 ton, 35 stories
TC Tower	Kao Hsung, Taiwan	1996	85 stories
Nanjing Tower	Nanjing, China	1998	310 m

Box 2. Controlling Telescope.

The Keck astronomical telescope at Mauna Kea in Hawaii uses control innovatively. The basic objective of the telescope is to collect and focus starlight using a large concave mirror. The shape of the mirror determines the quality of the observed image. More light can be collected with a large mirror, and hence dimmer stars can be observed. The diameter of the mirror on the telescope is 10 m. To make such a large high precision mirror out of a single piece of glass would be difficult and costly. Instead, it uses a mosaic of 36 small hexagonal mirrors. These 36 segments must then be aligned so that the composite mirror has the desired shape. The control system to do this may also be depicted as a closed loop system similar to the one in *Figure 2*. In controlling the mirror's shape, it suffices to control the misalignment between adjacent mirror segments. For the mirror to have an ideal shape, the displacements should have certain ideal values that can be pre-computed; these are the components of the reference **R**. Behind each segment are three piston type actuators, applying forces at three points on the segment which adjust its orientation. The compensator **C** must be designed so that in the closed loop system, y is held close to **R** despite disturbing forces such as wind gusts, changes in ambient temperature, etc..

Brief History of Control Theory

When did mankind first attempt to intentionally alter the environment? Obviously, we cannot identify this specific event, but we know it has always been a human characteristic to seek to control objects with which we interact. Perhaps, the first control system was that of the Greeks and Arabs who invented water regulation devices so that they could more accurately measure time; but if we use a general definition of control systems, we suspect there are even earlier examples. Making fire is a way of controlling temperature! In what follows, we present a very brief history of automatic control. Technical details and the names of engineers/scientists are not given to sustain the flow of the narration. Motivated readers may find more details in [1]. The history of automatic control can be divided into four main periods as follows:

Early Control : Up to 1900

Most inventions and applications of this period were concerned with the basic activities of controlling temperatures, pressures, liquid levels, and the speed of rotating machinery. However, growth in the size of ships and naval guns, and introduction of new weapons such as torpedoes, resulted in the application of steam, hydraulic and pneumatic power systems to operate position control mechanisms. Further applications of control systems became apparent with the growth of knowledge of electricity and applications. For e.g. arc lamps require the gap between the electrodes to be kept constant, and generally it is necessary to keep the voltage of the electricity supplied to users constant.

The Pre-Classical Period : 1900 – 1935

The early years of the 20th century saw the rapid and widespread applications of feedback controllers for voltage, current, and frequency regulation, boiler control for steam generation, electric motor speed control, ship and aircraft steering and auto stabilization, and temperature, pressure, and flow control in the process industries. As applications multiplied, engineers became puzzled and confused; controllers that worked for one application, or for one set of conditions, were unsatisfactory when applied to different systems or different conditions; problems arose when a change in one part of the system (process, controller, measuring system, or actuator) resulted in a change in the major time constant of that part. (For a definition of the time constant please refer to the next part of this article).

The differential analyser provided a means of stimulating the behaviour of dynamical systems.

During the same period, extensive work was being done on mechanical analog computers at the Massachusetts Institute of Technology. This work resulted in the differential analyser, which provided a means of simulating the behaviour of dynamical systems and of obtaining numerical solutions to differential equations.

The Classical Period : 1935 – 1955

During the first five years of this period, advances in the understanding of control systems were made independently by three prominent groups working in the United States. A group at Bell Labs tried to find ways of extending the bandwidth of their communication systems. The second important group was mechanical engineers and physicists working in the process industries. They sought to establish a common terminology and tried to develop design methods. The third group was at the

Electrical Engineering Department at MIT. They used time-domain methods based on operator techniques and began to develop the use of block diagrams, and used the differential analyser to simulate control systems.

The advent of World War II steered control systems work on a few specific problems. The most important of these was the aiming of anti-aircraft guns. This is a complex problem that involves the detection of the position of the air-plane, calculating its future position, and the precise control of the movement of a heavy gun. Work on the 'systems' problem brought together mechanical, electrical and electronic engineers, and an outcome of this cross-fertilization of ideas was a recognition that neither the frequency response approach used by the communication engineers nor the time response approach favoured by the mechanical engineers were separately effective in designing position control systems. This led to the use of Laplace transform techniques. By the end of the war, researchers were concentrating on the nonlinear and sampled-data systems. The other major development to emerge from the fire control work during the war was the study of stochastic systems. During the same period, the teaching of control theory spread, initially through special courses run for practising engineers and graduate students and then absorption within the standard syllabus of many engineering courses.

The advent of World War II steered control systems work on a few specific problems. The most important of these was the aiming of anti-aircraft guns.

The Modern Control Period: Post 1955

Although the direction of some post-war work was influenced by the insights and new understandings developed during the war, the trajectory of development was largely determined by two factors : First, the problems that governments saw as important – the launching, manoeuvring, guidance and tracking of missiles and space vehicles; and second, by the advent of the digital computer. The first problem was essentially control of ballistic objects, and hence detailed physical models could be constructed in terms of differential equations, both linear and nonlinear. Further, measuring instruments and other components of great accuracy and precision could be developed and used. Engineers working in the aerospace industries, following the example set by Poincaré, turned to formulating the general differential equations in terms of a set of first-order equations, and thus began the approach now known as the *state-space* approach.

In the later part of the 1950s, Bellman began working on optimal control theory, at first using the calculus of variations but later, seeking to formulate deterministic optimization problems in a way in which they could be solved by using dynamic programming. The generalization of Hamilton's approach to geometric optics by Pontrjagin in 1956 in the form of his maximum principle, laid the foundations of optimal control theory. This and Bellman's insight into the value and usefulness of the concept of state for the formulation and solution of many control and decision problems led to extensive and deep problems of automatic control. The growing availability of the digital computer during the late 1950s made a recursive algorithmic solution possible.

By the early 1960s, the digital computer had been used on-line to collect data for optimization and supervisory control and for a limited number of applications of direct digital control. A leading advocate for the use of digital computer in the process industries was Donald P Eckman. He persuaded several companies to support research based at the Case Institute of Technology, Cleveland, Ohio. The programme was initially called *Process Automation* but it was renamed later as *Control of Complex Systems*. This was because Eckman wished to distinguish what he was doing from the popular image of automation. By the end of the decade Eckman was arguing in support of *Systems Engineering* with the idea that what industry needed was engineers with 'a broad background across conventional boundaries of the physical, engineering and mathematical sciences' and with 'an ability to approach problems analytically, to reduce physical systems to an appropriate mathematical model to which all the power of mathematical manipulation, extrapolation, and interpretation can be applied'.

Suggested Reading

[1] *IEEE Control Systems Magazine*. Special issue on the evolving history of control. Vol.16. No.3, 1996.

[2] *IEEE Control Systems Magazine*. Special issue on breaking through with emerging technologies. vol. 17. No. 6, 1997.

[3] Maciariello and Kirby, *Management Control Systems*. 2/e. Prentice Hall of India, 1998.

[4] M Gopal. *Control Systems: Principles and Design*. Tata McGraw Hill, 1997. (An exhaustive list of classified references may be found in this book).

[5] P R Be Langer. *Control Engineering*. Saunders College Publishing, 1995. (This book was reviewed in *The Hindu* recently).

Address for Correspondence
A Rama Kalyan and
J R Vengateswaran
Department of Instrumentation and Control Engineering
Regional Engineering College, Tiruchirapalli 620 015, India.
Email: vkalyn@rect.ernet.in

Self Assembled Monolayers

T Pradeep

In recent years, a number of systems have been shown to spontaneously assemble on appropriate solid surfaces to form films with thicknesses of molecular dimensions. The special techniques needed to prepare, characterise and study such monolayers are described. The potential applications of these systems as catalysts, sensors, electro-optic and other materials are summarized.

T Pradeep is at the Department of Chemistry and Regional Sophisticated Instrumentation Centre, IIT, Chennai.

Introduction

Surface and interfacial phenomena have attracted scientists of diverse disciplines. Of the variety of processes occurring at surfaces, adsorption and concomitant catalytic transformations have been widely explored and constitute some of the most intensely pursued areas of contemporary research. Perhaps, the most important step in these chemical processes is adsorption of a molecular layer of the incoming reactant at the surface. Although the concept of monomolecular films is nearly two centuries old, the deliberate organisation of molecules on surfaces to form layers of molecular dimensions happened only in this century. Today, scientists are devising methodologies to assemble molecules of their choice on different surfaces for certain specific properties. This area is especially active, because of the funda-mental issues as well as due to the potential applications. This article provides a bird's eye view of the fascinating area of monomolecular films.

History

It has been known from the time of sea voyages that when oil is spilled over water, the film formed dampens surface waves and ripples. The holiness of water has been attributed to this phenomena rather than to the oiliness of oil! Aristotle explained

Direct measurements of molecular dimensions became possible with the methodology introduced by Pockels, who worked mostly on her kitchen table.

it as due to the slipping of waves on an oil covered water surface, so that they do not make an impression. The first scientific studies of oil films on water were conducted by Benjamin Franklin (1706–1790), the American statesman scientist. He seems to have got interested in oil films during one of his frequent travels to Europe. An account of Franklin on oil films was published in *Philosophical Transactions of the Royal Society* in 1774, although his original interest in the subject dates back to 1757 during his travel to Europe. An interesting aspect of the account is that on careful reading, it implies the molecular nature of matter. Several scientists took up investigation of the subject in the 19th century, but the breakthrough occurred in 1890, with the publication of a series of papers by Rayleigh (1842–1919) on water surfaces. In one paper he discussed the effect of olive oil films on the surface tension of water and concluded that the films of olive oil are 10–20 Å in thickness. Rayleigh appeared to have the belief that oil spreads on the water surface to such an extent that the film thickness approaches molecular dimensions. Direct measurements of molecular sizes became possible with the methodology introduced by Pockels (1862–1935) and her 1891 letter to Rayleigh which was forwarded to the leading journal *Nature* for publication, describes the basic apparatus which constitutes the bare essentials of monolayer research even today. Pockels worked for about 18 years on the problem and mostly on the kitchen table, without any professional training. As she remarks to Rayleigh, "with regard to your curiosity about my personal status, I am indeed a lady"! Then came Langmuir (1881–1957), who developed theories of adsorption although his contributions spread into a number of areas. Langmuir developed a number of new techniques for studying films, the famous one being the film balance which bears his name. Blodgett (1898–1979) transferred the floating films on liquid surfaces (the Langmuir films) to solid surfaces (the Langmuir–Blodgett or L–B films). These films have interesting properties and serve as model systems to study a number of fundamental properties.

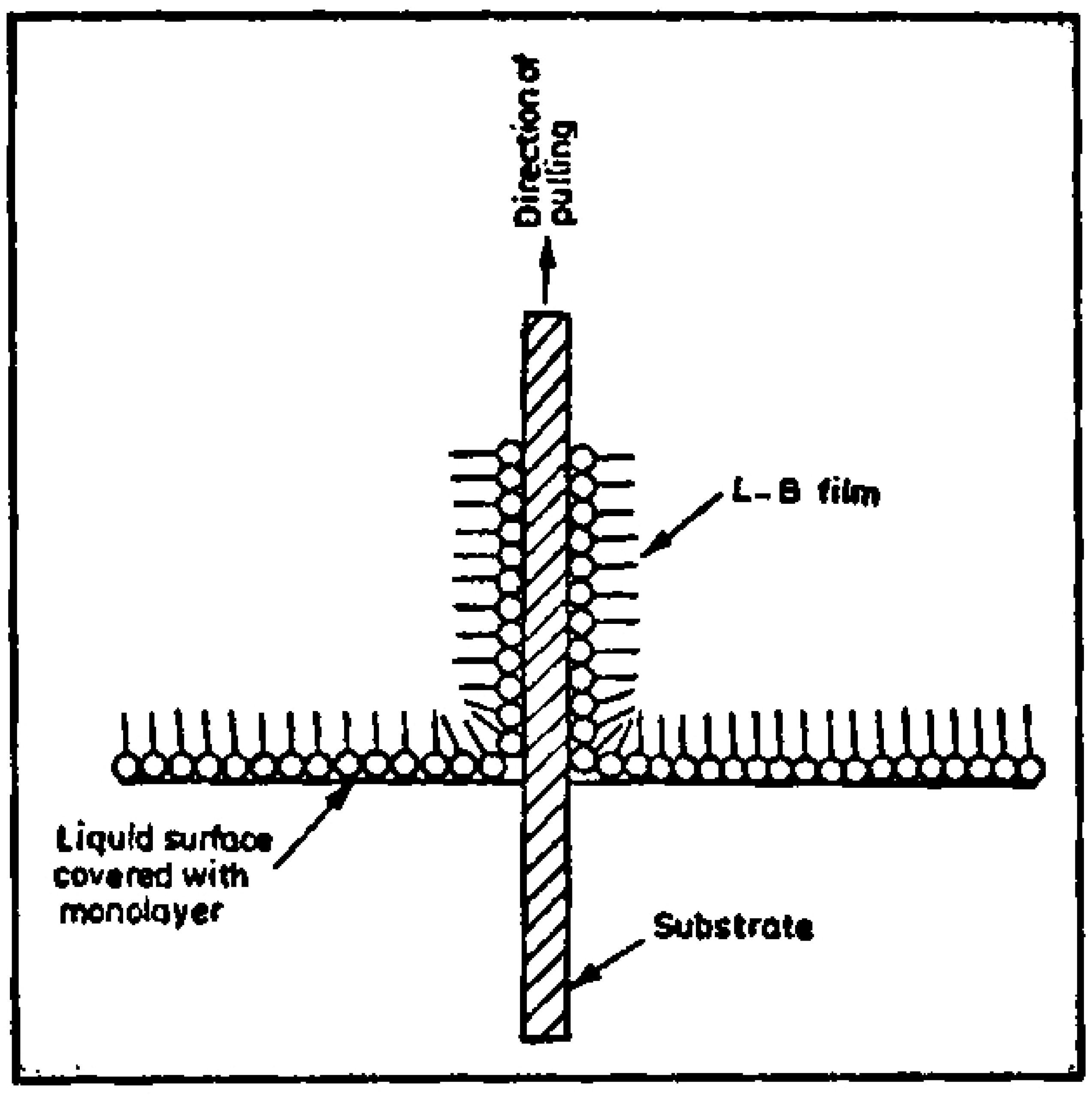


Figure 1. Schematic diagram of the preparation of a Langmuir–Blodgett film. When the substrate is removed from a liquid whose surface is contaminated with a layer of surfactant, a monolayer clings onto the surface. Upon application of a constant pressure on the film floating on the liquid surface and maintaining the rate of pulling a constant, it is possible to transfer a well-ordered film on to the surface of the substrate.

L–B Films

A monomolecular film assembled by the L–B methodology is easily visualised. If one lifts a glass plate through a barium stearate [$C_{17}H_{35}COO-$] monolayer spread on water, the film which holds onto the plate will have the hydrocarbon tail pointing outward as shown in *Figure 1*. The surface after film preparation will be hydrophobic (water repelling). The film covered surface can be dipped back in the liquid depositing another layer of the film back-to-back and such a film will be hydrophilic (water liking). This kind of overlayer growth is called Y type growth. There are also other kinds of growth such as X and Z, in which similar surfaces are exposed all the time. Several tens of layers can be grown this way. The application of such films range from fundamental science to technology.

Self Assembled Monolayers (SAMs)

Self assembled monolayers are recent additions to the family of molecular films. These films are different from L–B films because they are self assembled to form an ordered molecular film, unlike L–B where they are transferred from the air-liquid interface to the surface. The SA monolayers are thus defined as molecular assemblies which form spontaneously by the immersion of a surface into a solution of surfactant. Thus depending on the surfactant and the substrate, monolayers vary.

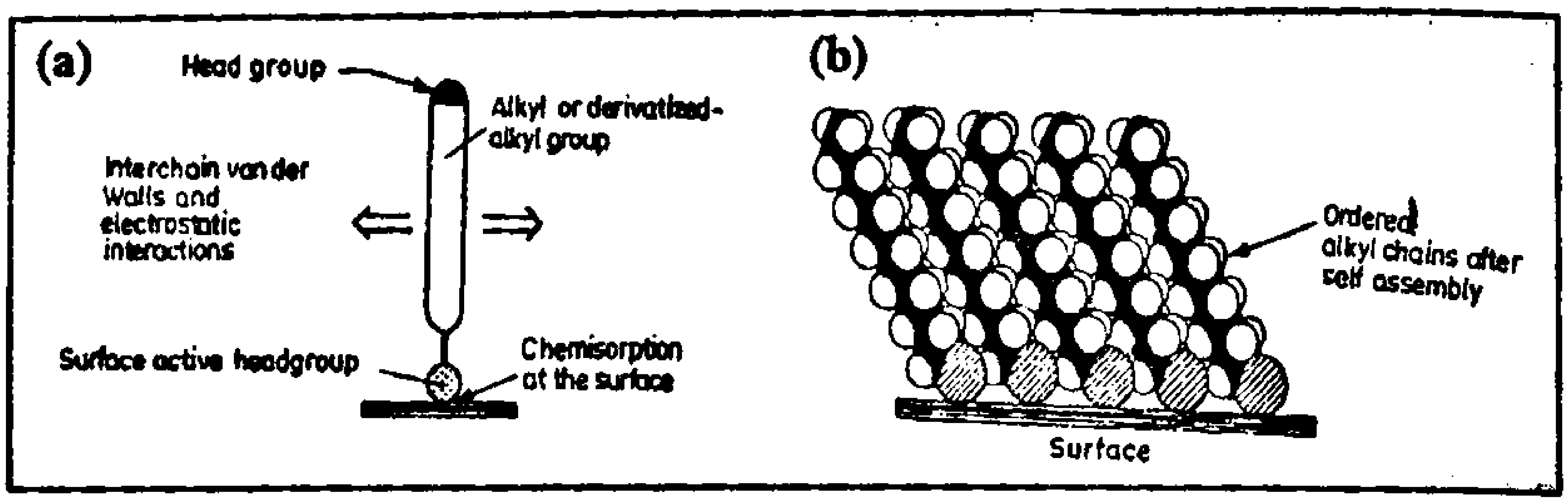


Figure 2. Schematic events during self assembled monolayer formation. (a) The surfactant having a head group which preferentially chemisorbs at the surface. Since the head group-surface interaction is strong, all the possible surface sites are occupied. The initial adsorption occurs quickly, within minutes. (b) Such a surface binding brings molecules closer so that inter-chain van der Waals interactions become important and the molecules stand up. This is a slow step in the monolayer formation, taking several hours.

The most common monolayers are formed with organosilicon derivatives, alkane thiols, dialkylsulphides, alcohols, amines and carboxylic acids on different surfaces. All molecules will not self organise on all substrates. The affinity between the molecule and the surface is an important factor. From a number of investigations, it is now clear that the first event in self organisation is the chemical bonding of the surface active group (the head group) to a surface site. It so happens that since the chemical formation reaction is highly exothermic, all the available surface sites are occupied. Since the kind of binding brings the molecules close to each other, the short range van der Waals forces become important. These interactions make the molecular chains attached to the head group stand up vertically, although with a tilt. An assembly of these molecules can extend over several hundreds of angstroms and an ordered oriented monomolecular layer results. The organised assembly can be schematically represented as in *Figure 2* with the corresponding events.

When one says *order* in an organised assembly, normally it implies only translational symmetry in a two dimensional surface. Close-packing in monolayers is related to the density of the film with reference to crystalline polyethylene. Monolayers are more ordered than liquids or amorphous solids. The terminal group has orientational disorder. It has been shown by studies that the orientations of terminal groups freeze only at fairly low temperatures of only a few Kelvins. It is also possible to exchange the molecules at the surface with molecules in solution. This could produce a monolayer of mixed molecules. This implies that monolayer ↔ solution equilibrium is dynamic.

Variety of SAMs

There are several different kinds of SAMs. Carboxylic acids on aluminium oxide, silver and such other surfaces constitute one class. Long alkane chains terminating in a -COOH group are the amphiphiles and upon the monolayer formation, the surfaces become hydrophobic. This is the most immediate change that one can observe (note that monolayers cannot be seen, since they are only of molecular dimensions). The other class of monolayers are alkylsilane derivatives, $RSiX_3$, R_2SiX_2 or R_3SiX where X is chlorine or alkoxy and R is a long alkyl chain with or without a functionality. A surface bearing hydroxyl groups such as SiO_2, SnO_2 or TiO_2 is exposed to a solution of these molecules for a few minutes. What occurs is that the surface hydroxyl groups react with Si–Cl bonds and a chain of Si–O–Si bonds forms at the surface leading to a compact monolayer as shown in *Figure 3*. The surface does not contain any chlorine at all. This monolayer can be further converted to form a multilayer as shown in *Figure 3*. The difference between this and the LB methodology is that the molecules are chemically bound to the surface and to each other. This makes the surfaces thermally and mechanically stable so that applications can be thought of.

Figure 3. (a) A monolayer of alkyltrichlorosilane on a hydroxylated surface. It is possible to make (b) a hydroxy terminated surface and then to make (c) a multilayer from such a surface. Adapted from A Ulman, An Introduction to Ultrathin Organic Films: from Langmuir–Blodgett to Self Assembly, Academic Press, New York, 1991.