CEYLON ASSOCIATION

FOR THE

ADVANCEMENT OF SCIENCE

PART II

PROCEEDINGS

OF THE

TWENTIETH ANNUAL SESSION

22nd to 25th September 1964

COLOMBO 1965

CEYLON ASSOCIATION

FOR THE

ADVANCEMENT OF SCIENCE

PROCEEDINGS

OF THE

TWENTIETH ANNUAL SESSION

22nd to 25th September 1964

CONTENTS — PART II

COLOMBO 1965

3546—1

TWENTIETH ANNUAL SESSION
of
The Ceylon Association for the Advancement of Science
Tuesday, 22nd September, 1964

Opening Address

by

SIR NICHOLAS ATTYGALLE,
Vice-Chancellor, University of Ceylon.

Mr. President, Your Excellencies, distinguished guests and members of the Association,

I consider it a great privilege extended to me by the members of the Association in inviting me to open the Twentieth Session of the Ceylon Association for the Advancement of Science. I believe this is the first occasion on which a member of the Association has been called upon to declare open the sessions.

In Ceylon, in the past years, more stress, has been laid in recounting our past history. For many years our teaching has been mainly devoted to the liberal Arts. It is only in the last few decades that special attention has been paid to the teaching of natural sciences. The teaching of modern social sciences has only been started very recently.

All these years we have mainly devoted our time to developing tea, rubber and coconut industries and rice. In spite of all these we are importing 45% of the rice requirements of the island, which forms our staple diet. There is another factor which should find an important place in future planning, that is the explosive increase of the birth rate in this country. It is reported to be the highest in the world. The family planning project has been in existence in this country for some time now, and with the discovery of oral medicines to prevent birth and the use of other appliances it should be possible to control the population increase in such a way that a reasonable limitation of the population could be aimed at. It is to be clearly understood that we are living in a small island of about 25,332 square miles. The recent trend has been to industrialise our country. The Government has launched many projects for small scale industries. Today a large part of our requirements have to be imported from foreign countries. Even dried fish and salt are imported in a small island like ours.

The free education scheme was introduced in 1947 and education is compulsory up to the age of 14. Under this scheme education is free from the kindergarten to the end of the University course. Even the professional courses like medicine, engineering and law are free. It is a pity that more attention is not paid to the establishment of science laboratories in the rural areas, and the establishment of schools of technology. This should have followed the introduction of free education. Even today we find only a few schools of technology. Science laboratories where science could be taught up to the University level are limited. A few technological schools have been established but it is difficult to attract students to these schools. The remedy is very clear, that is to pay technologists higher salaries than white collar workers.

I have for a long time advocated the establishment of laboratories in rural areas and to do away with the system of having small laboratories in each school.

It is my sincere hope that a large number of scientists will be trained in the near future and be paid adequate salaries so that our industries and scientific projects can be properly manned and worked in a manner so that it will meet the needs of our country.

Our standard of living is low. Our only salvation lies in new discoveries and inventions in science, which offer undreamt of standards of living and comfort.

With these few words I declare open the 20th session of the Ceylon Association for the Advancement of Science.

General President's Address
22nd September, 1964

CEYLON NEEDS A FRESH ORIENTATION FOR ADVANCEMENT

by

N. D. Wijesekera

This is an occasion when the General President of the Ceylon Association for the Advancement of Science has the privilege of meeting all its members. Therefore, it is fitting that, I, should thank you for the honour done me, and the Section I represent, for electing me General President for 1964. A statutory obligation, as it were, of the General President is that he should deliver an address at the Annual Sessions.

On this occasion I do not propose to speak on any speciality either mine or anyone else's. All of us are aware that a tremendous force is developing throughout the world for the spread of science and technology. A new generation of men and women have been ushered into a new world after World War II. The oft-repeated need of this generation has been reconstruction, regeneration and restoration. Every demand seems to be directed towards creating a better world for everyone to exist and a safer world for all to co-exist. I wish to state at the very outset that Ceylon needs a change. It is a vital and dynamic change that is essential. It is a change of outlook and attitude of mind, and a change of direction, in fact, a complete re-orientation of the aims and objects of advancement for this country.

Before proceeding to speak on this change I may be permitted to preface the address with a few remarks. In this part of the world we have always accepted change as an inexorable law of nature applicable to natural phenomena, living beings and society. It is not merely a change yielding place to regeneration and renaissance but also a willingness to accept the change as leading to progress. Whatever causes may bring about such changes in nature or society, be they inventions or discoveries, these changes tend to create a chain of events. Such changes as are conceived in the mind through knowledge and science have in the end been conducive to the greatest good of mankind. Change in cultural elements, too, has affected the complexion and structure of society. For example, fire is such a simple discovery that revolutionized development of early man. Domestication of animals and agriculture led to further social developments. These changes appear so simple today but they have in fact revolutionized human advancement.

There is no doubt that man has today surpassed his ancestors by a series of changes, some revolutionary others evolutionary, still others technological, economic and social. The potential of the mental intake and output far surpasses that of men of any other age. Man has invented machines more complicated and efficient than his own brain. These are capable of doing intricate work which the human brain may take years. His way of life is differently adjusted for peace and not for aggression. His mind is better attuned for co-operation on a global basis and better disposed towards greater human understanding. All this has been possible by an increase of knowledge and techniques and processes in using that knowledge for definite ends—i.e. by education.

The whole complex of the operative mechanism of advancement is now called "Science", and "Technology". There is a tendency to separate science from another department of knowledge generally known as arts. But happily for those in Ceylon and India men have always used 'vidyā' to mean knowledge which included both arts and science. In fact the UNESCO concept too, I believe, is the same. What then is the purpose of science or knowledge. Knowledge enables man to overcome his environment and control it at will for his benefit. Science not only brings about man's advancement 'vidyā vṛiddhiṁ avāpnōti' but it also endows man with discipline. "Vidyā dadāti vinayaṁ". Ceylon, fortunately is heir to such a tradition. Unfortunately, continuity of the scientific tradition has been broken for periods longer than is good for the health of the population. History of this Island explains this better than any other record available. Nevertheless, that same record gives the present generation hope and inspiration to strive once more to attain a state of advancement through modern science.

Two thousand five hundred years ago the people of Ceylon were faced with a serious problem. Professor Arnold Toynbee reminds us of the "arduous feat of conquering the parched plains for agriculture by the missionaries of the Indic-Civilization in Ceylon who once achieved the *tour de force* of compelling the monsoon smitten highlands to give water and life and wealth to the plains which nature had condemned to lie parched and desolate" (Toynbee, 1956). Science as we know and understand the concept today seemed hardly to have existed in those days. But there seems to have existed a body of knowledge in the possession of men of those times which helped to respond magnificiently to the challenge. "Nothing is known today of how these engineers of old and the technicians under them set about their work . . . All this information must have been contained in text books (University History of Ceylon, 1959). The people by recourse to such traditional knowledge have by their own efforts exploited the natural resources, tamed the rivers and developed the land to provide abundance of rice and food sufficient for a large population to subsist and prosper. Along with this they developed an art and architecture inspired by the religious impulses and spiritual values against a traditional background. Some of the objects of beauty created by them express with eloquence the ideals of that age. That body of science is lost today. The art is no more. It is not possible to revitalize a body of traditional knowledge so built up through centuries of experience when once the genius of a people has been suppressed and their ethos is chilled. What was obtaining could not and may not have faded out entirely. But with the impact of conflicting forces the knowledge accumulated through centuries and transmitted in pupillary and family succession receded into the backwoods of the mind until near extinction.

The mechanism of advancement adopted a different method, technique and medium. The men who should contribute towards success came from a different strain and, if from the same, from a strain of society which perhaps lacked hereditary potential for transmitting any dominant carrier patterns. The hoped for advancement benefitted only a minority of the population. Knowledge lingered with these new elements whilst presumed ignorance prevailed amongst the larger majority of the population whose minds, too, got colonized. In this strange and new set up the intelligence of the entire population could not be utilized but remained, perhaps latent in the mental make up, tape-recorded as it were, to be recaptured through more reactive methods and techniques. A minority of the people of Ceylon have been thus forced to advance along a devious path through a different medium to respond to a challenge of an artificial environment. The response by this new group of people to such a challenge could not be natural, vigorous and successful.

In such a group a century of new experience is available in the records pertaining to every aspect of science and technology. Practically every aspect of the scientific sphere was embraced by the individual research workers as well as those at Institutes. These included medical and veterinary sciences, agriculture and forestry, engineering, natural sciences, physical science and social science. The scientists engaged in the various fields of study are the products of a system of training prevailing for more than a century. The contributions made individually and institutionally may not be publicly known, or, if known, may not appear to be spectacular when compared with the attainments of other developed nations. It does however indicate the existence of a body of scientists and a body of scientific knowledge. The inquiries conducted into scientific progress has revealed so many encouraging features in the prevailing

outlook, facilities and organisations. If the development of the land through scientists and scientific knowledge is to be real and beneficial then the hitherto prevailing pattern must change for the simple but obvious reason that the fabric is not woven out of the warp and weft of national threads. To realize this the minds of the population have to be decolonized first.

The scientific workers have not yet adjusted themselves to the natural environment and the people of Ceylon are faced with a problem no different, nor more difficult, than that confronted by their ancestors. True, it is more complicated and intricate now than before since the problem our ancestors had to face related to an early stage of the nation's struggle. The solution must be different and the approach too must needs be different because the population, social organisation, political set up and technological processes are also much more complex. Therefore, a fresh orientation must be introduced. Advancement in the modern sense may require a complex mechanism. In the society of Ceylon advancement must take into account, not only technological and technical processes, economic and social orders, but also traditional values of an agricultural community. Material advancement must necessarily be induced through technical and technological processes and for that science is the most important agent. That is not enough. Moral and spiritual advancement has to be realized through discipline by a spiritual way of life. The tree of national culture has roots sunk deep into the past and cannot be hurriedly uprooted either by science or technology. Science and culture are different facets of the same process of evolution. One without the other cannot progress but comes to be outdated through its own inertia or disharmony. A synthesis of the two yields the highest potential for ultimate progress. Lack of the one without the other has caused most of the ills of present day societies some of which have advanced unilaterally beyond a stage of possible correction. Such materially and technologically advanced societies do feel and fear that humanity might destroy itself by its own folly unless the thoughts of one track minds are re-oriented and redirected to react to two impulses which will sustain the harmony of mind and body by a judicious synthesis of bilaterally contending forces of material advancement and spiritual contentment.

"Is it necessary for every society to pass through all the historical stages in advancement? Can't we skip many non-essential stages and shorten the process of advancement?" Torres Bodet, Director General of UNESCO, answered this. "I say you can". He went further and said, "The destiny of mankind is an indivisible responsibility which we all must share". If this is true then it is equally true that those who are advanced must give freely to those who are less advanced but are developing. Knowledge of science with its technology and know-how should be readily available in order to inspire

the hearts and minds of a developing people. Out of the minds and hearts spring the desire for confidence of the giver and the gift. Therefore it is in the minds of men that the seeds of advancement must be sown.

In different countries the path of advancement has been through different avenues all leading to one ultimate goal. Today the movement is accelerated by the mechanism of the approach. It is a universally accepted fact that science has provided both the mechanism and the power to motivate that mechanism. Power is the most important factor that has to be mobilized. In mobilizing this various societies have devised various means. The level and rate of advancement can be determined by this factor of power. The rate of advancement can be judged or measured by the quantum of energy developed. Today it is electrical energy. Tomorrow it is atomic energy. In the future it may be by some other agency through which power will be generated. We have to develop from the present stage. Here a change of direction is essential and the mechanism of that change has already been determined. In Ceylon we have to motivate the body as well as the mind and if the mind can be conditioned to the required change advancement can be assured.

I mean science for advancement of a society has to fulfil certain prerequisites. It is the duty of the present generation of scientists to ensure that these determinants are satisfied. How can this be done? This is the challenge the scientists must take up. This is the moment for society to co-operate with the scientists. This is the opportunity for the State to change the minds of the people and the hearts of the scientists to combine their talents for development. All these years of foreign occupation and even after attainment of independence the management of scientific personnel and orientation of science programmes had been the responsibility of civil secretariats as agencies for supplying the *ad hoc* essential needs of administration. A fresh reorientation is a *sine qua non* for all advancement, be it through science, technology or art. What should that be. A journey of discovery into the rural countryside brings to mind a remarkable phenomenon. One sees tanks, big and small, ruins standing in desolation and smiling faces of men and women of a fast growing population living in that environment. A change of a fundamental nature has taken place to disintegrate these elements, which had been closely integrated. The population has increased fast and is increasing at an alarming rate. The once irrigated and populated countryside has relapsed into the jungle.

The challenge before the nations is "can we continue the human exertions to maintain the miraculous transformation of the face of nature in Ceylon". Can we grow enough food to sustain a fast growing population which doubles every generation or can we

limit or plan families to maintain the population within manageable limits. We of the present generation could do both better than in the past because of science provided we are bold enough and brave enough to change the outlook, attitude of mind and deviate from the path so far followed. It is a remarkable fact that the pattern of this transformation has been the same throughout history. The miraculous change seemed dramatic when it was first effected in about the 3rd century B.C. There have been lapses and relapses during cycles of 800 years. But whenever those in authority decided on the change the nation responded. A new era of prosperity and plenty dawned on the ancestors of the men and women who are still smiling. Another such period of prosperity was that of the 5th century A.D. when King Dhatusena and Mahasena promoted irrigation. For the third time the cycle was completed in the 13th century A.D. when King Parakrama the Great changed the even course by his wisdom to build tanks and conserve every drop of rain for the benefit of man.

It will be apparent that one fact emerges boldly. The pattern for advancement has been essentially and basically the same. Whenever this pattern was changed or could not be followed due to a clash of political, social or commercial interests the nation's exertions became paralysed. The people witnessed the relapse of that "once irrigated and cultivated and populated countryside into its primeval barrenness and destruction". History provides the solution and it lies in irrigation and rice cultivation and agriculture. It is the only remedy to cure the ills of despondency and frustration of unemployment and increasing population.

This does not mean that the nation is to abandon all the other projects for development such as industrialization, power production, technological improvements, plantation industries, animal husbandry and food production. It simply means that all these techniques and knowledge must be harnessed to make the change possible without delay. For the transformation to be successful there must be a plan and it must have a new orientation. It must be worked with the co-operation of the people as a whole. No development can be successful unless the individual becomes an active agent. The village should be the unit of development and the family the constituent member of that village. The personality of the individual being must be ennobled, and the dignity of men must be restored. Their standard of living must be raised. Here again the key to the problem is found in our social organization. The salvation lies in small tanks. There were village tanks scattered throughout the Island. At least about 10,000 may have existed. There may have been more. Every village practically had a tank of its own. "The general emphasis should be on the smaller schemes. To avoid such lasting damage to future development, increased attention should be

given first of all to the remodelling of the village schemes . . . The efficiency of irrigation is often higher in the village works where everyone is interested in the better use of the available amount of water" (Report of International Bank, 1952). Moreover, large irrigation channels have raised the water table and increased the salinity of the soil thereby rendering it less fertile for cultivation. Today there are nearly 20,000 villages with a population of 10 million. 100 families of 5 persons for each must be reorganised to be active co-operators. The emphasis must change. The approach must be realistic. Responsibility must be devolved on the village and the individual outlook, attitude of mind and response to the challenge must be socialized. The individual must be the most important factor in the enterprise of national advancement through development.

One may wonder how an ancient pattern or a past culture complex can be revived. It is not the culture that is being revived or restored. It is the eternal values underlying the fundamental ideas that are expected to be revived, remoulded and re-fashioned to fit the times and this hope is to be achieved by energising the minds of men vitalizing their bodies and adopting techniques and machinery that were never available before. Development is progressive advancement and different men in different countries at different times achieved this goal by stimulating the interest and ennobling the genius of its people in accordance with stupendous plans whereby all the resources of the land are exploited for the supreme benefit of men of that land.

Then again it is popularly believed that Ceylon is committed to an age old way of life and a social organisation that resists change. Such popular fallacies exist. Nevertheless, Ceylon like other countries with a very ancient culture and civilization presents several seeming impediments, prejudices and difficulties. No one likes change. Both men and society resist change. Tradition is sensitive to change. Old beliefs die hard. Nevertheless, all societies change. Such changes as are beneficial to man constitute progress. Attempts were often made by those less informed to develop Ceylon in other ways. The attempts made during the past to get back to the right path have confirmed the fact that rice cultivation must take pride of place and priority of attention over all other forms of development. How else can the paradox of prosperity through tea, rubber and coconut in the midst of appalling poverty be explained. The

obvious answer is that the plantation industries are not in the line of natural development. Natural development is through agriculture and rice cultivation by land utilization with irrigation.

Traditionalism has often been blamed by those who fail to grasp the significance of values embodied within traditional culture. Traditional culture associated with ancient races of mankind has helped to preserve the eternal values that helped such races to survive. In the traditions themselves are embodied the potentialities for development. These have acted as shock absorbers, national genii for safeguarding and preserving the moral stamina and spiritual strength of the people. These formed a bulwark against dangers from within and without.

Industrial revolutions and social evolutions have effected radical and quick changes in young countries by sweeping the old and reducing everything—people, culture, way of life and thought into a drab pattern. The people have drifted further away from their moorings and have lost their bearings. Their variety has been submerged and their vivacity has faded out of recognition. Men are thinking in terms of machines and not the minds within the men who drive them. But, in the East and Ceylon, tradition is more than a mere accretion through ages. It is the spirit that gives life to the body national. Knowledge constituted a part of this tradition and it was dynamic and evolving all the time. The people have been encouraged to resort to scientific investigation and mental analysis. They have been told not to accept anything merely because the teacher said so, merely because tradition preserved it, but to accept after testing whether it agrees with reason and logic. Such a tradition, surely, cannot be an impediment to science. It is certainly a scientific attitude. Until these criteria of judgement are fulfilled the society firmly adheres to the tradition. This tradition is preserved in myths, legends, fables and stories which have become the repository and storehouse of people's knowledge. The authority for change must be sanctified by the assurance of a balance between the traditional way and new way of life. The balance leads to harmony of body and mind which alone can lead to advancement. The ethos of the nation pulsates within the framework of the traditional environment whose depths can be illumined by the light of knowledge. Thus genius and ethos come into free play generating the latent energy to power national development. Knowledge alone will help to remove this veil of superstition and ignorance to be able to penetrate the armour protecting the face of the environment.

Another method of preserving such ancient lore is the organisation of society into functional groups. It is true that in course of time the functional structure became stereotype into so-called castes, arranged in an order of superiority according to function.

Strangely enough the highest of these categories was the agricultural community. From that point of view it did not hinder cultivation. But it lead to other unpleasant differences affecting human dignity and equality of man.

Education in the past has not helped to eliminate this social distinction. In fact, it has helped to perpetuate it. Moreover, the type of education imparted in a foreign medium has tended to produce a new type of snobbery which led to abandonment of cultivation and agriculture as being unworthy of the educated. The only honourable work could be in the offices generally as clerks. The inherent aptitudes, manual dexterity, technical skills had no opportunities for development. There was a mass exodus from rural areas of nearly 70% of the population who shunned work, who disliked agriculture and who became not only unemployed but unemployable. Education has also to be re-orientated to be attuned to the exertion of maintaining development through cultivation. That is the real need and the most urgent one at that.

Therefore, education has to be re-orientated as quickly as possible. It has to undo what has been done badly. It has to generate a new attitude in relation to the natural environment. This can be done in two ways as there are two groups to be served. One group is the grown up population, so-called adults. The other is the school going population and youth. Fortunately for this country the percentage of literacy has always been reasonably high. This is the result of the ancient tradition of learning. New ideas, new techniques, new scientific knowledge has still to be popularized. It is here science with its technological advances comes to the aid of society. Mass communication media such as radio, television, sound films should become essential requisites and not dispensable luxuries. In this campaign television certainly is the most effective and fruitful medium. Ceylon has to catch up with time and make good its lost ground. The other sector is far more important as the persons concerned will be the executors, administrators and scientists whose exertions will be essential for development. The technologists and technicians, research scientists and engineers will be trained from among them. Trained personnel should not be wasted nor encouraged to leave but should be attracted to return and work in Ceylon. The dichotomy that exists in the fields of knowledge has tended to separate arts and sciences and helped to develop two kinds of citizens, viz., others and scientists. Between them there is such a lack of understanding that they seem to work in a sphere totally unrelated to the environment and the community whom they are supposed to serve. Therefore all those receiving education in schools, universities; scientific and research institutes should be made fully aware of their obligation to society, of the need to know the background of the people and their environment.

A change long over-due but which has since been introduced is the medium of education. National languages have been accepted as the media at all levels including the University. This will provide the highest incentive and the greatest appeal as years pass by. Science can be learnt better, quicker and more in consonance with the traditional environment. Consequently, the new ideas will be expressed more precisely and conveniently. There is no alternative short cut and without this change no advancement could have been possible. To maintain this standard of literacy, knowledge must be made easily available in these languages. In this regard scientific knowledge needs to be translated into the languages of the people so that science will benefit everyone and not remain a prerogative of the affluent.

Later, British period appreciated the problem of the advancement of the nation somewhat on correct lines but failed to pursue the course relentlessly and with sufficient courage and foresight. The idea was seized by the Ceylonese too, but man-power and finance were not deployed to achieve success in the transformation of the face of nature and the mind of man. If success is to be assured and in good time to recapture the lost ground and time then the entire nation should be galvanized along the path of development. The concentration must be on science. Research, both pure and applied, should be geared to aid by fresh knowledge. Personnel should be given specific assignments according to urgency and priority. There has to be a tremendous effort with the supreme emphasis on agriculture. Self-glorification and individual advancement should find no place. Anonymity is the tradition of our ancient artists and scientists. Self-sacrifice, self-denial, self-reliance and self-sufficiency should be the ideals. The challenge is not new. It persisted through centuries and has manifested itself once in every cycle of 800 years. The challenge was accepted and responded every time by developing agriculture through irrigation and increasing rice cultivation. And thus the transformation has been effected. It has resulted in self-sufficiency in food and efflorence of a culture and civilization. History *does* repeat itself and if it has done so three times during the last 2400 years it can do so again.

References

Report of Mission: International Bank for Reconstruction and Development (1952). *Economic Development of Ceylon*; p. 430. Baltimore: Johns Hopkins.

Toynbee, A. (1956). *A Study of History*, Vol. 2, p. 5. Oxford University Press.

University History of Ceylon (1959). Vol. 1, p. 355. Colombo: Ceylon University Press.

SECTION A—MEDICAL AND VETERINARY SCIENCES

Presidential Address
22nd September, 1964

"IN SEARCH OF THE NATURE OF DISEASE"

by

W. D. Ratnavale

Since Health and Disease are the chief concerns of this Section, I have selected disease as my subject with the title "In Search of the Nature of Disease".

I therefore propose to review in broad perspective man's attempt to understand disease and to trace the paths progressed in reaching our present position of knowledge. There is much that we know about disease—indeed quite a lot—but there is very much more unknown. Of some diseases we know much more than of other diseases but in almost all we have yet to know the final answer. We have not got there yet, and hence the title of my subject—"In Search of the Nature of Disease".

I have used the word 'Nature' of disease to make us look at disease in the way it should be studied—not as a thing or an object but as a continual process from beginning to end; because when we come up against it, face to face as individuals we see only just one scene in a whole act or drama and unless we know what went before and what comes hereafter our grasp of disease cannot but be adequate.

Disease is life under abnormal conditions and therefore is as old as life itself. It is not a new thing. It has travelled side by side with life. It is not restricted to man alone. It effects animals, plants and all living things. It has had a tremendous influence on life, so that the life we see today is the result of millions of years of adaptation of primordial life to altered circumstance. There are many forms of disease, many of them have common features but there is no single cause for all disease. The only common feature of all disease is the process of still-keeping alive in spite of it; and if life can be described as a happy state among a group of molecules, disease as we shall see later, is a disturbed or unhappy state among them.

Disease then is a deviation from normality producing either morphologic manifestations, functional manifestations or psychic

manifestations. It may be temporary or permanent and its awareness may be restricted to the individual concerned or only to those around him or to all and sundry.

It is not my purpose in this talk to take any disease or group of diseases and work out causative factors. Nor do I intend to try and give any explanation for the numerous forms of disease. I do not propose to describe any scheme that could be used or has been used in elucidating the nature of any specific disease. All I hope to do is to try and build a composite picture of the type of thinking, and describe a small fraction of the methods employed to support that thinking down the ages, during man's search for the nature of diseases in general.

Primitive man must have had his own idea of the cause and nature of disease. He certainly must have believed in supernatural beings, beings responsible for the strange happening around him. The weird jungle sounds, lightning and thunder, earthquakes and floods could not have been otherwise explained. So also disease. He probably thought that disease was the work of evil spirits. Fear must have played a big part in his thinking. Man's mind is such that the only way in which he can comprehend the variety of things happening around him is to classify them into experiences in the light of his understanding and with early man evil spirits must have seemed to be the most reasonable and likely cause. Of course to many even in this country and also elsewhere, they still seem to be very powerful agents of disease. They have to be driven away by 'thovil ceremonies' or devil dancing ceremonies and kept at arm's distance by wearing amulets and charms. Evil spirits and magic go hand in hand and you have no doubt heard how some diseases are still thought to be due to evil objects buried around a patient's house.

Magic was part of primitive religion and later since religion seemed to provide the answers to things unseen, many if not all disease conditions were soon placed on a religious basis. Disease was then thought to be divine punishment or an expression of the wrath of God. It was a punishment for sin or sometimes as the Greek and Hindu mythology would have it, diseases were sent on men for the mere amusement of the Gods.

Eventually however groups of men, broke away from magico-religious bonds and formulated their own theories of disease. The Greek Pythogorean School believed that the human body was composed of four substances, earth, fire, air, and water and that health resulted from their balance while disease from their imbalance. Here were men seeking to understand the nature of disease by introducing hypothetical components. Subsequently the latter Greeks

produced the theory of balance of humors which Hippocrates further expounded. They said that the body contained four humors, blood, mucus, black bile and yellow bile. When they were present in balanced proportions there was health but should one of these be in excess the body underwent disease and after a variety of manifestations the excess humor boiled over and was excreted in the urine, in the stools, or in the vomit or as pus. These Greeks had not only broken away from the idea of supernatural influence but they had also introduced some form of reasoning into their thinking. Hippocrates in describing epilepsy—the so-called sacred disease said "It is not my opinion any more divine or more sacred than other disease but has a natural cause its supposed divine origin is due to man's inexperience".

Unfortunately with the fading away of the Greek civilisation this liberal approach to the study of the nature of disease disappeared. Rational thinking as to the nature and the origin of disease was lost. Galen in the 2nd century propounded his own theories of disease without any rational background and man made very little attempt to understand disease, but merely accepted it as will of God.

After almost 14 hundred years of sleep during the dark ages, mankind awoke with the Renaissance and ushered in the Scientific era. There was re-awakening in literature and in art and all the numerous branches of science. Many workers were interested in disease and progress was made in all the aspects of medicine. The famous Leonardo de Vinci himself dissected thirty corpses and made excellent diagrams therefrom. Anatomy was studied extensively and there was great interest in Physiology.

Towards the end of the 15th century the search for the nature of diseases had recommended. Post-mortem examinations were done with a view to finding out the cause of death, and a new concept was born—the concept that diseases were due to alterations from the normal structure of body tissues. Thus evolved the beginnings of the subject of Morbid Anatomy. Morgagni (1682-1771) was the first to make a serious study of this aspect of desease and in 1761 he produced his famous book, *De Sedibus* in which he described in detail the changes seen in gross morbid anatomy and tried to explain the symptoms of diseased persons in the light of changes in the diseased organs. This a very important land mark in man's search for the nature of disease. Prior to the Renaissance man's concept of disease was mainly humoral but with the Renaissance one can justly say that a new era had dawned—the Era of Anatomical alterations—the Organic Era.

Post-mortem examinations revealed in many diseases, gross abnormalities in certain organs. Alcholics died leaving behind hard knobbly livers described as the hobnail liver of cirrhosis. People

who had rheumatic fever had abnormal hearts, some were described as the bovine hearts resembling that of an ox. There were a variety of picturesque description for many diseased organs. There were strawberry gall bladders, geographical tongues, flea-bitten kidneys, sago spleens, leather bottle stomachs, pipe-stem arteries and several other less picturesque description of a wide variety of abnormalities. Every departure from the normal was a pathological state and these were carefully studied. Here was a tangible evidence of abnormality. No longer was disease a question of vague hypotheses. Specimens obtained from dying persons in different stages of disease helped in piecing together the natural history of any disease. Of course there were many incorrect conclusions drawn, but with improvements in techniques and better data, the causes and nature of several diseases became more clear. There were also several diseases in which no gross anatomical changes were detected and these had to wait other methods of examinations. Thus we see that towards the end of the 18th century man's idea of disease had begun to be orientated on the basis of lesions seen in gross morbid anatomy. The why and the wherefore of these alterations were not always known but least they knew that alterations in the body organs were associated with many diseases.

Although simple microscope systems were being used by biologists about this time, it was not till the 19th century that the better type of instruments using apochromatic lenses were available. Human tissues up to then, examined by compound lens systems were believed to be composed of globules until Schleidon and Schwann showed that all animal tissues were composed of cells. The light microscope then opened up a vast field of discovery showing in detail how the large organs of the body were composed of small cells arranged in definite histological patterns. Virchow used the microscope to study the gross anatomical specimens and after seven years of very fruitful study published in 1859 his lectures in Cellular Pathology. Here he described the changes in cells and postulated the theory that all diseases had their origin in disturbed cell function. Virchow's work ushered in the era of cellular pathology and with the increasing use of the microscope, the next one hundred years opened out an enormous store of distorted cellular patterns which now form the basis of modern histopathological diagnosis. The search for the nature of disease had focussed itself down to the cell as the basis of all disease. Though this may not be strictly true this theory had a profound influence on the subsequent study of almost all disease processes.

Improvements of microscopic techniques helped workers to study the reactions of cells to a wide variety of conditions. The histological features of inflammation, circulatory disorders, tumour formation were revealed in hitherto unimagined detail. The microscope helped in the discovery of parasites, heleminthic and protozal,

and finally large numbers of bacteria. Bacteriology made rapid strides as a result of the increasing use of the microscope and several hundreds of organisms were discovered. Many of them were pathogenic to man and it was soon discovered that though the cells of the body reacted in a general pattern to almost all organisms there were in many instances slight variations in the cellular content and morphology of these reactions sufficient to give an indication as to the nature of the invading organisms. Disease processes began to be sorted out into aetiological groups. There were those that occurred as congenital manifestations, others were the result of trauma, others due to alterations of circulation, others were inflammatory in type and still other neoplastic. Some diseases appeared to be metabolic disorders and others were the result of extraneous agents, both biological and otherwise.

Much had been learnt about the causative agents of disease and much about the disease processes themselves. There still remained however, several gaps in the understanding of many. Almost all the information that could be gleaned both by macroscopic as well as microscopic examination of diseased organs still left some conditions unexplained and the search for the nature of such diseases had to be extended to the newer fields of experimental pathology. Using experimental animals under controlled conditions of experiments considerable amount of information was obtained in many diseases and with a certain allowance for species differences further knowledge was added to the nature of diseases in man.

Over the last 20 years there was an increasing drift from the study of tissues and cells. It was being realised that although the body was mainly composed of cells, there was also a considerable if not greater quantity of intercellular substance. Much of this intercellular substance was to be found within the fixed tissues but quite a considerable amount was met within the body fluids of plasma and lymph. With the discovery that these fluids may show considerable chemical alterations in diseased states there had been a phenomenal interest and work in the biochemical field.

Virchow's hypothesis placed the seat of all disease in the cell. However within recent years there have been found several circumstances that did not fit into this picture. All over the body, lying between cells as well as between groups of cells is an intercellular matrix, known as connective tissue. This connective tissue consists of protein fibres and interfibrillar ground substance. The fibres are collagen fibres, reticulum fibres and elastic fibres while the ground substance holds a large amount of body water and many types of acid mucopolysaccharides. We now know that there is an increasing group of diseases where these non-cellular elements have been found to be the seat of morphological and biochemical change and these diseases have acquired the status of being called collagen

disease. There are localised and diffused varieties and they include the well known diseases of rheumatic fever, rheumatoid arthritis, and scleroderma among quite a long list of several rarer ones. Moreover, increasing studies on the allergic diseases like asthma and hypersensitivity stress the role played not by cells but by independent chemical entities, like proteins, amines, and other metabolites. In addition there were several normal body reactions that took place without the immediate intervention of cells as for example the clotting of blood. Defective clotting processes often showed no morphologic abnormalities in body tissues. It became necessary therefore to look for disease, not only among cells but in the smaller units of macro-molecules, and molecules. Our search for the nature of disease had therefore narrowed down to units smaller than cells and we had entered the stage of subcellular pathology. This period has only just begun and at the most is only about 20 years old.

The advance into the subcellular phase opened up on two fronts. This became possible due to the spectacular advances in electronic engineering—but I would call an electronic revolution. This revolution has produced very elaborate and exceedingly complicated instruments to help us in our search for the nature of disease. Of the very many instruments in use I wish to make special mention of two. 1. ELECTRON MICROSCOPE and the other 2. The SPECTROPHOTOMETER. The former used by morbid anatomist and the latter by the biochemist in their joint advance into the subcellular field. It is difficult to follow these advances as separate activities because they are so closely interwoven. I will therefore progress along both pathways first stressing one aspect and then the other, returning every now and then to one or the other.

THE ELECTRON MICROSCOPE has placed pathologists of the 1950s in a similar position to what the light microscope had done to Virchow and the other workers a hundred years ago. Today in all parts of the world pathologists are seeing things they have never seen before. They are seeing things at a magnification of several hundred thousand times. They are amazed at the intricate structural details of the sub-cellular world and the medical literature of today carries a wide array of electron microscopic pictures of a variety of pathological states. Many hitherto unexplained problems have begun to appear in clearer light and a tremendous leap forward has been made in the search for the nature of disease. Bacteria which were only just visible with the light microscope are seen as very highly organised agents under the electron microscope and the viruses which were ultramicroscopic have been found to show a very great range of fascinating shapes and appearances.

In the meantime biochemists have taken tremendous strides and have shown in elaborate chemical detail the functions of living cells and cell extract. They have elucidated step by step the fantastic pathways of cellular metobolism; how food is used by living cells to produce energy to build up more cells and to control the functions of other more distant cells by a very intricate system of chemical messengers and controls. They have shown that the whole body is one vast chemical factory, performing a multitude of chemical reactions and electron transfers necessary to keep the huge mass of cells working and thus ensure the co-ordinated activity of all its component parts. They have also discovered how alternate pathways were evolved to cope with morphologic alterations as well as helped to explain the morphological alterations that occurred as a result of a disturbed biochemical processes. Within the last few years biochemistry and morbid anatomy have had to travel the same stretch of road as equal partners in the search for the cause of disease. The new subjects Histochemistry, Immunochemistry, Fluorescence Microscopy bear witness to this partnership.

Disease is no more a question of magic, no more one of spirits or humors or religious teachings. Nor is it purely due to deformed organs or distorted cells. Its basis now appears to be that of disturbed molecules. Molecules have the ability to initiate many reactions. In fact in the human body many molecules are introduced into the blood stream to initiate or even stop a reaction far removed from the site of the elaboration of the molecules. In the living animals we find many thousands of such substances and the great majority of them are called enzymes. It is normal regulated and co-ordinated activily of these enzymes that make the basis of life and it is the suppressed, deviated and abnormal action of these substances that produce the manifestations of disease. Thus the cell appears to be a mechanism or apparatus fabricated to regulate activity of these enzymes both quantitively and spatially.

Although this applies to every cell in the body certain specialised cells have been endowed with far more enzymes than others. The liver cell has a very great variety of them and is thus able to produce quite a number of proteins. Should one of these enzymes or proteins be defective we then have the stage set for disease. Let us take an example—the liver normally produces among many other substances a substance known as Anti Haemophilic Globulin (AHG). This substance is necessary for the clotting of blood. In some persons the liver cells do not make any AHG or only in feeble proportions. Such a person will suffer from Haemophilia and may bleed to death from trivial injuries. In our search for the nature of disease we have to find out why it is that this patient's liver cell does not produce enough AHG. It is in the search for this type of information that it has become necessary to look for data within the liver cell itself.

In cancer we know that a group of cells suddenly become disorientated and multiply at a rate much faster than the adjacent cells. The fault is obviously within the cell. There has been a disturbance of the normal intracellular mechanism controlling cell division and cell growth. Our search for the nature of cancer must therefore be to determine the intricate details of the normal mechanisms as well as how and what caused the change. Unfortunately cancer is not one disease. There are many different cancers. The nature of one will not be same as the other. In many we know that genes play a part. In some the genes changes are hereditary, in other genes changes are due to chemicals and still others due to viruses while in most we do not know. Nor do we know the exact nature of the change. However the search goes on relentlessly.

We thus see that in the ultimate analysis of many diseases it will be necessary to appreciate fully the workings within the cell both in normal state as well as in abnormal conditions. The electron microscope has revealed very highly organised internal structure of all cells and the biochemist with ultra speed centifuges has broken down each of the intracellular structures and worked out their chemical nature. It is now known that the working of the normal cell is exceedingly complex. The normal cell cytoplasm manufactures the necessary proteins and enzymes on receipt of a specific code of instructions from the nucleus. These instructions are already available in something like tape recorded form on long molecular filaments known as chromosomes loops. One or more of them constitute genes and special molecules known as messenger RNA's entwine themselves around these genes and having taken the imprint, pass through the nuclear wall and hasten to the ribosomes on the endoplasmic rectculum of the cytoplasm, to start the machinery required to produce the necessary proteins or enzymes. In the example given earlier with the liver cells, it starts to produce AHG. It will thus be seen that if any part of this complicated programme gets stuck, no AHG will be produced. Thus in our search for the nature of disease we find ourselves rummaging so to say within the cell for the answer, in and among chemical entities of the nature of macro-molecules.

Of late exceedingly interesting discoveries have been made among the structures inside the nucleus of the cell viz. chromosomes these have shown variations in certain diseased states. Every normal somatic human cell has twenty three pairs of chromosomes in its nucleus and by using relatively simple techniques it is possible to separate them, classify them and look for abnormalities. In the normal male the twenty third pair consists of X and Z chromosomes whereas in the female the twenty third pair is XX. Deviations from this normal have been found in abnormal human beings.

In the condition of Mongolism children are born with defects and chromosome analysis in these cases have almost always shown an extra chromosomes in either the twenty first or the twenty second pair making a total of forty seven instead of forty six chromosomes.

Interest in chromosomes is steadily rising, and let us look at them further. If suitably treated with chemicals they swell up and when examined under the electron microscope they are found to be quite complex structures. In a picture of a tiny segment of a chromosome there is a fine feathery structure it has received the name "Lamp brush" chromosomes. All those tiny strands are DNA molecules and these constitute the genes. These genes carry the pattern of the molecules that must be manufactured for the needs of the living organisms. Should the pattern not be correct then the molecule produced will not be normal.

I have now reached one but the last station in our search for the nature of disease and would like to mention here the case of the disease sickle cell anaemia. In this disease little children become anaemic and suffer from many symptoms. When their blood is examined under certain conditions of lowered oxygen tension their red cells take on a peculiar elongated and curved appearance discribed as sickle cell while normal red cells are nice and round. The abnormality is due to the fact that the haemoglobin within the red cells of these children is a little different from the normal haemoglobin and it crystallises into elongated needles while normal haemoglobin does not. The normal haemoglobin is called haemoglobin A and sickle cell haemoglobin is called haemoglobin S. The only difference between the two rests in the globin parts of the molecules. The globin is a chain of polypeptides and each polypeptides is made up a number of amino acids. Although the whole molecule has about 50 polypeptides it has been found that the fourth polypeptide had an amino acid composition slightly different in the two haemoglobins. In haemoglobin S the seventh position is occupied by valine instead of glutamic acid which occurs on the normal haemoglobin. This difference is believed to be due to the fact that the blue print or master plan of the gene responsible for the formation of haemoglobin has been altered.

We have come a long way along the road in the search for the Nature of disease. We have left a period of mystic ideas, religious opinions, passed though humoral pathways of "pithe" and "vathe", along diseased organs and cell patterns and find ourselves staring at abnormal molecules. We stand at present almost helpless in front of them—in front of these two molecules and several similar known and unknown ones and want to ask the question why and what made the change in the genetic code. What rubbed off the original

picture. We have almost come to the end of our search for the nature of at least this disease and find no answer yet for the last question. There is no need for us to hide our ignorance in the word mutation. Is it the effect of cosmic rays or emanations from outer space?

Is it the effect of drugs or molecules or do we have to complete the full cycle and ask whether this is not divine will, destiny, magic, evil spirits or just sheer mathematical probability or chance in a mass of colliding molecules.

I would like to end by quoting a few words from Fitzgerald's Omar Khayyam:—

For in the market place—one Dusk of Day
I watched the Potter thumping his wet clay
And with its all obliterated Tongue
It murmured "Gently, Brother, gently, pray!"

Listen again one evening at the close
Of Ramazan, ere the better moon arose
In that old Potter's shop I stood alone
With the clay population round me in rows.

And strange to tell among the earthen lot
Some could articulate while others not
And suddenly one more impatient cried
"Who—the Potter, pray and who the Pot?"

Then said another—"Surely not in vain
My substance from the common Earth was ta'en
That he who subtly wrought me into Shape
Should stamp me back to common Earth again."

None answered this; but after silence spake
A vessel of a more ungainly Make;
"They sneer at me for leaning all awry;
What! did the Hand then of the Potter shake?"

SECTION B—AGRICULTURE AND FORESTRY

Presidential Address
23rd September, 1964

THE ECONOMICS OF FORESTRY IN DEVELOPING COUNTRIES

by

A. E. K. Tisseverasinghe

Introduction

The title of my address needs some explanation. I did not select this title merely to be in the fashion and beat the economic development drum nor because I am a Specialist in the field of forest economics but because this is a most important subject for us in all developing countries and because it has been grossly neglected in the past.

There are two reasons for this neglect. The first reason is that Forest Economics began, and continued to be studied in, developed countries, and in the field of forestry at least, the problems facing developing countries are very, very different. In developed countries much stress is placed on the financial yield from forestry as compared with the rate at which money can be borrowed. Other factors are of little importance. I will give just one instance to illustrate this. In a scheme of reforestation, if the financial yield is 8% and the rate at which money can be borrowed, or the interest rate, is 6%, then the investment is considered to be a desirable one. The time taken for the investment to mature or what planners call the gestation period is of little account. This attitude is quite sound in the developed countries of the West where a large proportion of the forests are privately owned and where forest owners operate under conditions of more or less perfect competition and where investment capital is freely available. In developing countries the position is totally different. In most of these countries State Forestry is so predominant as to make conditions monopolistic. Very often the rest of the economy is also to some extent or other controlled by the State. Further, investment capital is not freely available and the little that is available has to be allotted according to a system of priorities. This position does not in any way invalidate the argument that the investment must be profitable in terms of financial yield, but this is merely one of the factors which are

taken into account. In Ceylon for instance, retail prices of timber sold at Government Depots are fixed by the State. The price is not fixed according to supply and demand or according to the capacity or willingness of the consumer to pay, but in such a way that timber is made available to all sections of the population at reasonable prices. The effect which factors of this kind will have on forest economics has not been studied.

The second reason for the neglect of the subject of forest economics in developing countries is that planners generally lump forestry together with agriculture. In developing countries generally, agriculture is an extremely important sector of the economy and tends to overshadow forestry. Man must eat to live, hence from the dawn of civilisation agriculture has been the most important and widespread human activity. The position today is no different. One of the reasons for the relative stagnation in Ceylon's economy is the fact that so much money has to be sent out of the country to purchase food. Planners therefore tend to give agriculture the highest priority and forestry suffers as a result. Planners also find that in terms of money yield per acre, forestry in developing countries is relatively inefficient and thus does not merit attention. For instance studies made in India show that forestry which uses 22% of the land contributes 0·6% to the National Income while the 40% of the land devoted to agriculture contributes 50% of the National Income—(Von Monroy 1960). In Nigeria, a country which even exports timber, the value of total timber production (excluding firewood) was less than 1% of the Gross Domestic Product. (Prest & Stewart 1953, quoted by Macgregor 1959). I am not aware how these calculations were made and hence cannot comment on their validity. However, in addition to the bare National income calculation if output per person employed were taken into account or income in relation to expenditure, forestry may score over agriculture. In any case, in developing countries where the people live at subsistence level, most of the domestic product must necessarily be agricultural in origin. To take a simple case: If Ceylon produced sufficient rice to feed 10 million people at the rate of one pound per person per day, the value of the annual production of rice alone on the basis of the present guaranteed price would be in the region of Rs. 1,400,000,000! Forestry can never hope to produce anything like this value if considered purely as a primary product. If however, industries based on these primary products are taken into account, the picture will be very different. Since such industries do not play a very significant part in the economies of developing countries, forestry has tended to be neglected!

I have taken some time to indicate why forestry in developing countries has been relatively neglected but the time has not been wasted because what has been said so far is relevant to the main theme of my address.

In developing countries generally, land is the most important factor of production. Leaving aside other minor forms of lands use, agriculture and forestry are the two main uses to which land is put. For various reasons which it is not necessary to go into now, a certain proportion of the land must be left in forest. In the words of a previous President of this Section, "forests, like the poor, must always be with us"! In Ceylon, the latest inventory based on aerial photographs of 1956 revealed that 44% of the land was under forest (Andrews 1960). The position has of course, changed since 1956 and will continue to go on changing but yet one can assume that a fairly substantial proportion of the land is going to remain as forest. Several questions now arise, are we to let the forest remain as it is, harvesting only the meagre increment which we are able to remove and use or are we to spend money to build roads into inaccessible forests, or invest in various plant which will help to make unusable species usable, or create factories to convert timber into other products such as particle board, fibre board, pulp or artificial fibres; should we convert the natural forest into high yielding plantations and so on.

Investment criteria in relation to forestry

There are several criteria to be used in deciding whether to invest money in forest development in preference to, or as complementary to other forms of investment. Some of these criteria are rather simple and are mentioned here only for the sake of completeness.

Since forestry is a form of land use one of the first criteria which will come to mind is the comparative income per acre of forestry as against other types of land use. I have already said that studies in India and Nigeria show that the income from Forestry is very small in contrast to the income from agriculture.

However, even without going into these two cases individually it is possible to say that such statements have to be considered in relation to several other factors and not in isolation. For instance, it is generally true to say that Agriculture is practised on the most fertile land and forestry on the least fertile. The annual expenditure on Agriculture per acre is generally very much higher than on forestry. The reason for this again is that in a subsistence economy all other activities are subordinated to food production. Relatively long term projects such as forest development got much less attention. If comparable amounts of money are spent on forestry, the results may be quite startling. The annual increment in the forests of Ceylon, taking all species into consideration, has been calculated at 4-16 cubic feet per acre for the Dry Zone and 16-95 cubic feet per acre for the Wet Zone (Andrews 1960). A good Eucalyptus

plantation on the other hand can produce more than 300 cubic feet/acre/year on a 10-year rotation. The expenditure for this would have been about Rs. 30/- per acre per year. Thus, a relatively small sum of money spent on plantations in a good site can produce yields per acre comparable with the annual value of some agricultural crops and also produce a very high income in relation to expenditure.

Another criterion which is used in deciding investment priorities is the extent to which the product can substitute for imports. Ceylon imports a very high proportion of her annual food requirements and hence much money is spent every year on encouraging local production of this food e.g. the relatively high proportion of the total investment which is spent on providing irrigation facilities, the high guaranteed price paid for rice produced in the country, subsidised fertilisers, agricultural extension services and so on. However, the extent of import substitution as a means of assessing the value of *industrial* projects has fallen into disfavour recently. With a tight foreign exchange situation other criteria have become more important. Every imported product need not necessarily be substituted for by a locally manufactured product—instead one can learn to do without. Mr. Maithripala Senanayake, Minister of Rural and Industrial Development addressing a meeting of technical personnel recently said: "We have now come to the stage of thinking in terms of a strict priority of industries rather than an earlier policy of import substitute industries. You will find gentlemen, that in at least 60% of our industries the major portion is imported in the form of industrial raw materials. While there is a foreign exchange saving, yet with increasing demand and increasing population, the foreign exchange saving resulting from industrial manufacture had disappeared".

The extent of possible import substitution in the forestry field is not high in terms of value compared to agriculture but it is much easier to achieve a break-through in this field e.g. 18 million rupees were spent in 1962 on the import of timber and timber products such as Tea Chests and 25·5 million rupees were spent on the import of paper and paper products. One should bear in mind that 1962 was a year when import restrictions operated very severely and hence these imports should be considered as being absolutely essential. In the case of agriculture, increased production involves land tenure reforms, the use of new methods of cultivation, the opening up of new land and a host of other factors which tend to make the process slow. In forestry, as Ceylon's short term implementation plan has recognised, the position is quite different. "Forests provide an existing resource which only awaits exploitation. The investment required to bring this resource into a condition where it could be used is relatively small. The Capital/

output ratio is also low while its gestation period is comparatively short (except in the case of reforestation). Priority should therefore be given to investment in forestry, and every effort should be made to increase the rate of exploitation in conjunction with a more rational utilization of local timbers based on the processing of a wider range of species". I will not labour the point about the comparative advantages of forestry as against Agriculture—the key to economic development in Ceylon is more productivity from the land and this includes both Agriculture and Forestry.

As far as industrial projects are concerned a whole host of criteria are now used in deciding priorities. These have been stated by the Minister of Rural and Industrial Development in the same address quoted above, in the following terms: "The Planning Secretariat in consultation with the Ministry, is also planning a set of criteria on which future approvals for industry will be based. Such factors as the gestation period, capital output ratio, capital labour ratio, capital foreign exchange ratio and capital profit ratio will receive increasing attention. But one fact has clearly become evident. Unless an industry makes use of a large percentage of our local raw materials, the chances of it being approved are slim". The reasons for this are obvious. In a period of extreme difficulty, imports of manufactured products can be stopped at the cost of temporary inconvenience, but if raw material imports have to be stopped the effect on the economy may be serious. Hence the stress on using local raw materials and hence also the importance of forest industries using local forest products as raw material.

One highly important criterion to be taken into account when deciding on the priorities for investment is the influence of particular types of investment on economic growth. The tends to be forgotten if planners and industrialists concentrate on the purely book keeping aspects of investment but for a developing country this should be the all important criterion.

The Indian 5–year plans have been severely criticised by some economists on the ground that excessive stress is placed on heavy industry at the expense of agriculture, education and etc. This kind of imbalance can happen only if the broad aims of development policy are lost sight of in the maze of immediate priorities. Reddaway (1962) commenting on the Indian economy says:

"Long range strategy must have one overriding objective: to secure a progressive rise in the level of consumption per head, by methods which will make possible an indefinite continuance of the process and which also satisfy certain social and political ideals . . . Other economic objectives such as increased capital formation and industrialisation, or a viable balance of payments—

are essentially dealing with means and not ends". Again Friedman (1955) (quoted by Bauer 1961) says about India: "There is a tendency not only in India but in most of the literature on economic development to regard the ratio of investment to National Income as almost the only key to the rate of development, to take if for granted that there is a rigid and mechanical ratio between the amount of investment and additions to output. This seems to me a serious mistake. There are two reasons why the amount of investment and the increase in output can be, and empirically are only loosely connected: first the form and distribution of investment are at least as important as its sheer magnitude. Second, what we call capital investment is only part of the total expenditure on increasing the productivity of the economy.

"The first reason needs little additional comment. The second is perhaps less clear. In any economy, the major source of productivity is not machinery, equipment, buildings, and other physical capital; it is the productive capacity of the human beings who compose the society. Yet what we call investment refers only to investment on physical capital; expenditure that improve the productive capacity of human beings are generally left entirely out of account".

Galbraith (1958) expressed this same idea much more pungently when he lamented that education was considered unproductive but the provision of a toilet to a school lavatory was productive!

You might wonder where forestry fits into all this. I have quoted these economists to show that purely mechanical criteria are not sufficient for deciding investment priorities. Merely because other criteria are difficult to assess it does not make them less important. The Indian 5-year Plans have been criticised on these grounds also. Two Indian economists—Vakil and Brahmanad (1956) posed these questions: "Is the industrial structure that is emerging, evenly balanced both in relation to each of the industries included within the structure and also in relation to the other sectors of the economy. Should not the plans of expansion take into account the needs and requirements in specific fields? Is it true to say that that industries, once they are started, find that the problems of adequate markets and of effective utilization of their capacity are automatically solved? These are somewhat inconvenient questions, but no one can deny that they are important". Investment in forest industries is desirable whether looked at from the point of view of using an existing raw material as an industrial base or from the point of view of substituting for imported manufactured products.

Thus the type of economic development proposed should fit naturally into the social and economic circumstances of the country—

it is only then that development can take place on all fronts. Merely transporting western factories into developing countries will not solve the problems of economic growth. Excessive investment in one or two fields to the neglect of others will only produce a lopsided result. One example of this can be seen in the oil rich Arab States where Cadillacs and Camels are parallel means of transportation.

In developing countries a high proportion of the work force is engaged in Agriculture and this cannot be changed overnight. Planned national development should therefore start with this base and from there extend to other fields. In fact it seems necessary at different stages of economic growth to concentrate on different areas of investment. According to Rostow (1960) at the early stage of growth it is necessary to "apply quick-yielding changes in productivity to the most accessible and naturally productive resources . . . In a generalised sense, modernization takes a lot of working capital and a good part of this working capital must come from rapid increases in output achieved by higher productivity in agriculture and the extractive industries", and again: "It is therefore an essential condition for a successful transition that investment be increased and—even more important—that the hitherto unexploited backlog of innovations be brought to bear on a society's land and other natural resources, where quick increases in output are possible". Since a large proportion of the work force is engaged in Agriculture (in which I include forestry), small increases in productivity will have a substantial total value. The amount of investment necessary to achieve increases in productivity need not be high—this is specially true of forestry.

Developing countries usually want a return for the money invested to be obtained very fast. In forestry, investment can be in several different fields. It may be in the field of timber extraction where provision of mechanised methods or new roads will increase output or in the field of timber utilization, where the installation of seasoning kilns and pressure impregnation plant will help to use timber which might otherwise not be used, or the installation of more efficient sawing machinery which will help to reduce waste in sawing, or in the field of reforestation. Finally investment could be on factories whose raw material requirements are supplied by forestry say, Paper mills, Plywood, Hardboard or Particle board factories and so on. In none of these things is the gestation period unduly long except perhaps in reforestation. Even here, an acquaintance with the economics of plantation forestry in Ceylon will show that in some cases returns can be expected after as little as seven years—this first return may even repay the cost of planting. Also in the case of fast growing species such as Lunumidella (*Melia dubia*) *Albizzia moluccana*, or *Eucalyptus* the final crop might be harvested

even before 10 years. The idea that forest plantations only pay for themselves after very long periods of time is a misconception that must be abandoned.

Considering all the criteria I have mentioned above, it will be seen that investment in forestry is desirable in developing countries. The investment should be in the direction of firsty exploiting existing resources and secondly in making the resource more productive.

A Forestry Survey of Developing Countries

Before I discuss the possible lines of forestry investment in developing countries it is necessary to take a quick look at the position of forestry in these countries. The Table below gives some basic statistics of a selection of developing countries of America, Africa and Asia together with similar statistics for some developed countries.

Country	*Total Forest Area '000 ha.*	*Forests in use*	*Total Forest as % of land*	*Removal of industrial wood per ha. of total Forest*
Fed. Rep. of Germany	7,038	7,038	29	3·5
Denmark ..	444	436	10	3·2
Switzerland ..	976	897	24	3·0
Japan ..	23,040	12,411	63	2·1
Finland ..	21,874	21,874	72	1·7
Phillipines ..	13,173	5,001	44	0·5
Uganda ..	1,607	117	77	0·47
Br. N. Borneo ..	6,306	355	83	0·44
Malaya ..	9,518	220	72	0·24
Pakistan ..	2,552	2,469	3	0·20
Ghana ..	14,103	1,724	59	0·13
Br. Honduras ..	1,813	1,378	80	0·06
India ..	72,752	37,273	22	0·06
Indonesia ..	90,177	61,268	61	0·06
Burma ..	45,274	24,087	67	0·04
Ceylon ..	3,546	1,045	56	0·04
Nigeria ..	31,960	1,078	33	0·04
Br. Guiana ..	18,130	260	84	0·01
Tanganyika ..	51,282	27,195	58	0·01

Figures computed from "Year Book of Forest Products Statistics" (1963) and "World Forest Inventory" (1958) both published by F.A.O. Rome.

Much variation can be seen in the basic picture that is shown in the Table. For instance, variation in size ranging from Indonesia with 90,177,000 hectares of forest land to Uganda with 1,607,000 hectares; also in the forest area as a percentage of total land area, ranging from British Guiana with 84% to Pakistan with 3%. However in one feature there is not much variation and this is in the depressingly low production per unit area, computed on the basis of total forest area, as compared with the developed countries. Just to take an example, the production of industrial wood per hectare in West Germany was 350 times the production in Tanganyika or British Guiana. There could be several reasons for these low production figures but it is rather curious that the developing countries most of whom enjoy climatic and soil conditions which are very favourable for forest growth should exhibit this common factor of low production. It seems to suggest that all the different reasons for low production can be reduced to a single basic cause. The cause is of course, insufficient investment in forestry and by investment I mean not only capital investment in machinery or plantations but also investment in research and education.

This, to an economist will seem a truism. He might contend, with some truth, that if any field of economic activity in developing countries is selected, insufficient investment will be found to be a basic problem. However, not all forms of investment are equally attractive; forestry is a unique field of investment in that firstly it is a naturally occurring resource only awaiting exploitation, secondly because planned exploitation does not reduce it, but will in fact, increase its productivity, thirdly because it provides a basic raw material on which other industries are built, and fourthly because the gap between potential and actual productivity is very wide.

It must be remembered that except for one or two countries such as Britain, high productivity in Agriculture and forestry preceded economic growth. This is not entirely a coincidence. In Sweden and Finland economic development was based on forestry and its allied industries. In the early years of the U.S.S.R. foreign exchange was earned by vigorous exploitation of its vast forest resources and sale on the world market. To such an extent was this done that world prices were depressed and the U.S.S.R. was accused of 'dumping'.

Many of the developing countries also have rich forest resources which can be developed in this way. It may perhaps surprise many of you to know that the growing stock of timber in Ceylon's forests has been estimated at about 4,500,000,000 cubic feet ! Of course, included in this figure are species which are not used today and also species like Wira (*Drypetes sepiaria*) which cannot be used

in its natural form in the foreseeable future. Yet it gives an idea of the vast potential of forestry which only awaits planned exploitation and development.

Lines of Investment in Forestry

I shall now deal with the type of investment that can be made in forestry in developing countries.

(1) Rendering more forest areas accessible. I have placed this first because in Ceylon at least this will be the quickest yielding investment. Also this is a type of investment that can be labour intensive or capital intensive at will. If we consider road building for instance, it can be done using labourers with pickaxes and shovels or using bulldozers and other heavy equipment. The other circumstances in the country will determine the method to be used. In Ceylon the richest forests are found in the Wet Zone in the South Western part of the island. The type of country in which the forest occurs is very rugged and quite a proportion of it is not served by roads. Thus, some of these forests are, as near as may be, virgin forests awaiting exploitation. Most of these areas have been mapped out and typed using aerial survey methods and the location of rich forests is known. One such Compartment in Kanneliya Reserve of 267 acres was enumerated last year prior to felling and was found to contain 250,000 cubic feet of timber in the form of trees over 5 ft. in girth. Exploitation of this Compartment only became possible because a forest track 9 miles long had already been cut over the previous few years. If the exploitable timber in this one compartment is valued at only Re. 1/- per cubic foot, the production from this Compartment alone would pay for the whole road ! Of course, the road has in the course of its 9 mile advance into the forest, already paid for itself several times over.

I quote this example to show how under certain circumstances road building into forests can be extremely profitable—it may not apply to other countries where the growing stock per acre or the level of utilization is much lower than in Ceylon.

(2) Mechanisation of logging. This is a field where Ceylon is particularly backward due to historical reasons. In the Western countries a high proportion of the forests are privately owned and have supplied timber for industrial use for several centuries. Hence investment in forestry

grew together with production. In most of the tropical countries timber exploitation was on the concession system where private firms assured of long-tenure leases invested heavily. In Ceylon, most of the forests are State owned and have never been exploited on the concession system. This may be a good thing because this system tends to promote local over-exploitation but one of the disadvantages was that investment becomes the responsibility of the State. In the last few years Government has invested in road building, impregnation plant, seasoning kilns and modern saw milling machinery but the field of logging and extraction is difficult for the Government to enter and has thus been relatively neglected.

At the present time, with very few exceptions, private contractors use only elephants or sling carts for extraction of logs from the forest. The daily output of elephants and sling carts is very small compared to that of say crawler tractors with logging arches or winches, and since the availability of elephants is not unlimited haulage costs tend to rise. This tends not only to restrict the radius of economic exploitation from a consuming centre and thus leave productive forests unexploited but also to prevent the speedy exploitation of forest areas released for development. I will not say anything more about this because such difficulties only arise in countries where forest policies and organisation do not encourage investment in this field and this is a problem which we will have to solve ourselves.

(3) The third field of investment is in ways of rendering utilisable, species of timber which have not so far been used. I would like to stress this because this is one problem which is common to all developing countries. I have in an earlier paper (Tisseverasinghe 1963) tried to show that this is a field in which most developing countries can increase production per acre without a great deal of investment and effort. Very few species of timber need be considered as completely useless, yet due to traditional and other reasons many species remain unused in the developing countries. To take a few examples: in British Guiana in 1959 two-thirds of the total timber production comprised of a single species—Greenheart (*Ocotea rodiaei*), and nearly one-fourth by 5 other species. In Uganda according to Swabey (1954) of 200 species growing to timber size, 92% of the production from Reserved Forests came from 20 species, five of these species contributing 70% of total production. In British Honduras 4 species

constituted 90% of the total production. The extent to which production can be increased by better utilization is seen from the statement of Dawkins (1954) that in Budongo (in Uganda), sustained yield can be increased by 400% if only four additional species viz. *Celtis*, *Chrysophyllum*, *Alstonia* and *Cynometra* were saleable. Species which are not in general use can be conveniently classified into three groups:

(*a*) Timbers which are inherently durable but are not used due to lack of knowledge or their properties.

(*b*) Timbers which are not inherently durable under certain conditions but whose durability can be improved by proper seasoning or preservative treatment.

(*c*) Timbers which are not suitable for use in their natural form but which will have to be converted into some other product before they can be used e.g. as plywood, particle board, hard board or pulp.

One might imagine that timbers in the first class i.e. those which are inherently durable, will present no problems in an effort to make them used. In actual fact the solution is not at all easy. Firstly, long experience or research is needed before one can definitely specify the purposes for which a timber is suitable or to what extent it is inherently durable. This does not suffice—it is also necessary to convince the consumer of timber that particular timbers which had not been used in the past are in fact, suitable for various purposes. Whatever the difficulties of this are, yet it cannot even be attempted unless research on timber utilization is done. Hence this is an essential part of the forestry programme in any developing country.

The second class of unused timbers are those which are not inherently durable *under certain conditions*. I stress this qualification because durability is such a relative term. There are many timbers which are not durable under extreme conditions such as constant exposure to sun and rain or in contact with the ground but which are quite durable when used indoors. Certain others are not used because if used unseasoned they will warp and shrink even when used indoors.

Even in countries with advanced technologies the ordinary consumer of timber will not normally wish to be burdened with the problem of having to season it before use. The problem will naturally be more difficult in developing

countries. Again, in the case of timbers which cannot be used unless they are treated with preservatives it would not normally be possible to pass this burden on to the consumer. The reason for this is that the only effective and practicable method of preservative treatment is pressure impregnation and this treatment needs expensive plant which is not usually accessible to a consumer.

In these cases therefore, it is incumbent on the seller of timber to sell seasoned or pressure impregnated timber. To do this investment in seasoning kilns and pressure impregnation plant is necessary. I had earlier examined (Tisseverasinghe, 1963) the economics of pressure impregnation and shown that the capital/output ratio of the impregnation plant operated by the Forest Department was 1·43 as compared with the average capital output ratio for the various industrial enterprises in Ceylon of 3·5. The ratio of 1·43 was derived by calculating output as the value added to the raw timber by the treatment process. If the value of the final product based on selling price is considered as output, the capital output ratio would come down to 0·5. The gestation period of this investment is also very short. On all these counts therefore this type of investment is eminently desirable.

Industries based on forestry products

The third class of timbers which remain unused comprise those which cannot be used in their natural state and which have to be converted into other products before they can be used.

The first industry which I mentioned was Plywood. One of the disadvantages of timber in its natural state is the fact that its strength along the grain is several times its strength at right angles to the grain. As you will be aware, this difficulty is overcome in plywood by combining veneers in such a way that the grain in one veneer is at right angles to the grain in the succeeding veneer. Also, timbers with an unattractive appearance can be used as core veneer while figured words can be used as face veneer. Many species which may not be used in their natural form can be used as Plywood. Even if no Plywood is manufactured, a veneering machine is always a useful investment because many of the developing countries export veneer quality logs which are processed by other countries. Veneering these logs in the country of origin and exporting the veneer will definitely be more profitable. Ceylon exports, for instance, a certain quantity of Satin logs which are used in veneer form by the importing country. Timber is a bulky commodity and hence freight charges account for a high proportion of the cost

of the timber. If instead of transporting logs with the attendant waste, only the veneer were transported a much better price could be obtained for the product. Further, the work involved in veneering will be done in the country of origin and not in the importing country.

Ceylon imports a large number of Plywood Tea Chests in which export tea is packed. The Ceylon Plywood Factory at present produces only 20% of the annual requirement of Tea Chests. However, plans are now under consideration to increase production at the factory and it is hoped to produce all the requirements of the country without building new factories. This will mean that the factory will be using 5 times as much timber as it is doing at present. This order of increase of production of timber from the forests can only be achieved if more money is invested in roads, mechanised logging equipment and so on. The investment will be worthwhile because of the ready made market that exists.

The Manufacture of Particle Board is only in its nascent stage in developing countries. Koelmeyer (1959) in his Presidential address before this Section gave figures to show that with a capital investment of 3 million rupees of Chipboard Factory producing 10 million square feet per year could be established. At a price of -/50 cts. per square foot this represents an annual production in value of Rs. 5,000,000—thus making the capital/output ratio a little more than 0.5. Under present day conditions this figure may have to be revised but even if the capital cost goes up to 5 million rupees the capital/output ratio will still be only 1. Particle Board manufacture usually started off in Western countries—as an ancillary to saw mills and other wood using industries in order to make use of waste wood. Today it is an industry in its own right competing with other industries for timber. Particle Board is not very choosey about species of timber and is hence particularly suitable for tropical countries which have a large number of species in their forests—many of which are not used at all. Particle board has certain advantages over timber in its natural form. For one thing it does not warp or shrink and for another its strength is uniform in all directions. It can be given a very pleasing appearance by the addition of a veneer surface.

Fibre board manufacture is the next stage in forest industries development. The technology is more complicated than in particle board since timber has to be reduced to its fibres instead of being cut up into chips, however, a much bigger variety of products can be manufactured to any required density. The consumption of fibre board which includes soft board and hard board is still rising very rapidly and although fears of possible market saturation were expressed at the early stages of the industry, that level has still

not been reached. In Ceylon there have been suggestions that hard board be manufactured from Rubber wood and a hard board corporation has been set up, but so far work has not started on it. Although rubber wood is not strictly a forest product it deserves a lot of attention because it is a comparatively uniform raw material available in known quantities generally in areas well served by roads.

Paper pulp manufacture using mixed tropical woods has not yet been shown to be a feasible proposition. Pulp mills in the temperate countries depend mostly on conifers for pulp. Recently, the timber of broad leaved species has been used to an increasing extent but yet no mill has used arbitrarily mixed species.

In the newer countries the tendency has been to raise forest plantations for pulping. Australia, South Africa and Brazil to name three widely scattered countries produce large quantities of pulp from species of *Eucalyptus*. In India and Pakistan, bamboo and grasses are used, and in Ceylon, paddy straw. Feasibility reports on the use of Rubber wood and *Eucalyptus grandis* (wrongly termed *Eucalyptus saligna* in the past) have been prepared by the Eastern Paper Mills Corporation and it is likely that a newsprint mill will be set up soon. Unfortunately Eucalyptus has a relatively short fibre in the region of 1 m.m. and will thus have to be mixed with coniferous pulp, which of course will have to continue to be imported. However, the future need not be bleak—experiments are under way to grow various species of Pine in Ceylon and if this is successful, Pine plantations may contribute the coniferous pulp required. Bamboo can also substitute to some extent for long fibred pulp. Studies in India by Lundquist (1964) suggest that *Pinus insignis*, *Pinus insularis* and *Pinus merkusii* can be successfully grown and *Pinus insularis* has a very long fibre—more than 3 m.m. Past experiment with Pine planting in the Up-Country in Ceylon have not been very successful—perhaps due to lack of suitable mycorrhiza and perhaps because planting was confined to high elevations. Continuous experiment by the Forest Department is going on in this connection.

One of the difficulties facing developing countries in their effort to start forest industries is that machinery has been designed in Western countries to make use of processes based on raw materials available in those countries. Thus paper mills are generally designed for very large scale operations using coniferous wood as raw material. Such large scale operation does not suit developing countries and in any case the coniferous raw material is not always available in quantity. Special processes and machinery may have to be designed to make use of local raw materials and to be able

to do this the industry must make a start, learn from mistakes and then perhaps improve the process. I will mention two examples of how new processes have helped developing countries. The first is the attrition process for pulping hardwoods which is now being used in many parts of the world. A new mill has been established in Pakistan based on a local hardwood. Similarly dissolving pulps for the artificial fibre industry used to be made only from Spruce-type timbers. The recently developed "prehydrolysis sulphate" process has made it possible to use a wide range of other raw materials. A mill producing artificial fibre is under erection in India using Bamboo as raw material. (Greitz, 1963).

The next stage of forest industries development is the use of the Cellulose in timber as artificial fibre. This has not even been attempted in Ceylon and hence it may appear to be a strange idea. However, timber is a very important fibre producer. It may surprise you to learn that a mature forest contains as much fibre as six hundred cotton crops can produce from the same acreage. Looked at in another way, an acre or forest will produce annually as much fibre as 5 acres of good cotton land. (Glesinger, 1950). In countries like ours which have to spend large sums of money annually on the import of textiles, there is much scope for the establishment of an artificial fibre industry. At present private industrialists have started a rayon weaving industry using imported fibre. The time may not be far off when the fibre itself is produced within the country.

Plantation Forestry

I have not given this subject its customarily important place because I feel that in underdeveloped countries the stress should be on using what is available rather than on replacing natural forest with plantations. Of course, planting must go together with exploitation in cases where the net return from natural forests is low or if industries are established which require special, uniform varieties of timber. The subject of plantations is usually quite adequately covered in the literature of forestry and this was another reason why I did not expand on it. There has been a tendency in some developing countries to embark on ambitious planting programmes purely on the ground that the percentage net yield was higher than the rate at which money could be borrowed. I think I have said enough to show that this narrow view of forest economics is unsuitable for developing countries. In such countries, when planting programmes are organised stress should also be placed on planting high increment, short rotation species. In Ceylon we have recently increased our softwood planting programme to cope with the growing demand for match wood, packing case wood so on.

Minor Forest Produce

This is an aspect of forestry which has tended to be neglected. Some countries like India set much store by their minor forest produce such as canes, bamboos, grasses, medicinal plants, seeds, herbs and so on. These items contribute substantial sums to forest revenue in India. In Ceylon however, this subject was neglected in the past and we have been content to import the various kinds of produce. This was taken to such an extent that in one year even cow dung was imported! Even at the present time Aralu (*Terminalia chebula*) and Bulu (*Terminalia belerica*) seeds are imported although these occur in abundance in this country and in fact were once exported from Ceylon. I do not have the time to expand on this subject but I am glad to say that we have now woken from our apathy and are now making every effort not only to collect all our requirements but even to build up an export market.

We are now organising the collection of Beedi leaves (*Diospyros melanoxylon*) locally and are making investigations on Gammalu (*Pterocarpus marsupium*) resin and Weniwelgeta (*Coscinium fenestratum*) for which there is an export market. The male bamboo (*Dendrocalamus strictus*) is now being propagated with a view to using it for paper pulp in the future.

It has been possible within the time at my disposal only to touch on some of the important fields of forestry investment, but I hope this will suffice to show you to what extent forestry can stimulate economic development. I must warn you that because foresters are a relatively neglected community, we tend to make exaggerated statements about the value of forestry. In this particular instance however, being faced with an audience comprising the Scientific elite of the country, I have studiously avoided exaggeration—if I have erred in any way, it is in the direction of conservatism.

References

Andrews. (1960). *A Forest Inventory of Ceylon.* Ceylon Government Press.

Bauer. (1961). *Indian Economic Policy and Development.* p. 152, London: George, Allen & Unwin Ltd.

Dawkins. (1954). The Productivity of Uganda High Forest. Uganda Forest Department. Tech. Note No. 3 - p. 3.

Galbraith. (1958). *The Affluent Society,* Penguin books. p. 298.

Glesinger. (1950). *The coming age of wood.* London: pp. 15 and 279.

Greitz. (1963). *Production of Pulp and Paper. Unasylva,* **17** 69.

Koelmeyer. (1960). *Proc. Ceylon Ass. Advanc. Sci.* **2**, 77.

Lundquist. (1964). Draft Report to the Government of India on the Expansion of Plantation Forestry in India. F.A.O. Rome.

MacGregor. (1959). Some Economic Problems of Forestry in Western Nigeria (Unpublished) pp. 11 and 184.

Reddaway. (1962). *The Development of the Indian Economy*. London. p. 216.

Rostow. (1960). *The Stages of Economic Growth*. pp. 12 and 179. Cambridge University Press.

Senanayake. (1964). Address given at a meeting of technical personnel on the subject of "Co-ordination of a research programme (offered by CISIR) as a base for industrial development. (Lin ited circulation).

Swabey. (1954). Forest Productivity in Uganda. Forest Department Uganda Tech. Note No. 9/54. p. 9.

Tisseverasinghe, A.E.K. (1963). The contribution of improved forest utilisation to productivity in developing countries. *Ceylon For.* 6.

Vakil and Brahmanand (1956). *Planning for an Expanding Economy*. Bombay: pp. 30 and 404.

Von Monroy. (1960). Report to the Government of India on integration of forests and forest industries. F.A.O./ETAP. Report No. 16, p. 205.

SECTION C—ENGINEERING

Presidential Address

24th September, 1964

HIGHWAY RESEARCH AND ITS APPLICATION FOR CONSTRUCTION AND MAINTENANCE OF ROADS IN CEYLON

by

A. A. D. O. P. Saparamado

In any scientific field research has been a major factor in achieving progress. The methods used in engineering and, in fact, the entire industrial field, are constantly being improved through the development of new materials, new ideas, new procedures and new and better ways of doing things. In recent years, the great demand for highway transportation and the large sums that had been required for highway construction have resulted in a greater emphasis upon highway research and development. Highway research has been directed to find ways and means of economic and efficient use of man-power, money, materials and the development and utilisation of new techniques.

The rapid industrial and agricultural development of this country and the expansion of administrative, educational and other national services that follow, warrants the need for a highway transportation system that is adequate, efficient, economical and safe. This is a serious challenge to the highway engineer specially with the limited funds at his disposal. It is only with the contributions made through highway research that the engineer is able to solve the many problems he is faced with.

In highway research there are three basic fields of study, namely Planning Research, Physical Research and Operational Research.

Planning Research is directed to evaluate highway needs and benefits with a view to laying down priorities for improvement so that the investments made may give the maximum returns.

Physical Research is concerned with the study of materials and methods of construction, in order to develop techniques, to suit conditions prevailing in each country, with a view to making the most advantageous use of materials available locally.

Operational Research is directed to study problems arising out of actual traffic operation, such as congestion, accidents, delays etc. with a view to devising a system for the convenient, economic and safe movement of vehicles and pedestrians.

Research has considerable contributions to make in all these fields of study. This paper, however, is concerned particularly with some of the contributions made in the field of Physical Research which has an application on our immediate problems. It is proposed to illustrate this by some of the examples taken from the work carried out at the P.W.D. Research Institute. I have to acknowledge with thanks the permission given by Mr. H. R. Premaratne, D.P.W. to present this work.

It should be recognised that every country has conditions of climate, geological origin, natural resources, industrial development, finances, traffic etc. that are different from those of any other country. Therefore, the results of research carried out in one country have to be modified (before application in that country) to suit the different conditions in another country. In some cases these may not be applicable at all. Therefore, additional research, both in the laboratory and in the field, is required to determine the modifications needed for each set of local conditions.

Components of a Highway Pavement

A highway is an engineering structure constructed to carry a moving load without deformation or failure. In highway engineering terminology such a structure is known as a highway pavement. Pavements may be constructed with concrete or with earth, gravel and metal with a bituminous surfacing. The former is called a rigid pavement while the latter is called a flexible pavement which is the type used in this country.

A flexible pavement consists of three components, the subgrade, the base course and the surface or the wearing course.

The subgrade is that layer immediately below the base course and is normally constructed with the natural soil, generally by compacting a 6″-9″ layer of soil, so that the soil will attain the maximum resistance it is economically capable of offering. The more stable the subgrade, the thinner the structure required over it and, as such, the cheaper the cost of the pavement.

The base course is that part of the pavement which carries the applied load. The primary function of the base course is to distribute the wheel load over such an area of the subgrade that the bearing capacity of the subgrade is not exceeded.

The surface or the wearing course is placed over the base course and its principal function is to provide a smooth riding surface free from dust, and to protect the material below from the effects of such destructive forces as abrasion due to traffic. The wearing course usually consists of some combination of aggregates with a bituminous binder.

The Study of Soils

Soil is the most important material for the road engineer. All highway pavements are supported on soil foundations. Soil is also the cheapest material available at construction sites that can be used for construction purposes. Therefore, one of the basic needs for any country in the field of highway development is to build up a scientific knowledge of the nature and engineering properties of soils and also to ascertain the occurrence of various types of soils so that the most economic use may be made of these local resources.

Although soil is one of the oldest construction materials, the study of the engineering concept of soils for highway construction has occurred only within the past 25 or 30 years. By the early road builder, the study of soils for highway purposes was not considered important as the traffic was very light and the intensity low. In recent years the increased loads and intensity of traffic on existing roads adversely affected the soil foundation, as a result of which many failures occurred on the subgrade. It was, therefore, found necessary to determine the properties of soil as related to the behaviour under pavements and also to find ways and means of classifying them into groups according to their engineering characteristics.

One of the main contributions that had been made in the study of soils is the development of a method of identifying and classifying soils for highway purposes. In soil classification, soils of like characteristics are grouped together according to their properties and in respect of their uses. The performance of one soil could be predicted from the behaviour of others in the same group. There are various systems of soils classification, of which the one commonly used is the 'Casagrande' system of soil classification. This system is followed in countries like India, Malaya and Japan.

The 'Casagrande' system, originally developed in 1942, takes into consideration almost all types of soils. There are 21 different groups in this system based on the physical characteristics, such as gradation, liquid and plastic limits which are very easily determined in the laboratory.

The Research Institute of the P.W.D. has carried out some work with a view to classify the soils found in Ceylon. As a preliminary step it was decided to adopt a simple form of classification on the lines of 'Casagrande' system but limiting the soils to a lesser number of groups. Statistics of soil data were collected from the results of routine work carried out in the laboratory. These soils were grouped into the following categories:

G—1	=	Gravel with little fines.
G—2	=	Gravel with a high percentage of fines.
S—1	=	Non-plastic sand.
S—2	=	Clayey sand (P.I. less than 15%).
S—3	=	Sand with a high percentage of fines (P.I. greater than 15%).
C—1	=	Clays (more than 50% passing the No. 200 mesh sieve and P.I. greater than 15%).
C—2	=	Clays (more than 50% passing the No. 200 mesh sieve and P.I. greater than 25%).

Altogether 202 samples of soils had been analysed.

Although the number of soils tested is not sufficient to prepare any type of soils map for the island, they still represent a cross-section of the soils in the country and give a general idea of the types of soils which the road engineer will encounter. It is seen from these results that most soils fall into S—1, S—2 and S—3 groups and are categorised as sandy soils.

The majority of S—2 soils have a P.I. less than 12. These are generally accepted as good subgrade material and could be economically improved for use in road bases carrying light to medium traffic. The majority of soils have a C.B.R. between 10 and 20 and a small percentage of soils falls within the clay groups which have low bearing values and high plastic indices.

Established procedures in making soil surveys include the use of test pits and had auger borings and these are generally being adopted in this country for highway soil survey. Such procedures are costly, tedious and somewhat slow. Rapid advances made in recent years have developed methods which are less expensive. The most significant of the newly developed techniques is the use of aerial photographs for predicting engineering soil conditions. An aerial survey of this island has been carried out by a Canadian team under the Colombo Plan Aid and aerial photographs covering every part of this island are available. The Survey Department is also equipped to take further photographs of any desired location. Therefore, the basic requirements in adopting this technique to highway soil surveys is available in this country.

An aerial photograph has three distinct applications with respect to highway soil surveys. It can be used for delineating boundaries between soils of unlike characteristics, from which a generalised engineering soils map could be prepared. It could be used to predict the engineering characteristics of soils. It could also be used for locating field sampling areas in making a soil survey for a highway project. The use of aerial photographs for making these soil surveys is a rapid and more economical method than the conventional method which will require, for the same aerial extent, several months of field work and countless borings.

The aerial photograph gives a pictorial representation of the earth surface. The amount of details obtainable from an aerial photograph depends upon the experience and training of the engineer. A working background of geomophology, geology and soil science is desirable. This may be obtained from reference to information already published. Many years of highway research on the co-relation that exists between air photo pattern and characteristics of soils has been conducted by the Purdue University, U.S.A. This has led to the development of procedures in aerial photo interpretation and the photo keys and photo guides, the preparation of which serves as valuable tools to the engineer in the interpretation of aerial photographs.

Earthwork and Soil Compaction

In the early days of road building very little attention was paid to the compaction of soils especially when preparing subgrades. There was no systematic compaction in the construction of embankments and no method of checking compaction was available. As the weight of traffic increased, the larger and heavier loads that were imposed on the pavement, when transmitted to the subgrade, caused stresses over and above the subgrade soils could carry. For example, the stress produced at a depth of 8″ below the road surface by a 2–3 ton vehicle is of the order of 9 lbs/sq. inch, whereas the stress produced at the same depth by an 8–ton vehicle is as much as 19 lbs/sq. inch. As a result, non-uniform compaction of the subgrade on existing roads took place causing either deformation of the pavement surface or complete failure of the pavement structure. This type of failure is observed in many sections of the main and trunk roads in this country.

Research has shown that one of the simplest and economical method of improving the properties of a subgrade soil is by compaction. During compaction the soil particles are brought closer to each other resulting a lower percentage of air voids and, therefore, a higher density. Increasing the state of compaction of a soil is also found to increase the soil strength, reduce the possibility of settlement and offer resistance to water absorption.

The importance of subgrade compaction has been demonstrated by one of the research studies undertaken by the Central Road Research Institute, New Delhi. It has been shown that a substantial portion of the thickness of a pavement could be provided merely by proper compaction of a corresponding depth of the subgrade or embankment soils immediately below the base course. The extra amount involved in the compaction of this soil is very small in comparison to the alternative of providing an extra thickness of an extensive base course. In their paper on "The Importance of Subgrade Compaction" Mehra and Uppal say "The saving as a result of improving the subgrade by compaction and thereby reducing even by 9″ the thickness of hard crust required by standard practice (this refers to water-bound macadam pavement) would be off the order of Rs. 31,500/- per 10 ft. wide mile".

The main factors affecting the state of compaction of a given soil are its moisture content at the time of compaction, and the amount of energy put into the soil during compaction. Investigations made by R. L. Proctor in 1939 in connection with the construction of earth dams in California have shown that a relationship exists between the density attained by a soil, and its moisture content, for a constant amount of energy of compaction. The results of compaction of a few typical soils in this country, show such density relationship curves. Referring to these curves it is seen that the density increases with increase in moisture content and after reaching a peak value it decreases with further increase in moisture content. The density and the moisture content at this peak are known as the maximum density and the optimum moisture content respectively. This relationship affords a convenient basis for defining the state of compaction of a soil in the field. Laboratory methods for determining the maximum density and the optimum moisture content based on the above concept have been developed. According to the Proctor compaction method the soil is compacted in three layers in a cylindrical mould 1/30 cubic feet volume, each layer receiving 25 blows from a 5½ lb. hammer falling 18″. This came to be known as the Proctor method of compaction.

The knowledge concerning compaction brought with it problems of specifying the degree of compaction in the field. Although a higher state of compaction is desirable, still greater energy is required to achieve this. If costs are to be kept to a minimum, it is important that the state of compaction should not be greater than what is really needed. In earlier specifications much thought was given in specifying compaction by the type of plant to be used for compaction, the maximum thickness of layers, and the number of passes of the plant to be provided. Such a specification, though having a number of advantages, has also many limitations especially to be adopted in countries where little experience is available in the use of compaction equipment and where it is difficult to determine what type of plant should suit the different types of soil. Later specifications defined the desired end-results to be achieved, such as the compacted unit weight. The most widely used method of specifying compaction is by the minimum density that should be achieved in the field. A dry density of 90-95 per cent of the maximum density obtained in the Proctor compaction test is usually specified for subgrades and 100 per cent of the maximum density for base courses. The experience of the compaction of soils in this country has shown that these values are easily attainable and are being adopted in normal specifications.

Pavement Design

The conventional method of road building is the macadam type of construction using rubble and crushed stones and no design method is employed to determine the thickness of the various layers. The same thickness of construction is used over a strong gravel subgrade as well as over a weak clayey subgrade. This will lead to an uneconomical design in the former case. The highway engineer was faced with the problem of designing pavements of adequate thickness to withstand the traffic loads without failure, at the same time avoiding over-design that would lend to excessive costs. Some reliable method to evaluate the economic thickness of construction was therefore required.

The general problem of flexible pavement design consists of two parts. The first is the thickness design of the component layers—the subgrade, the base, and the wearing surface—so that sufficient strength be build up in these layers to carry the imposed load. The other is the material design of these layers to ensure that the material used in construction is sufficiently strong enough to withstand the stresses caused.

Every highway pavement design is governed mostly by the local conditions of climate, subgrade soil types and type and intensity of traffic. There is no accurate method of analysing the stress distribution caused by a wheel load. Therefore, unlike in other branches of engineering, the problem of pavement design has been too complex to evolve a method that could be generally accepted. Numerous efforts have been made from time to time for evolving design procedures, some of which are based on theoretical considerations and some on emperical data coupled with experience. It is largely on the results of full scale experiments that the current standard of design in use have been based.

One of those design methods which is commonly used in many countries including Great Britian and India is the California Bearing Ratio method commonly abbreviated as C.B.R. It is fairly simple both in technique and equipment. It was originally developed

by G. J. Porter of the California Department of Highways and later modified by the Corp of Engineers, U.S.A. It consists of a C.B.R. test to evaluate the strength properties of the soil from which the thickness of construction would be determined from a set of design charts. The C.B.R. test consists of causing a cylindrical plunger 3 square inch area, to penetrate a sample of soil under test and measuring the load required to penetrate ·1″ into the sample. This load expressed as a percentage of a standard load viz. 3,000 lbs is known at the C.B.R. of the soil tested.

Emperical design charts to determine thickness of the different layers of the pavement for different types of loadings are available. The charts used in U.S.A. are based on vehicle wheel loads. Similar charts have been prepared in England based on traffic classified according to the number of heavy vehicles. These curves are also adopted in India.

The Research Institute, P.W.D. has been adopting these curves as a guide for evaluating thickness requirements of highway pavements.

Soil Stabilization

The use of crushed stones for road construction became more and more expensive due to the high cost of labour and longer distances of transport required. Therefore, an alternative material to crushed stones that would be cheap still having the necessary strength was considered necessary. Research has developed methods where soils, the cheapest material available, is used for construction of pavements of all types. Most soils in their natural state do not have sufficient stability to withstand the stresses caused by heavy traffic. Cohesionless sands ravel and lose their supporting power when dry. Cohesive soils such as silts and clays become very plastic when wet. Research has shown that the load bearing characteristics of a natural soil could be improved for use in highway construction. This technique is known as soil stabilisation.

Soil stabilisation has been carried out to a large extent in Western countries and since the war millions of square yards of pavements have been constructed. Soil stabilisation has a great scope in this country to cut down costs in road construction as well as to increase the pace of construction. Experience gained in other countries could be made use of, but it is to be recognised that methods and techniques of more advanced countries have to be modified to suit local conditions and needs. The works so far done in the P.W.D. Research Institute has shown encouraging results both in the laboratory and in the field.

Several methods of stabilising soils are available, of which the most important is the method known as Mechanical Soil Stabilisation. This method involves the blending of two or more soils in order to produce a stabilised soil mixture which has the desired strength and durability characteristics. Soils derive their strength from the internal friction of the coarse particles and the cohesive strength of the silt and clay fractions. It has been shown that the internal friction of a compacted mass is proportional to the area of contact of the particles. Therefore, the maximum stability is derived when the particles are so distributed to give the minimum air voids content and hence the maximum density. Fuller has shown that this condition is reached when the gradation of the mass follows a certain law given by the equation:—

$$P = 100 \frac{d}{D}$$

where P = percentage passing a sieve of diameter d.
D = maximum size of particles in the mass.

The curve represented by this equation is known as the Fuller's curve for maximum density.

Based on the above concept and from observations of experimental works in the field, limits for gradation have been recommended. In mechanical stabilisation, two or more soils having gradation outside these limits, could be mixed together in certain proportions to obtain a soil mixture having a gradation within these limits. This soil mixture would then be compacted to a maximum density to obtain a mechanically stabilised soil.

Two of the major projects where mechanical stabilisation has been carried out in this country are the Batticaloa Airport runway and the Anuradhapura Airport runway. In the construction of the Batticaloa airport runway the soil that was available within the proposed runway was mixed in equal proportions with a gravel transported from nearby. This mixture compacted to its maximum density gave a satisfactory base course over which a 2″ premix carpet was laid. In the construction of the Anuradhapura airport runway investigations revealed the presence of a soil which had already a grading closer to the limits recommended according to U.K. practice. Although the gradation fell outside the recommended limits and the plastic characteristics a little high, still considering the very limited amount of traffic using the runway, this soil was considered suitable for a stabilised base course in its natural form. A low type surface using bitumen covered with sand was used. This runway which was open to traffic in June 1961 is behaving very satisfactorily. This type of construction has eliminated the use over 5000 cubes of rubble and metal in each runway, which would have otherwise been required if conventional methods of construction were adopted. The production of this quantity of metal would also have taken a considerable length of time.

Soil stabilisation is also carried out by the use of admixtures such as lime, bitumen or cement. The use of lime as an admixture for improving the properties of soil has been practised from very early ages. In China lime has been used in rural roads for years, but very little was known about the science involved in the process. It was not until 1940s that stabilisation of roads with lime was taken up on scientific lines modified by laboratory and field tests to fulfill the variety of stabilisation requirements. The use of lime is specially recommended to improve the characteristics of clayey soils. Natural clays are plastic and lose almost all their load bearing properties under saturated conditions. Lime acts chemically with the available alumina and silica in the clay to form complex compounds of silicates and aluminates imparting strength to the clayey soil. This action is a long term one and the soil continues to gain strength for a long period of time. Lime also reduces both the plastic properties of clayey soils and the swelling characteristics. Lime stabilisation was carried out on 2 miles of the Paranthan Pooneryn Road in Jaffna District, where 2% of lime was mixed with gravel and sand to form a 4" sub base. A 2" metal layer was laid over this as a wearing course. This section of the road, completed about 3½ years ago, is behaving very satisfactorily.

Stabilisation with Bitumen

Small percentages of a bituminous material mixed with certain soils could be used to improve their load bearing characteristics. Asphalt is the most commonly used bituminous material for soil stabilisation. Asphalts are useful for stabilisation because of their cementing and water proofing properties. The cementing property is most effective in providing increased stability in sandy soils. The water proofing qualities assist in the preservation of the natural stability of cohesive soils such as silts and clays. The increased resistance to absorption of water in a clay soil is due to the coating of individual soil particles by the thin films of asphalt. These films of asphalt resist the entry of water from outside.

Experiments were carried out in the P.W.D. Research Institute to determine the effectiveness of asphalt for stabilising some of the typical soils found in this country. One of these soils is a sandy clay soil found along the coastal belts. This is a non-plastic soil 100% passing the No. 4 sieve and about 14% passing the No. 200 sieve. Two grades of asphalt, namely a rapid curing cutback and a medium curing cutback were used. The stability of specimens prepared with varying percentages of asphalt and the percentage of water absorption of these samples were determined. The stability of the samples was determined when dry as well as after soaking in water for 24 hours.

Cement is also another admixture that could be used for stabilising soils. The high strength obtained by adding cement to soil has been a factor in the use of cement for the construction of road bases. In the U.S. alone over ten million square yards of cement roads have been constructed.

Almost all types of soils that could be mixed with cement will respond to cement treatment. However, economical considerations limit the use of cement in highly plastic soils. Cement could also be used as an admixture in soils merely to modify their properties rather than to produce a hard structural material. Cement decreases the plasticity and the shrinkage characteristics of clayey soils thus imparting a greater load bearing capacity over a wide range of moisture content.

Wearing Surface

The wearing surface or the surface course is an important structural component of a flexible road pavement. The function of a wearing surface is to protect the base from the abrasive action of traffic, waterproof the base and provide a smooth and non-skid riding surface. Wearing surfaces vary from a simple bitumen painted surface to a scientifically developed bituminous carpet. The type of surface to be used on a road depends on the type and intensity of traffic using the road. The higher the traffic intensity, the higher the type of surfacing to be provided and greater the cost. Wearing surfaces are also provided on existing roads as a part of road maintenance to prolong the life of the road and protect it from early failure.

The conventional type of wearing surface that has been used in this country from the early days is the bitumen painted surface covered with sand aggregate. This is the cheapest type of surface and the lowest in quality. This surface offers very little resistance to the base material specially on roads carrying heavy traffic. In the normal macadam base the aggregate particles get rounded destroying the interlock that is vital for stability of such a base; as a result of this constant failures occur forming many pot holes.

The asphalt concrete or premix surface was introduced into this country within the last 10 years or so. Asphalt concrete surfaces are expensive to lay and cost as much as Rs. 50/- per 100 sq. ft. However, a properly designed asphalt concrete surface laid on a structurally sound base course should last from 15 to 20 years. Asphalt concrete is a mixture of metal, aggregate and bitumen. The success of an asphalt concrete carpet depends on the selection, grading and proportioning of the material used. Bitumen is one of the most important components and its presence in the correct proportion is often the criterion of success or failure of the carpet.

Premix carpets were used in the Western countries as early as 1900. Extensive research has been carried out by manufacturing firms and other organisations on this type of surfacing and the large amount of data collected has made it possible to design and lay carpets of high quality capable of carrying the heaviest traffic.

In the design of an asphalt concrete mix the engineer seeks to produce one that will be readily workable, compactable and relatively impermeable to water, and have the stability to withstand traffic stresses. The design of an asphalt concrete mix is largely a matter of selecting and proportioning materials to obtain the desired qualities in the finished construction. The overall object is to determine the economical blend and grades of aggregates and the corresponding asphalt content. There are several methods of design of asphalt mixes that in use for which criteria have been established through laboratory and field correlation. Of these methods, the Marshall method of mix design has enjoyed wide acceptance for the design and control of mixes. This is mostly due to the simplicity of the test, and the equipment used.

When premix work was first started in this country there were no facilities for design and control of mixes. Since the establishment of the Research Institute, P.W.D., it was possible to undertake scientific design and field control using the Marshall method of mix design. The earlier mixes used were open graded with no filler. Filler is the finest portion of aggregate, predominantly that portion of material passing the No. 200 mesh sieve. Filler is a very important constituent in producing a dense graded mix so that it will be less liable to be acted upon by water and less liable to interior displacement or movement. Mineral dust, limestone dust and cement are generally used as filler. This material was not obtainable at economic prices when premix was started and, as such, no filler was used in the mix. Experiments were later carried out by using a waste product from local quarries, namely quarry dust. This has about 15 to 20 per cent passing the No. 200 mesh sieve. Using this material it was possible to produce a mix that is dense and more effective, having a low percentage of voids. It was also found that the bitumen content in the mix could be reduced by 2% which itself has brought about a saving of nearly 7000 gallons of bitumen in a 2″ carpet on a mile of road 30 ft. wide.

An asphalt concrete mix using only sand, filler and bitumen is known as sand asphalt. In sand asphalt, filler is a critical factor for stability. Specifications normally recommend 10 to 20 per cent material passing the No. 200 sieve. This filler is needed to bring the void content to the minimum 3 to 5 per cent recommended. Although sand is available in plenty in this country, still the use of such a high percentage of filler is economically impracticable. It was, however, decided to lay a sand asphalt carpet designed

in the laboratory using the minimum percentage of quarry dust as filler to give sufficient stability to carry the traffic. A mix was designed but due to the low quality of filler used the voids in the mix were found to be 15%. An experimental section of this carpet was laid on the Galle Road and gave satisfactory performance. After one year cores were cut and tested for voids and it was observed that in some cases the voids were reduced to as low a value as 7%, by the action of traffic. Though this carpet has sufficient stability to take the traffic immediately it was laid, it was observed that it took about a week for it to get completely hardened. During this period kneading action of the traffic would have brought further compaction and reduced the voids. Several sections of the trunk roads and some sections within the city of Colombo, were later laid with sand-asphalt carpet and they are behaving quite satisfactorily. The cost of a sand-asphalt carpet is less than 50% of an asphalt-concrete carpet for the same distance of transport of materials.

A surface dressing or a seal coat is another type of wearing surface that has been recently introduced to this country. The process is very simple and consists of spraying a coat of bituminous binder on the existing surface, applying a layer of uniformly spread aggregates, and compacting. This is called a single seal coat. A second layer of bituminous aggregate can be applied over this, and is termed a double seal coat.

A seal coat is one of the most versatile and useful type of road surfacing. It is an essential aid in maintaining existing roads. It could be used as a temporary surface in the development of a road by stage construction. It provides an excellent sealing for the base, gives good bond and reasonably high stability. The average cost of a single coat treatment in this country is Rs. 9/- per 100 sq. feet and a double coat treatment is Rs. 15/- per 100 sq. feet. It is more advantageous to lay a seal coat in lieu of asphalt concrete on roads when the base is structurally inadequate to take up increased loads expected in the future.

Although the laying of a seal coat is a very simple process, the durability of such a surface depends on various factors such as quality and size of aggregates, quality and type of bituminous binder, the rate of application, method of application etc. Extensive research of these subjects have been carried out elsewhere and have made many contributions. Before this method was adopted in Ceylon, several experimental sections were laid in different parts of this country for studying the performance under local conditions as well as methods of construction using hand labour.

A double seal coat was carried out on a length of 3½ miles along Chilaw-Ambakandewela Road having a stabilised gravel base. This section of the road is behaving satisfactorily for the last two years. The comparative cost of construction is well below 50% of the conventional method of construction using a 2 inch metal surfacing.

A bituminous surface dressing is essentially a problem on adhesion in the presence of water and the surface ceases to exist, if the coarse aggregate gets stripped off the binder.

In a country like Ceylon where the rainfall is heavy and persistent all the year round, the adhesive characteristics are an important factor for a seal coat treatment even in an asphalt concrete carpet. This factor has not been realised until recently. Some aggregates have a greater affinity for water than for bitumen. Such aggregates called hydrophilic aggregates when coated with bitumen has a tendency to strip. Experiments carried out in the P.W.D. Laboratory on some aggregates commonly used for road work has shown that charnokites, gneiss and limestones have good resistance to stripping. Quartzite, however, shows heavy stripping properties. Rocks containing a high percentage of quartz will, therefore have the tendency to early stripping.

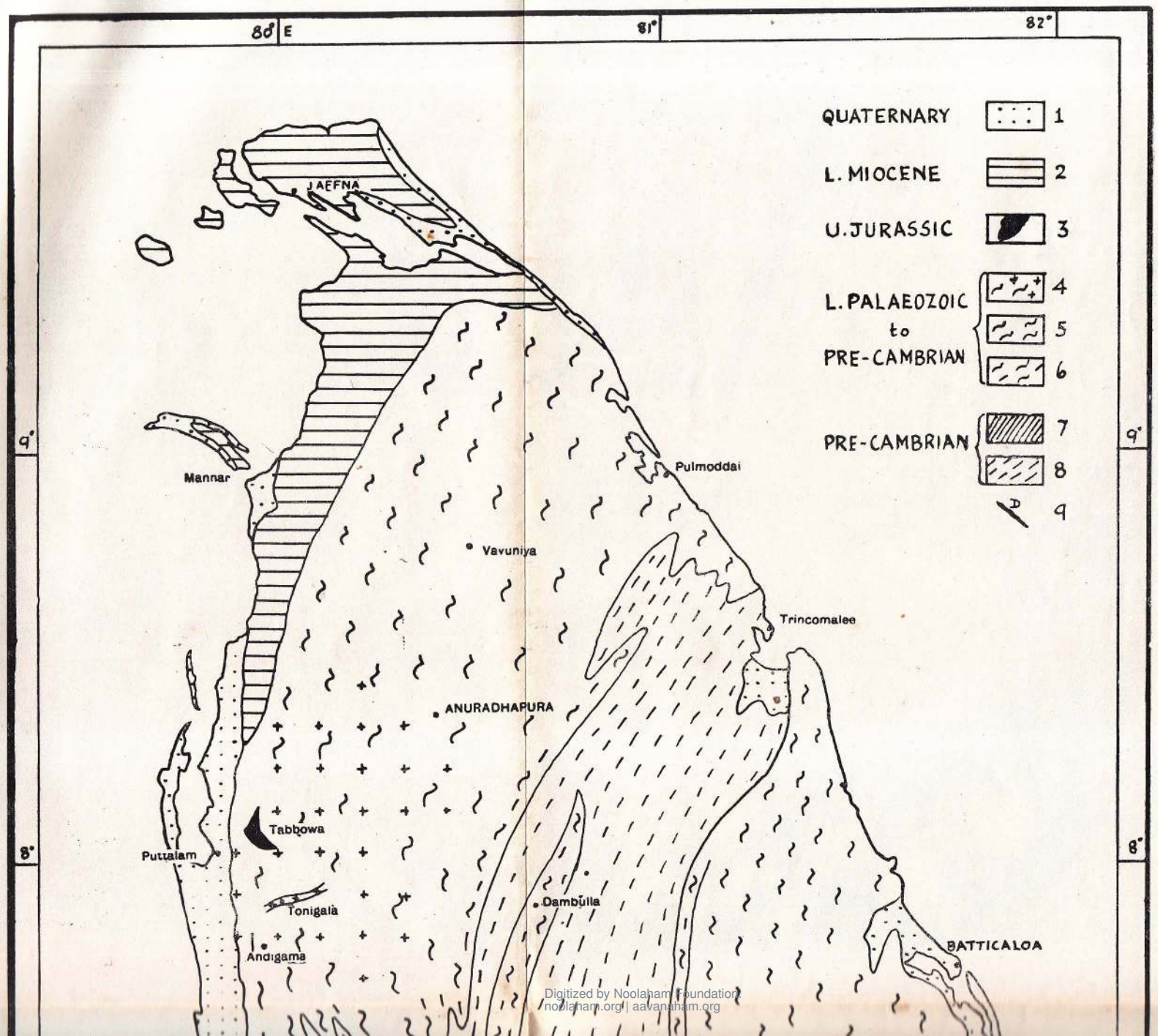
It has been shown in many experiments elsewhere both in the laboratory and in the field that one of the methods to improve resistance to stripping is by using a high viscous bitumen i.e. a harder grade of bitumen. This would present some difficulties in the initial wetting of the aggregate which is also an important factor. This would be, however, overcome by fluxing the bitumen with an oil such as kerosene. The use of 10% kerosene with the bitumen has shown very satisfactory results in seal coat work in this country.

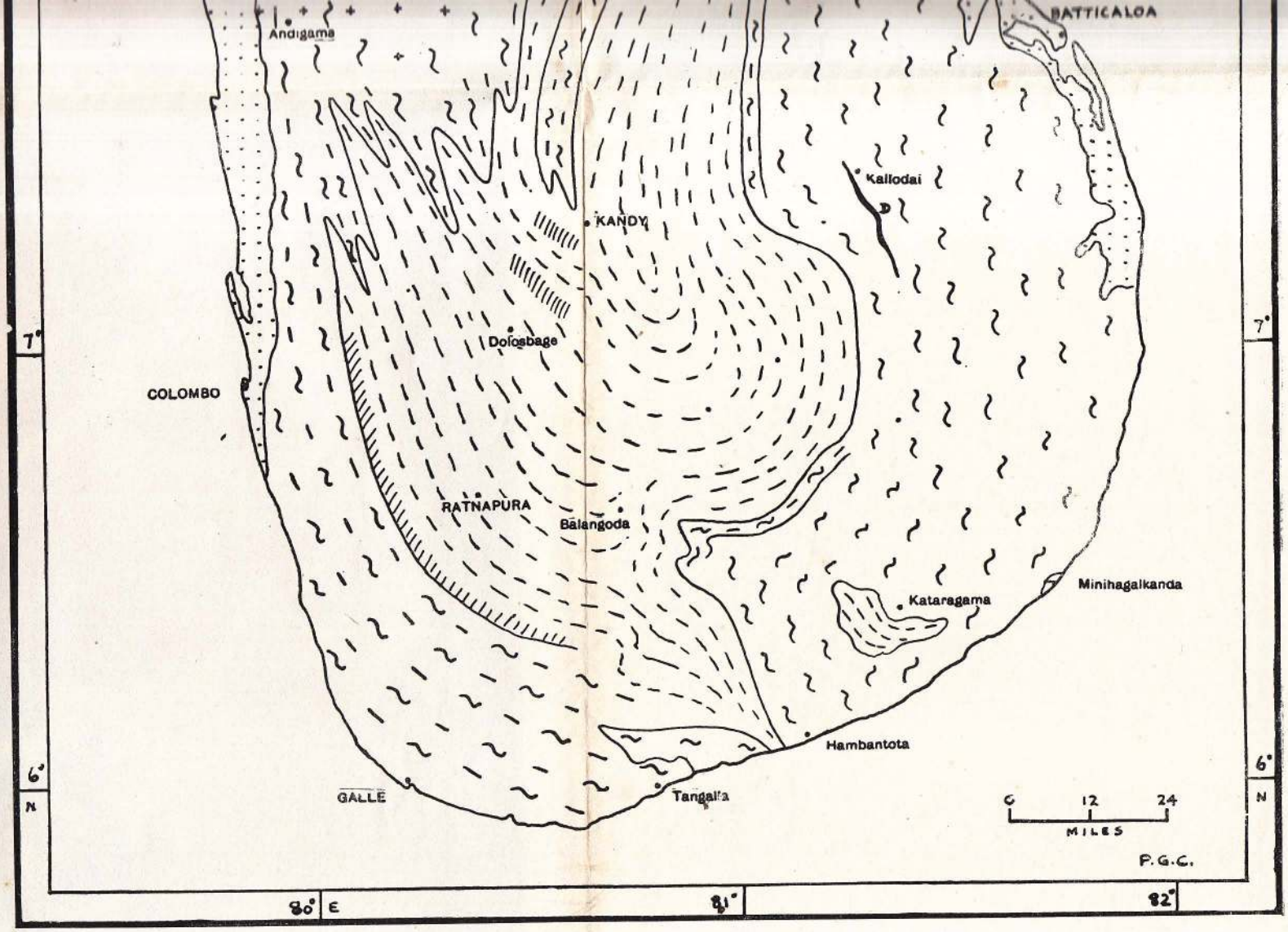
The use of alkaline powder such as lime or cement is also found effective in improving adhesive characteristics. Such powders can also be used as a filler in premix to improve adhesion. There are also certain surface active agents in the form of complex chemical compounds. These are available in the market. The addition of these agents to the bitumen even in a small percentage as low as 1% improves the adhesive properties. The addition of rubber has shown to improve the adhesive characteristics of bitumen. Rubber could be added in the form of rubber latex. Experiments conducted in the P.W.D. laboratory showed that 2% of rubber latex added to hot bitumen improved adhesive properties.

Conclusion

With the present economic conditions in the country the engineer should have more recourse to the use of less expensive material and methods of construction in the development and maintenance of roads. Traditional methods which relied on rule of thumb procedures and personal judgement should give way to more scientific methods based on the engineering properties of materials. Therefore, a scientific knowledge regarding the materials and techniques on road building in the background of local conditions should be systematically acquired and applied. The need for adequate research facilities to provide the know how for the construction of cheaper more durable, and safer roads cannot be over emphasised.

Plate I





GEOLOGICAL MAP OF CEYLON (modified from Coates, and Fernando).

1. Gravels, sands, clays and alluvium (*Older* and *Younger Groups* of N.W. Coast)
2. Limestone (*Jaffna Limestone*)
3. Arkose, feldspathic sandstone, siltstone, and mudstone (*Tabbowa Beds*)
4. Pink granites, granitic gneisses, and migmatites (*Tonigala Complex, Kathiraveli Gneisses, Kirinda Gneisses*)
5. Granitic gneisses and migmatites (*Vijayan Series*)
6. Granitic gneisses with charnockites, charnockitic rocks, and metasediments (*Dambulla-Habarana migmatite belt, Transitional Zone, South-West Region*).
7. Basic rocks: amphibolites and basic charnockites (*Kadugannawa Gneisses, Sinharajah basic zone*)
8. Metasediments and charnockites (*Highland Series, Kataragama Complex*)
9. Dolerite dykes.

SECTION D—NATURAL SCIENCES

Presidential Address

Wednesday, 23rd September, 1964

THE GEOLOGY OF CEYLON—SOME RECENT ADVANCES IN KNOWLEDGE*

by

P. G. Cooray

INTRODUCTION

Although it is customary for a Presidential Address to be concerned with the one's special field of work, I have decided to break with custom on this occasion, for the reason that my views on the charnockites of Ceylon have recently been set out elsewhere and need only brief mention here. I am, however, following a precedent, for L. J. D. Fernando, in his Presidential Address to this Section in 1948 (Fernando, 1949) discussed some aspects of the geology of Ceylon and reviewed our knowledge of the Pre-Cambrian rocks at that time. I have set myself the somewhat more ambitious task of reviewing the whole field of Ceylon geology, partly because periodic reviews of this nature are very necessary, and also because I have, of late, interested myself in other aspects of the island's geology besides the charnockites.

At the time when Fernando's review was made, the Geological Survey Department (then known as the Department of Mineralogy) had a staff of 4 geologists, 2 chemists, 3 field assistants, and very little equipment. Geological mapping had been largely of a reconnaissance nature, aimed at delimiting certain major geological boundaries which were thought to exist.

Today, sixteen years later, there are 9 geologists, 3 chemists, 1 geophysicist, 20 field assistants, 9 technical and laboratory assistants, 6 laboratories, and a large amount of equipment. Systematic geological mapping on the scale of 2 inches to 1 mile commenced in 1952, and since then 13 geological sheets comprising 5,165 square miles have been completely mapped, and 6 sheets comprising 2,625 square miles are currently being mapped (see Table I and Fig. 1). In addition, several hundreds of analyses of rocks and minerals have been carried out, several thousands of feet

*Published with the permission of the Director, Geological Survey Department, Colombo.

TABLE I

PROGRESS OF GEOLOGICAL MAPPING IN CEYLON, 1952 - 1964

Sheet	*Area (Sq. Mls.)*	*Period of Mapping*	*Mapped by*
1. Polonnaruwa*	480	1953—1957	P.W. Vitanage (1959)
2. Rangala*	480	1953—1957	P.G. Cooray (1961)
3. Panadura-Horana	490	1958—1959	P.W. Vitanage
4. Alutgama**	390	1958—1961	P.G. Cooray
5. Galgamuwa***	480	1952—1962	P.G. Cooray
6. Gampaha***	480	1961—1962	D.B. Pattiaratchi, J.W. Herath and D. J. A. C. Hapuarachchi.
7. Avissawella***	480	1962—1963	D.B. Pattiaratchi and J.W. Herath
8. Ratnapura***	480	1957—1963	J.W. Herath and D.J.A.C. Hapuarachchi.
9. Chilaw***	115	1963	P.S. Kumarapeli and L.K. Seneviratne
10. Battulu Oya***	150	1963	P.G. Cooray
11. Rakwana***	480	1958—1964	J.W. Herath and D.J.A.C. Hapuarachchi
12. Puttalam	180	1962—1964	P.G. Cooray
13. Dandagamuwa	480	1963—1964	P.S. Kumarapeli and L.K. Seneviratne
Total ..	5,165 sq. miles		
14. Elahera	480	1954—	D.B. Pattiaratchi
15. Wariyapola	480	1955—	D.B. Pattiaratchi
16. Kurunegala	480	1964—	L.J.D. Fernando and P.S. Kumarapeli
17. Ambalantota	425	1964—	D.J.A.C. Hapuarachchi
18. Ambalangoda	280	1964—	J.W. Herath and L.K. Seneviratne
19. Hatton	480	1964—	P.G. Cooray
Total ..	2,625 sq. miles		

*Maps and descriptive Memoirs published, year of publication in brackets.
**Map and Memoir being published.
***Maps being printed.
Descriptions of the areas mapped each year are to be found in the Administration Reports of the Director, Geological Survey Department (and previous to 1962, of the Government Mineralogist).

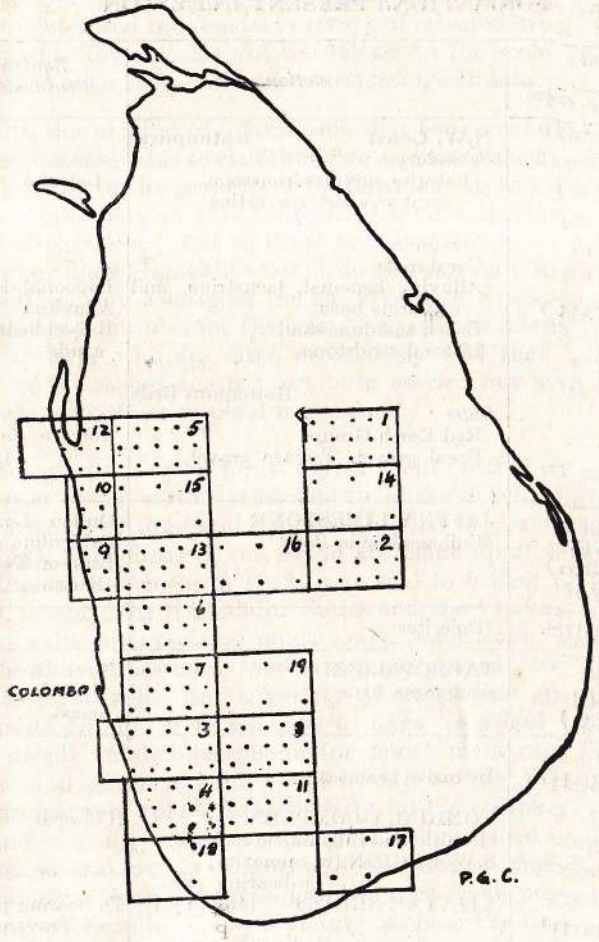


Fig. 1

Map of Ceylon showing progress of systematic geological mapping on the scale 2 inches = 1 mile, from 1952 to 1964.

(Figures refer to sheets named in Table I).
Mapping completed = close stipple.
Mapping proceeding = wide stipple.

TABLE II

FORMATIONS PRESENT IN CEYLON

Geological Period (approx. m.y. ago)	*Formation*	*Equivalent in Peninsular India*
QUATERNARY (2 m.y.)	**N.W. Coast** **Ratnapura** *Secondary :* Laterite, nodular ironstone, kankar, chert, travertine *Younger :* Coral reefs Alluvial, lagoonal, lacustrine, and estuarine beds Beach and dune sands Littoral sandstone Ratnapura Beds *Older :* Red Earth Group Basal gravel. Terrace gravel.	Laterite Lagoonal deposits Alluvium Raised beaches; dune sands Terrace deposits (Madras)
LOWER MIOCENE (Burdigalian) (20 m.y.)	JAFFNA LIMESTONE Minihagalkanda Beds	Quilon Limestone Orbiculina Shales, Lsts. of Kathiawar & Kutch
CRETACEOUS	?Dolerites	
UPPER JURASSIC (150 m.y.)	TABBOWA BEDS Andigama Beds	Upper Gondwana deposits of Madras coast
UPPER PALAEOZOIC (400 m.y.)	Intrusive granites TONIGALA COMPLEX Granitic and migmatitic rocks of S.W. REGION (w. monazite)	?Closepet Granite
LOWER PALAEOZOIC (600 m.y.)	Th-bearing pegs. VIJAYAN SERIES (485 m.y.) Zr-bearing pegs. (575 m.y.) VIJAYAN SERIES	Th-bearing pegmatites of Travancore ?Peninsular Gneisses
PRE-CAMBRIAN	HIGHLAND SERIES Kataragama Complex Kadugannawa Gneisses	Khondalites and charnockites of E. Ghats. (1650 m.y.)

have been drilled in the course of mineral surveys, engineering geology, and water-supply investigations, air-borne and ground magnetometer and radiometric surveys of selected areas have been made, and a Tectonic Map of the Island on the scale of 1 inch to 4 miles has been compiled from photogeological data.

With this amount of information that has accumulated, it does not seem unreasonable to claim that we now have a somewhat better understanding of the geology of Ceylon than we had 15 years ago. It is my intention, in this address, to review the main advances gained in that time, and to draw attention to those fields where gaps in our knowledge still exist. I do this not only for the benefit of myself and my colleagues (on the results of whose work I have freely drawn), but also for the sake of my fellow scientists and the public at large who are often erroneously told that a geological survey of the island has not yet been carried out and that very little is known of our mineral resources.

Geologically, Ceylon forms only a small portion of an ancient, stable part of the earth's crust known as the South Indian Shield, most of which is Pre-Cambrian in age (that is, over 600 million years old). Four-fifths of the island are made up of Pre-Cambrian crystalline rocks which, in 1948, were said to belong to two major groups, namely, the Khondalite Series and the Vijayan Series. The original sedimentary clays, sandy clays, sandstones, shales, mudstones, and limestones—now metamorphosed to quartzites, khondalites, marbles, and a variety of schists and gneisses of the Khondalite Group—were thought to have been laid down on an older, deeply eroded basement (or floor) made up of granitic gneisses and granites known collectively as the Vijayan Series; both groups were folded subsequently into a complex pattern of folds known as a synclinorium. Into this much-folded complex were intruded several groups of igneous rocks such as charnockites, the Kadugannawa Gneisses, zircon granites, rare-earth pegmatites, and the Tonigala Granite. Today's picture of these Pre-Cambrian rocks of Ceylon is, as I hope to show, rather different.

The rest of the island is made up of a few small patches of sediments of Jurassic age, an expanse of sedimentary limestone of Miocene age mainly in the extreme north, and a belt of unconsolidated deposits of Quaternary age along the entire coastal belt (Plate 1). Here there are no radical changes in the overall picture, but we do know more about some of these formations in detail.

A table of the formations known to be present in Ceylon is shown in Table II.

THE PRE-CAMBRIAN

Subdivision of the Ceylon Pre-Cambrian

The fundamental concept in modern stratigraphic studies of the Pre-Cambrian is that individual regional units with characteristic geotectonics can be fitted, in time and space, with the aid of radiometric data, to phases of one or more 'geological cycles' (Harpum, 1960). We have only just begun to apply this concept to the Pre-Cambrian of Ceylon, and for a start have recognised that the two major series can, in fact, be subdivided into a number of smaller lithologic-metamorphic units, each characterised by a particular type and grade of plutonism and tectonic style (see Cooray, 1961*b*). The relations of some of these units to each other are partly known. We know, for example, that the Highland Series (see below) and the Kataragama Complex belong to the main orogenic phase of the Taprobanian cycle, and that the Habarana-Dambulla migmatite complex as well as portions of the Vijayan have resulted from synkimematic granitisation of the Highland Series (see Plate 1). The Tonigala Granite complex can be recognised as a late-kinematic group of rocks (Cooray, 1958), and intrusives of palingenetic origin occur in both the eastern and western Vijayan regions. We also know that there were at least two periods of post-tectonic activity.

On the other hand, we do not yet know (*a*) the relation of the south-west region to the rest of the island, (*b*) the relationship between the eastern and western Vijayan, (*c*) the time span of the Taprobanian cycle, and (*d*) whether the Pre-Cambrian rocks are related to only one geological cycle or to more. The answers to these and many other questions must await further applications of the concept of 'geological cycles' to our Pre-Cambrian rocks.

The Highland Series

Earlier workers like Coates (1935), Oliver and Erb (1957), and Vitanage (1959*a*) recognized that the so-called Khondalite Series were intimately associated with charnockites in the field, and further, that the two classes really belonged to a single succession of rocks. The name 'Highland Series' was therefore proposed for this association in the Rangala area (Cooray, 1961*a*), and subsequent mapping in the Ratnapura, Rakwana, Gampaha, Avissawella, Horana and Hatton sheets (see Table I, footnote) has shown beyond any doubt that the association of the Khondalite Group of metasediments and charnockites extends throughout the Central Highlands of Ceylon and is predominant in that area. The name

Highland Series, defined as a succession predominantly of metasediments and charnockites formed in the granulite facies of metamorphism (both pyroxene granulite sub-facies and hornblende granulite sub-facies rocks), displaying little or no subsequent granitisation, and folded into a parallel Taprobanian system of symmetrical, overturned, and recumbent folds, (see Cooray, 1962), is fully justified.

Perhaps the most important new fact about these rocks is that they belong to the pyroxene granulite sub-facies of metamorphism (see Cooray, 1961*a*; Hapuarachchi, 1961) and were formed at high temperatures and pressures, in a dry environment, at considerable depths in the crust. Some parts of the Highland Series belong to a slightly lower sub-facies, the hornblende granulite sub-facies and more data are needed to confirm this (see Hapuarachchi, 1961). It is significant that a similar close association of metasediments and charnockites exists in other high-grade metamorphic areas of the world (see Cooray, 1962); the detailed mapping being carried out in Ceylon is providing useful information for workers in such other areas. There is no doubt that as detailed mapping proceeds we shall learn much more about this interesting group of rocks, particularly its lithological characteristics, its origins, and the conditions in which it was formed.

Charnockites

Until 1948, the charnockites of Ceylon were said to have been intrusive rocks which crystallised from a magma, but Fernando (1949) doubted this and drew attention to their close association with the metasediments. We have now ample evidence, both published and unpublished, that the charnockites of the Highland Series are metamorphic rocks which have undergone the same metamorphic and tectonic history as the metasediments with which they are associated (see Cooray, 1962; 1964*a*).

Outside the Highland Series areas, however, are a variety of charnockitic rocks which resemble charnockites in some respects but not in others. These have been found on the eastern side of the Highland Series belt, within the Transitional Zone which marks the limits of the Vijayan granitisation (see Vitanage, 1959; Cooray, 1962), and are also well developed in the south-west region. Field and petrographic evidence in the Alutgama area (Cooray, 1964*b*) suggest that the charnockitic rocks have been formed by retrogressive metamorphism of normal charnockites under varying pressure/temperature conditions and in the presence of water; the ultimate product is a leucocratic biotite-garnet gneiss with irregular patches

of greyish, greasy looking charnockitic rock. These changes have accompanied processes of granitisation in the surrounding rocks.

Mapping in the south-west of the island has also shown that small bodies of intrusive charnockite, porphyritic or pegmatitic in character, do exist (see Cooray, 1964*b*). Little detailed work has yet been done on these rocks but they are thought to have formed by palingenesis (or melting and recrystallisation) of pre-existing charnockites. It is important to note that such intrusive charnockites have been found, so far, only in the south-west, where granitisation of pre-existing rocks is fairly common.

The high degree of metamorphism and recrystallisation of the charnockites makes it difficult for their antecedents to be recognized. On their field relations, particularly their association with quartzites and calc gneisses, and on the presence of sedimentogenous zircons (Vitanage, 1957), rounded monazite grains, and accessory graphite, it is probable that the majority of the charnockites were of sedimentary origin. At the same time, the possibility that some may have been originally volcanic or basic intrusive rocks cannot be ruled out. We must go on searching for the evidence.

The Vijayan Series

Owing to the fact that the rocks are relatively barren of economic minerals, and because most of the areas in which they occur are inaccessible, the Vijayan Series still remain largely unmapped, particularly in the north, east and south of the island (see Fig. 1 and Plate 1). In these areas they must still be considered to be "a heterogeneous association of gneisses, migmatites, granites and granitic rocks" (Fernando, 1949). The few traverses carried out in the Polonnaruwa, Elahera, Rangala and Kirindi Oya areas indicate that certain petrological types can be distinguished; and that the conditions of metamorphism approached the almandine-amphibolite facies (Hapuarachchi, 1961). The tectonic map has also revealed the existence of several dome-like structures near Komari, Okanda and Buttala (see Fig. 2B); these may be occupied by late-kinematic intrusives, made plastic by the action of permeating fluids, somewhat like the augen gneiss complex near Tissamaharama (Oliver and Erb, 1957).

Mapping in the Galgamuwa, Wariyapola, Dandagamuwa, and Gampaha areas on the north-western side of the island has shown that although there is at first sight a bewildering variety of rocks types, these can be sorted out into three or four main groups; the division is based partly on stratigraphic relations and partly on lithological characteristics.

In the Galgamuwa area, for example, it has been possible to recognise three main groups (Table III). They are,

A. A supracrustal group which include charnockites, amphibolites and metasediments; these are the earliest recognisable rocks and now occur only as scattered relic bands.

B. An early granodioritic group; these are dominantly leucocratic rocks which have been complexly folded and occupy relatively narrow strips of ground.

C. A later microcline granite group with associated migmatites (see Plate IIA); these rocks are predominantly pinkish in colour, and have replaced most of the earlier formed rocks. To this group belong the late-kinematic Tonigala Granite, sensu strictu, and the complex of pinkish to reddish gneisses which surrounds it.

TABLE III

FORMATIONS PRESENT IN THE GALGAMUWA SHEET

A. Metasedimentary Group
Crystalline limestone
Calc granulite and gneiss
Quartz-feldspar granulite with garnets
Charnockite
Amphibolite
Fine-grained biotite granulite.
B. Leucocratic gneisses and granitic rocks
Leucocratic biotite gneiss and granite gneiss
Leucocratic gneiss and granitic gneiss with veins and bands of pink granite
Leucocratic hornblende granite gneiss, foliated or streaky, with enclaves of amphibolite
Leucocratic augen gneiss and augen granite
Leucocratic porphyritic granite
Fine-grained granite gneiss
Granodiorite (trondjhemitic in parts), gneissic to streaky, with numerous basic enclaves
Granodioritic gneiss, veined and small-folded.
C. Pink, microcline gneisses and granites
Charnockitic gneiss with veins and bands of pink granite
Biotite schist with veins and bands of pink granite
Pink microcline granite, unfoliated
Pink granite, foliated or gneissic
Pink augen gneiss and augen granite
Pink migmatitic granite or granite gneiss with enclaves of hornblende-biotite schist and gneiss
Pink syenitic granite gneiss.

We can therefore define the Vijayan Series, in broad terms, as a polymetamorphic assemblage of partly mobilised, migmatitic and granitic rocks of the almandine-amphibolite facies with relic metasedimentary bands, plastically deformed into an irregular pattern of folds with a high degree of flow folding (Cooray, 1962).

One of the difficulties in mapping Vijayan-type granitic gneisses lies in finding a reliable basis for subdividing them. Coates (1935) attempted to do this on a regional basis (for example, *Bintenne Gneiss*, *Wanni Gneiss*, *Ritigala Gneiss*) but such a subdivision is inadequate for detailed mapping. The basis of mapping adopted in the Galgamuwa area is one possible method since it can lead to the recognition of rocks belonging to different phases of the geological cycle, but there are others. Whatever the method adopted, it is certain that to make any sense of such polymetamorphic areas, we must, when mapping, adopt a different approach to that used for the Highland Series. Lithological subdivisions are insufficient —we have to take into account antecedent, degree and agent of migmatisation, stage of granitisation, structural characteristics and tectonics, and ultimately, if possible, radiometric ages.

Relation of the Highland Series to the Vijayan Series

Adams (1929) was the first to suggest that the Khondalite Series were laid down on a 'basement' of Vijayan rocks and Wadia (1942) subsequently developed the idea that both were folded into a synclinorium; subsequent erosion had left 'outliers' of khondalites in the Vijayan, and 'inliers' of Vijayan in the Khondalite Series. In fact, all mapping previous to 1948 was devoted largely to demarcating these 'inliers' and 'outliers' and to mapping the boundary between the two major Series.

Detailed mapping of the Polonnaruwa, Rangala, and Kirindi Oya areas between 1952 and 1957 showed, however, that there is no sharp break between the two major groups on the eastern side of the island. On the contrary, they are separated by a Transitional Zone, 6 to 7 miles wide and extending for over 70 miles, in which representatives of both groups are present (Plate 1). We are now inclined to the view that the Vijayan Series on this side of Ceylon are the highly migmatised and granitised portions of the Highland Series, only scattered relics of which are seen; the boundary between the two is a zone marking the limits of this granitisation (see Cooray, 1962). On this theory, the Highland Series was formed by an early orogenic metamorphism, and the Vijayan Series, a polymetamorphic group formed from them by subsequent processes. Although there is no clear-cut transitional zone on the west, the same type of relationship between metasediments and charnockites on the one hand and migmatites and granitic gneisses on the other, appears to be present (see, for example, Pattiaratchi and Herath, 1962).

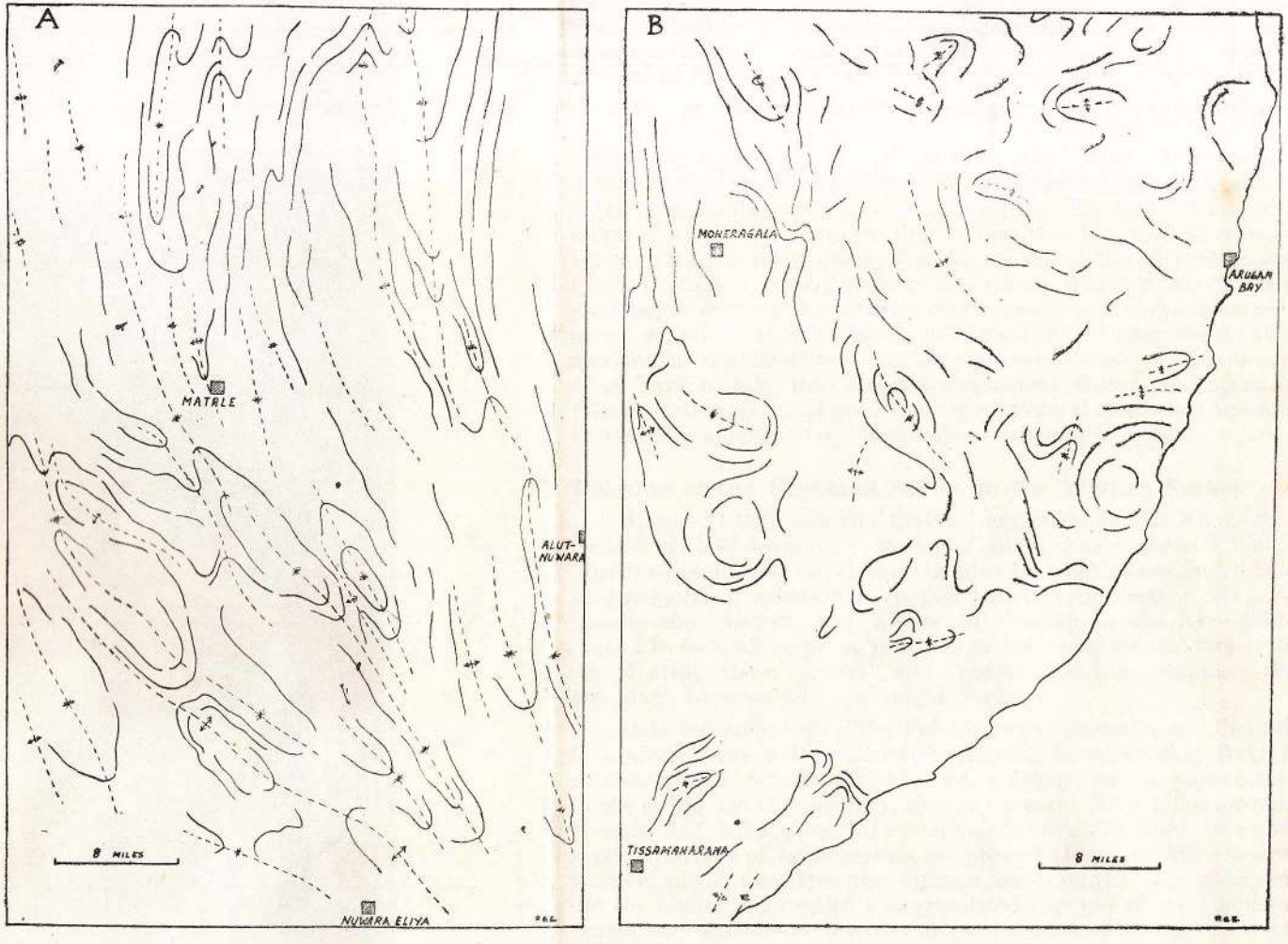
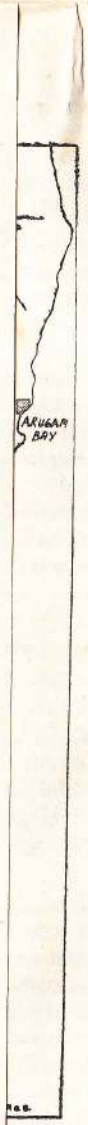


Fig. 2

A—Tectonic pattern typical of the Highland Series

B—Tectonic pattern in the Vijayan Series, south-east Ceylon

(From the Tectonic Map of Ceylon)
Foliation and bedding trends = continuous lines
Fold axes = broken lines.



Structure

The above ideas are closely related to the island's structure, and we may draw attention here to the advances brought about by the photo-interpretation studies carried out jointly by the Photographic Survey Corporation of Canada and the Geological Survey Department, a preliminary report of which was presented to this section some years ago (see Oliver, 1957). The Tectonic Map of Ceylon on the scale of 1 inch to 4 miles which has resulted from this joint work is now being published; it is a major contribution to our knowledge of the geological structure of the Island.

One of the most important facts revealed by this work is that there are marked differences in the tectonic styles of the Highland Series and Vijayan Series rocks (see Cooray, 1962). Whereas the former are folded into a number of parallel, regular, continuous folds which fall into a recognisable pattern, this is not so in the Vijayan (see Fig. 2). Here, foliation trends and fold axes become irregular away from the boundary, and considerable swirling of trend-lines indicates a high degree of plasticity of these rocks when they were deformed; they are seen, in the field, to have been "pushed and shoved about". Mobilisation, anatexis and palingenesis are common features in the Vijayan. Further, the Kataragama Complex, hitherto thought to be an outlier of the Khondalite Series, is now considered to be an enclave of Highland Series rocks which have escaped, to a great extent, the migmatisation and granitisation processes affecting the rocks surrounding it.

Modern methods of geological mapping involve the recording and use of as much structural data as possible. This, unfortunately, is a tool that has been inadequately used so far in this country, and until it is applied more intensively, our understanding of the Vijayan, and indeed of all Pre-Cambrian rocks, will be incomplete. Besides strike and dip of foliation, we need to map and interpret all visible lineations, the attitudes and shapes of minor folds, cleavages (where present), grain size, degree of homogeneity, attitude of banding and streakiness, and mineral elongation. We need also to begin petrofabric studies on properly oriented specimens.

The South-western Region

Geological mapping since 1958 has been largely concentrated in the south-west sector of the island, and from what is known about the Dandagamuwa, Gampaha, Avissawella, Horana and Alutgama areas (see Table I, footnote), certain rather unique features seem to be present here. That the geology of the south-west was somewhat different from the rest of the Island was, in fact, recognised by Coates in his geological map (1935) and even implied by Coomaraswamy (1902) in the term "Point-de-Galle Group", but the region

was included in the main 'Khondalite Series' belt when the Provisional Geological Map of Ceylon was published in 1948.

Some of the more distinctive features of this south-western belt are mentioned below:

(i) Cordierite, previously known only as a gemstone in Ceylon, was first described from the Horana and Alutgama areas in 1958 (Vitanage, 1959*b*; Cooray, 1959), but since then a variety of cordierite-bearing granulites and gneisses have been found throughout the region, some of them forming well defined bands. Their origin poses several interesting petrological problems.

(ii) Banded, metasedimentary ironstones of magnetite-hematite (Fig. 4A), though common in other Pre-Cambrian areas of the world, have only recently been found in Ceylon, in the Wilagedera-Panirendawa area (see Fernando, 1961). This is a discovery of major importance and was achieved largely by the use, for the first time, of geophysical methods (see Kumarapeli, 1964).

(iii) Wollastonite- and scapolite-bearing calc gneisses and granulites occur to the complete exclusion of crystalline limestones.

(iv) Many varieties of charnockites are found here, unlike in the Highland Series; they are altered, on a wide scale, to charnockitic rocks and ultimately to leucocratic biotite gneisses (see Cooray, 1964*b*).

(v) Intrusive hypersthene granites and charnockite pegmatities are common, as also are other intrusive granites like those at Arangala in the Alutgama area (Plate II B) and Ambagaspitiya and Loluwa in the Gampaha area (Hapuarachchi, Herath and Pattiaratchi, 1964).

(vi) The eastern margin of the region is marked by a wide, continuous zone of basic rocks (basic charnockites and pyroxene amphibolites) which runs through the Rakwana (Sinharaja Forest), Alutgama, Horana and Avissawella sheets. The relatively high magnetic intensity of this zone is due to the presence of magnetite in the basic rocks (see Pattiaratchi, 1959; also Hapuarachchi, Herath and Ranasinghe, 1964).

(vii) Primary monazite occurs in a number of granitic and gneissic rocks and has been concentrated by weathering in raised beaches, river terraces, and river mouths (Searle, 1962). Other radioactive and rare-earth minerals like thorianite, allanite, zircon, baddeleyite and some still unidentified minerals have also been recorded in pegmatites and granites in the area.

One of the major problems of Ceylon geology is the relation of the south-western belt to both the Highland Series and the Vijayan Series. In fact, throughout this south-western belt we find features that remind us of the Transitional Zone on the east. Charnockites and charnockitic rocks are associated on the one hand with metasediments, and on the other with plastically deformed migmatitic and granitic gneisses. Is it possible that we are here in an extremely wide 'transitional zone' in which the rocks may be given the status of a stratigraphic group and defined as a polymetamorphic assemblage of granulite facies charnockites and metasediments, partly granitised and migmatised along belts but retaining their early system of parallel Taprobanian folds? Or does it represent a change of sedimentary facies from the rest of the Highland Series, a change marked perhaps by an effusion of basic lavas? Are the two separated by a structural break? If the latter, are the rocks of the south-west more closely allied to those in the extreme south of India where wollastonite- and cordierite-bearing rocks are also found, and with which they appear to have a continuity of strike? Is it then part of a separate and later orogenic belt? Can we recognise here a cordierite granulite sub-facies of metamorphism? What is the true nature of the basic rocks? What is the age of the migmatisation? These are only a few of the questions that need to be answered—by careful analysis and correct interpretation of the large amount of data that have been collected.

Ages of the Pre-Cambrian rocks

Modern studies of Pre-Cambrian stratigraphy are increasingly based, as mentioned earlier, on the recognition of geological cycles, but for this it is necessary to have systematic data on the radiometric ages of rocks and minerals. Unfortunately, such data are still lacking for Ceylon, and only a few broad suggestions can be made on the ages that are available.

It has been customary to equate the Highland Series with the Khondalites of the E. Ghats belt of India, where ages of 1650 to 1750 million years have been recorded (Krishnan, 1960). However, though the correlation is geologically sound, owing to the remarkable similarity of the rocks types in the two regions, it must be remembered that they are several thousands of miles apart; further, there have still not been any ages recorded in Ceylon even approaching these figures.

Several K/Ar ages on micas were recently determined at the Institute for Pre-Cambrian Geology, Leningrad, by the late Professor A. Polkanov, and after his death in 1962, by Professors S.V. Obruchev and E.K. Gerling, the results being sent to the writer in a personal communication in 1963. These figures, as well

TABLE IV
RADIOMETRIC AGES OF MINERALS FROM CEYLON

Locality and Occurrence	*Mineral*	*Inferred Age*			*Source*
		U/Pb	Pb^{207}/Pb^{206}	*K/Ar*	
1. Ceylon, avg. of 21 gem-quality crystals	Zircon	574 ± 32			Gottfried & others (1956)
2. (not known)	Zircon		555 ± 20		Tilton & Aldrich (1957)
3. Balangoda District avg. of 8	Thorianite	490			Holmes (1955)
4. Rest of Ceylon, avg. of 10	Thorianite	490			,,
5. Balangoda District	Thorianite		485		Nier (1939)
Highland Series					
6. Nitre Cave, diopside-mica clot in cryst. lst.	Phlogopite			538	Obruchev and Gerling (Personal communication)
7. Norton Bridge, mica-spinel aggregate in cryst. lst.	Phlogopite			523[1]	,, ,,
8. Naranpanawa, mica schist lens in cryst. lst.	Phlogopite			512	,, ,,
9. Rangala, diopside-mica clot in calc gneiss	Phlogopite			481	,, ,,
10. Madumana, mica from pegmatite ..	Phlogopite			468[2]	,, ,,
11. Rangala, mica from quartz-feldsparintergrowth	Phlogopite			455	,, ,,
12. Arratenne, mica-schist w. associated graphite vein	Biotite			452	,, ,,
E. Vijayan (incl. Transitional Zone)					
13. Wilgomuwa, mica vein in calc granulite	Phlogopite			481	,, ,,
14. Mahawelakandia, granitic gneiss ..	Biotite			477	,, ,,
15. Mavaragala, biotite gneiss	Biotite			473	,,
S.W. Region					
16. Magala, granitoid rock	Biotite			459	,, ,,
17. Walpita, garnet-cordierite gneiss ..	Biotite			420	,, ,,
18. Beruwela, augen gneiss	Biotite			376	,, ,,
Others					
19. Uriya Wewa, near Anamaduwa, mica schist band in migmatitic gneiss ..	Phlogopite			438	,, ,,
20. Hambantota, biotitic selvage to granitic vein	Biotite			387	,, ,,
21. Bogala, mica schist near graphite vein ..	Biotite			349[3]	,, ,,

1. Average of 546 and 500 m.y.
2. Average of 476 and 460 m.y.
3. Average of 356 and 342 m.y.

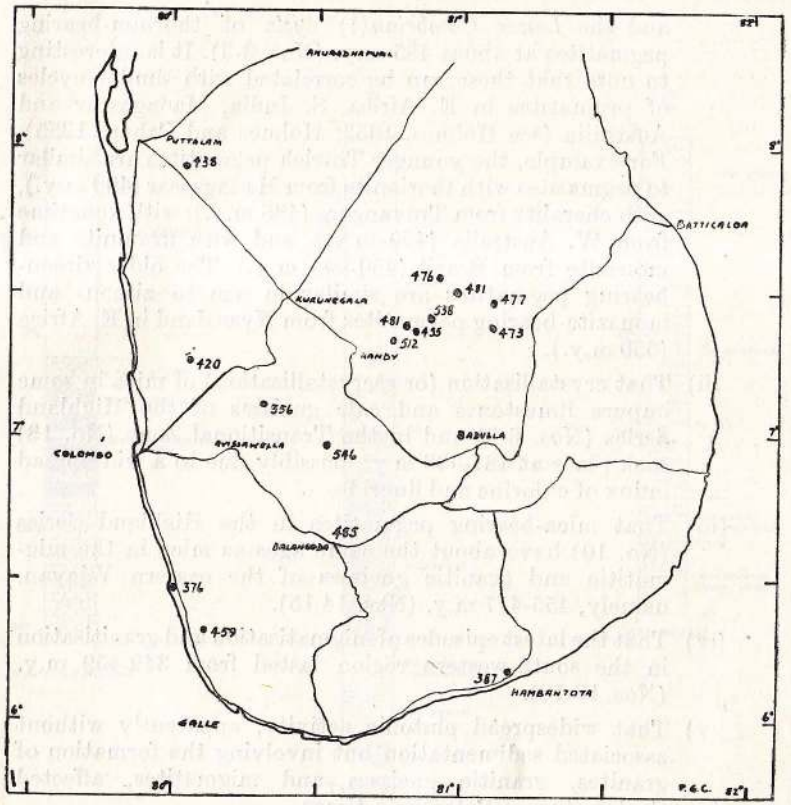


Fig. 3

Map of south Ceylon showing locations of samples of zircon, thorianite and mica on which radiometric ages have been determined (see Table IV).

as the more reliable of the earlier known figures are shown in Fig. 3 and Table IV.* It is not proposed here to have a full discussion of the radiometric ages but only to draw attention to certain significant facts. These are:—

(i) That two cycles of pegmatitic activity can be recognised, namely, the *Balangoda cycle* (Holmes, 1955) of zircon-bearing pegmatites at about 575 m.y. (Nos. 1 and 2),

* Earlier figures for feldspars and micas from Ceylon (quoted in Vitanage, 1959a) are omitted, as the incorrect decay constant for K^{40} was used in their determination (Obruchev & Gerling, personal communication).

and the *Lower Cambrian*(?) *cycle* of thorium-bearing pegmatites at about 485 m.y. (Nos. 3-5). It is interesting to note that these can be correlated with similar cycles of pegmatites in E. Africa, S. India, Madagascar and Australia (see Holmes, 1955; Holmes and Cahen, 1955). For example, the younger Th-rich pegmatites are similar to pegmatites with thorianite from Madagascar (490 m.y.), with cheralite from Travancore (485 m.y.), with xenotime from W. Australia (450 m.y.), and with uraninite and monazite from Brazil (450-500 m.y.). The older zircon-bearing pegmatites are similar in age to zircon- and monazite-bearing pegmatites from Nyasaland in E. Africa (550 m.y.).

(ii) That crystallisation (or recrystallisation) of mica in some impure limestones and calc gneisses of the Highland Series (Nos. 6-9) and in the Transitional Zone (No. 13) took place at 481-538 m.y., possibly due to a widespread influx of chlorine and fluorine.

(iii) That mica-bearing pegmatites in the Highland Series (No. 10) have about the same ages as mica in the migmatitic and granitic gneisses of the eastern Vijayan, namely, 455-477 m.y. (Nos. 14,15).

(iv) That the latest episodes of migmatisation and granitisation in the south-western region lasted from 349-459 m.y. (Nos. 16-18).

(v) That widespread plutonic activity, apparently without associated sedimentation but involving the formation of granites, granitic gneisses and migmatites, affected Ceylon during Palaeozoic times.

We are beginning to see that some of what has hitherto been regarded as 'Pre-Cambrian' contains rocks that have been formed during much later episodes of plutonism (see Table II). A systematic study of the radiometric ages of rocks and minerals in the main lithologic-metamorphic units is badly needed. Only then can their stratigraphic relations to each other and to the rest of Peninsular India be known.

THE JURASSIC

It was in 1925 that Wayland described the small basin of Jurassic sediments known as the Tabbowa Beds (Wayland, 1925) and several years later that the evidence of well-cuttings revealed the presence of a small concealed basin of Jurassic rocks at Andigama (Deraniyagala, 1939). Apart from the identification of the fossil plant impressions at Tabbowa and the establishment of their stratigraphic position as Upper Gondwana (Sitholey, 1944), no major work has been done since then on these rocks.

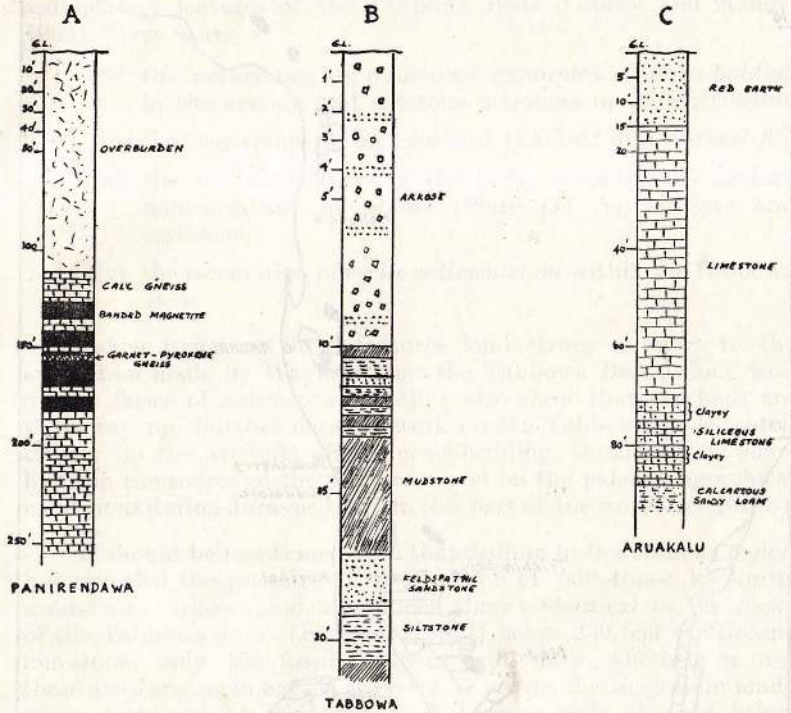


Fig. 4

Detailed geological sections.

A—Panirendawa, showing interbedded calc gneiss and magnetite beds (after P. S. Kumarapeli).

B—Tabbowa, showing alternation of coarse-grained (arkosic) and fine-grained (siltstone and mudstone) deposits.

C—Aruakalu, showing variability of Miocene limestone (after Herath, Pattiaratchi and Fernando).

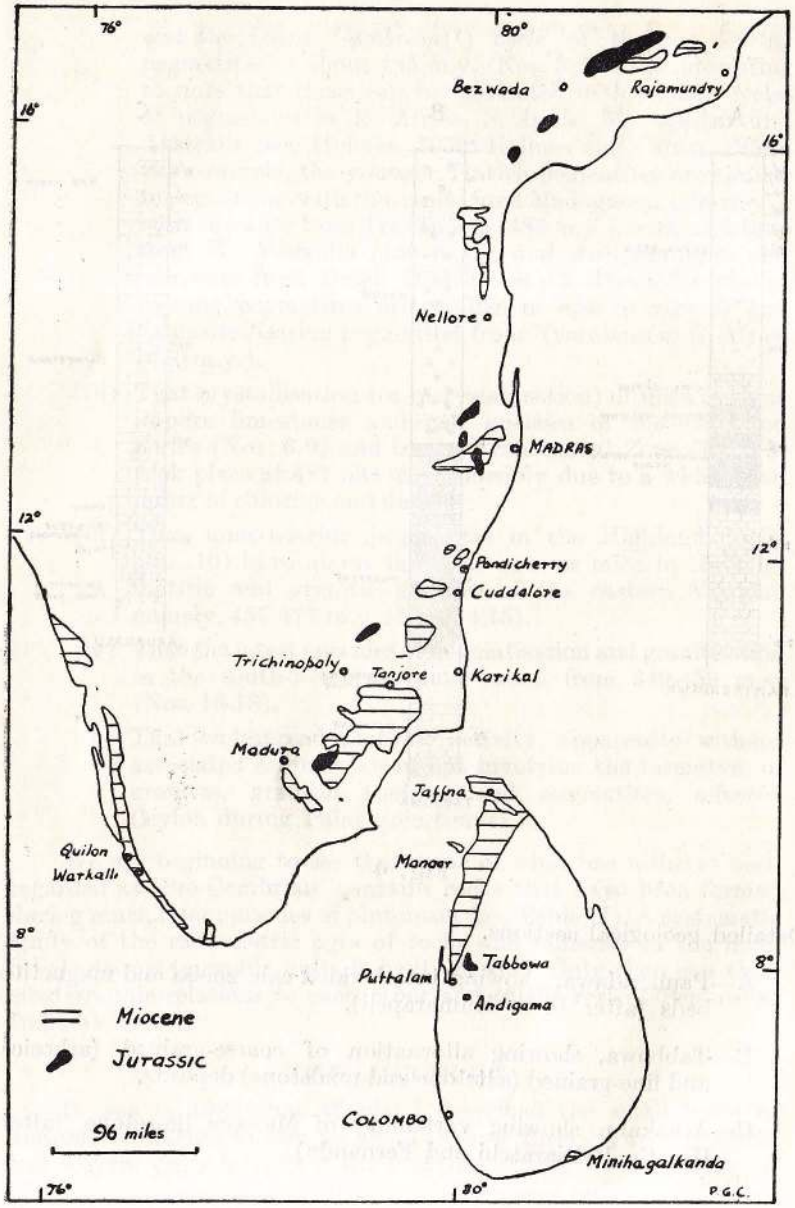


Fig. 5

Map of Ceylon and South India showing the distribution of Jurassic and Miocene beds.

A quick re-examination of the Tabbowa Beds was, however, made in 1963, when the detailed measurement of five exposed sections (see Fig. 4B) led to the discovery of some interesting sedimentary features of the Tabbowa Beds (Cooray and Money, 1963). These were:—

(i) the occurrence of numerous examples of *cross-bedding* in the arkosic and siltstone members of the succession,

(ii) the occurrence of an excellent example of a *channel-fill*,

(iii) the re-classification of the beds, according to modern nomenclature, as *arkose* (Plate III A), *siltstone* and *mudstone*,

(iv) the recognition of *cyclic sedimentation* within the Tabbowa Beds.

Taken together, these features lend strong support to the suggestion made by Wayland that the Tabbowa Beds belong to a deltaic facies of sedimentation; they also show that the beds are right-way up. Further detailed work on the Tabbowa Beds, particularly on the attitude of the cross-bedding, should throw some light on the source of the sediments and on the palaeogeographical conditions during Jurassic times in this part of the world (see Fig. 5).

It should be mentioned here that drilling in the Mannar region has revealed the presence of a succession of "siltstones, kaolinitic sandstones, arkoses and mudstones almost identical to the rocks of the Tabbowa area" (Fernando, 1964) below 250 feet of Miocene limestone; only the fossil evidence will show whether or not these are Jurassic in age. If they are, it proves the suggestion made many years ago by Wayland, that Jurassic beds do exist below and concealed by later Miocene deposits in N.W. Ceylon (see Fig. 6).

THE MIOCENE

Perhaps the most important new information on the Miocene rocks of Jaffna, first described by Wayland and Davis (1923), has been the re-identification of many of the fossils (Eames, 1950), the establishment of the correct stratigraphic position of the beds, and their correlation with other Tertiary formations in the region. We owe this new data to foreign workers like F.E. Eames, S.N. Rao and B.S. Tewari, among others.

The Jaffna limestone is now assigned to the upper part of the Lower Miocene, and specifically to the Burdigalian Stage (zones $f_1 - f_2$) (Rao *et al.*, 1957). It is therefore to be correlated with the Upper Gaj beds of Kathiawar and Kutch (mainly shales, marls, limestones and sandstones), and with the Quilon Limestones, all on the west coast of India (see Table V and Fig. 5). The Dj. Bichri

TABLE V

CORRELATION OF THE MIOCENE BEDS OF WESTERN INDIA AND CEYLON

(after Rao and others)

Age	*Jaffna*	*Quilon*	*Surat-Broach*	*Kathiawar*	*Kutch*
	Jaffna Limestone	Quilon Limestone		*Orbiculina* Limestone	*Orbiculina* Shales Naliya Beds
			Agate Conglomerates		
Burdigalian				Viswara Limestones Bhatia Limestones	
Aquitanian?					*Miogypsina* Beds

beds of Syria and the upper part of the Asmari Limestone of Persia are also Burdigalian in age (Table VI). Such index microfossils of the Burdigalian as *Austrotrillina howchini* and *Taberina* (= *Orbiculina* = *Orbitoilites*) *malabarica*, have recently been recorded from Nirukiri in Jaffna (Tewari and Tandon, 1960).

New information has also come to light regarding the lithology and extent of the Miocene beds during the course of drilling investigations carried out by the Geological Survey Department. In the Aruakalu area, north of Puttalam (see Fig. 4C), the limestone was found to be over 200 feet thick and to vary from pure limestone and siliceous limestones to calcareous sandstones and impure calcareous muds or marls (Herath, Pattiaratchi and Fernando, 1961). More recently, drilling in the Mannar district has shown that the highly sheared floor of crystalline rocks on which the Miocene limestone rests slopes rapidly westwards; the overlying Jaffna Limestone therefore thins out to the east, the boundary being in the vicinity of a line joining Madhu Road, Palampiddi and Tunukkai, where the most westerly exposures of crystalline rocks are seen (Fernando, 1964).

The most southerly surface outcrop of probable Miocene limestone was noted last year. It occurs about one mile south of the Mi Oya, on the footpath between Villuke and Sittaraveli. Drilling

TABLE VI

DISTRIBUTION OF FORAMINIFERA COMMON TO WESTERN INDIA, CEYLON, INDONESIA AND MIDDLE EAST

(after Rao and others)

		Eocene	*Oligocene*	*Aquitanian*	*Lower Miocene Burdigalian*	*Middle Miocene*
INDONESIA		a/b	c	d	e(1-4) e 5	f1 — f2	f2 — f3
Austrotrillina howchini							
Nephrolepidina							
Tryblio lepidina							
Miogypsina s.s.							
Lepidosemicyclina thecidaeformis							
L. bifida							
Miogypsinoides							
WESTERN INDIA AND CEYLON	Kutch			*Miogypsina* Beds	Naliya Lst. *Orbiculina* Shales		
	Kathiawar				Viswara Lst. *Orbiculina* Lst. Bhatia Lst.		
	Surat-Broach				Agate conglomerates		
	Quilon				Quilon Lst.		
	Jaffna				Jaffna Lst.		
MIDDLE EAST	Syria	Dj. Abiad, Biehri and Kalaai Jaaber	Dj.		Dj. Abiad	Dj. Biehri	
	Palestine		Ramleh Koubeibeh				
	Turkey	Asmari Limestones			Aintab		
	Persia						Fars Series

for the Puttalam Water Supply scheme a few years previously had also revealed what was probably Miocene limestone at 45 feet below alluvial sands and clays of the Mi Oya valley (Sirimanne, 1957).

Mention should also be made of a description of the Miocene occurrences at Minihagalkande on the south coast, by Deraniyagala (1961); it contains some interesting photographs of the actual occurrences and of fossil echinoids.

THE QUATERNARY

Although the Quaternary System of Ceylon is the source of much of its mineral wealth—mineral sands, clays, gem-bearing gravels, building sands and gravels, laterite, lateritic iron ore and water—very little attention has hitherto been paid to it, and few attempts made to map and classify it systematically. The distinction of specifically 'Pleistocene' deposits in early classifications was based largely on Wayland's description of the Plateau gravels with Palaeolithic stone implements which he described as 'Pleistocene' (Wayland, 1919). More recently there has been the description of the Pleistocene of Ceylon (Deraniyagala, 1959), but this is largely a systematic description of the Pleistocene fossils occurring in river valleys of the Ratnapura region; some account is also given of the lake and river deposits in which these fossils are found.

Within the last two years, however, three sheets along the north-west coast between Negombo and Puttalam have been mapped in detail (see Table I, Fig. 1), the main information regarding the Quaternary deposits of this region being presented in a paper to Section D of this Session (Seneviratne, Kumarapeli and Cooray, 1964). In the absence of fossiliferous evidence of specific age, these authors have preferred to group all post-Miocene deposits as Quaternary, and to sub-divide them into an Older Group and a Younger Group as follows:—

YOUNGER GROUP	Alluvium; lacustrine-lagoonal and estuarine deposits Beach and dune sands Littoral sandstone
OLDER GROUP	Red Earth Group (including unconsolidated grey and white sands) Ferruginised basal gravel. Terrace gravel.

At the base of the Quaternary are two distinct formations, the lower member of which is a basal, partly ferruginised, gravel with a predominance of angular to sub-angular quartz grains (see Fig. 6). It is fairly widespread, but is generally concealed by later deposits; its presence has been noted from well-excavations and drill holes. Overlying this gravel in places is the Red Earth Group which

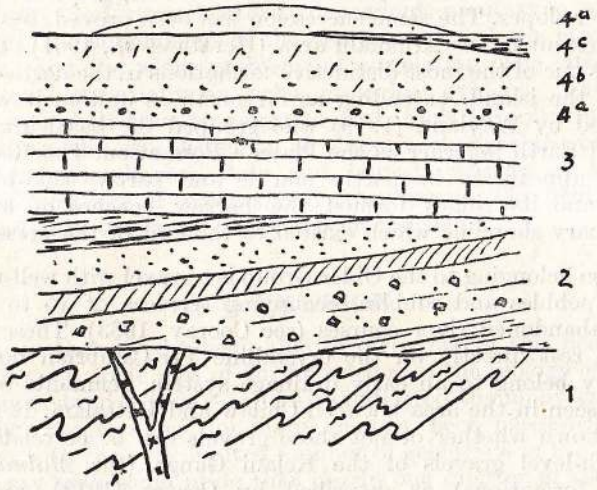


Fig. 6

Generalised succession of strata in north-west Ceylon (modified after Wayland).

1. *Tonigala Complex* with later pegmatitic intrusions (?Pre-Cambrian).
2. *Tabbowa Beds* (Upper Jurassic)
3. '*Kalpitiya Beds*'; Kalpitiya limestone including some argillaceous and sandy beds (?Lower Miocene).
4. '*Karativu Beds*' (Older Quaternary)
   (*a*) Basal ferruginous gravel
   (*b*) Red Earth (—Wayland's 'Red Sands of Pomparippu').
   (*c*) *Mundel Beds* and older lagoonal deposits (Younger Quaternary).
   (*d*) Blown sand, beach sand etc. (Younger Quaternary).

consists of loamy sands; these form elongate N-S ridges or plateaus and are brick-red in colour, but turn to brown, grey, or white on the lower slopes. The same succession has been proved, by drilling, to be present in the Aruakalu area (Herath *et al.*, 1961). The Red Earth is one of the most distinctive formations in the north-western part of the island, extending as far north as Jaffna; it was first described by Wayland (1919) who grouped the basal gravel and the Red Earth together as the Plateau Formation. The Red Earth deposit appears to be partly marine and partly wind-blown in origin, and its ridges formed the barrier beaches of an older Quaternary shoreline which existed to the east of the present one.

Also belonging to the Older Group is a gravel with well-rounded quartz pebbles and cobbles, occupying terraces of 15 to 25 feet above abandoned river courses (see Cooray, 1963). These can be seen to rest directly on the crystalline Pre-Cambrian floor, and probably belong to an early drainage system, remnants of which can be seen in the area between Chilaw and Puttalam. It remains to be shown whether or not these gravels can be correlated with the high-level gravels of the Kelani Ganga (the *Malwana* and *Ranale* formations) as described by Coates (1913), and with 'gravels' from several other parts of the island as recorded in the writings of Parsons, Wayland, Coates and Wadia in departmental files. Among the secondary formations, a nodular ironstone which develops alike on Quaternary gravels and on decomposed crystalline rocks, is seen, in places, to pass into a massive ferricrete crust (Plate III B); this formation is widespread in the north-west and other parts of the Dry Zone of Ceylon.

Regarding the Younger Group, it need only be said that an attempt has been made to recognise deposits of several sedimentary facies within it. In the absence of fossil evidence, it may be possible to build up a composite picture of the later Quaternary of Ceylon by correlating the lagoonal, estuarine, alluvial, and marine formations of one region with another.

That Quaternary geology is closely linked with physiographical evolution is becoming more and more evident. Changes in land and sea level, so well demonstrated in the Colombo area by Wadia (1941), older shorelines, raised beaches, older drainage systems and estuaries, the growth of barrier spits and bars, the formation of lagoons and their silting up, and the existence of older coral reefs can all be found, if looked for. Some of these features are clearly the result of Quaternary climatic changes, and the recognition of a cool pluvial *Ratnapura phase* and an arid *Palugahathurai phase*, during the 'Pleistocene' (Deraniyagala, 1959) is a helpful pointer in this direction. The only reliable date we have of any Quaternary formation is 3000 years for the coral bed below Hikkaduwa Rest House (Shepard, 1964).

There are many problems to be solved in the Quaternary geology of the island, among which might be mentioned those of Quaternary stratigraphy and nomenclature (dependent upon type areas or localities), the subdivision of the Quaternary and the recognition of stages, and the wider problems of correlation. The latter have to be based essentially on palaeontological evidence, but in its absence, we shall have to resort to such tools as pollen analysis, C-14 dating, the establishment of climatic changes (as indicated by lithology, landforms, flora and fauna), changes of land-sea level, the existence of terraces, and other means. The field is wide open, and only the fringe has yet been touched.

CONCLUSION

I hope that I have been able to give you, in this necessarily concise review, some idea of our knowledge of the geology of Ceylon as we understand it today—and also indicated the many gaps that still exist. Progress has, for many reasons, been necessarily slow, but even with the very limited number of geologists in the country we can say with some pride that nearly one quarter of the island has been geologically mapped within the last twelve years. This is no mean record, and no one should ever again be persuaded into believing that hardly anything has been done in the way of mapping the geology of Ceylon. Progress in the future will, however, be determined by how soon we can overcome the two serious shortages that we suffer from in the field of the geological sciences, and I would like to end this address by drawing attention to them.

One is the absence of geology from the content of secondary school education in Ceylon. Geology, being the science concerned with the earth's crust (or the topmost few miles of the planet), touches us at many points in our daily lives. It affects us, for example, through the physical landscape, in our search for and utilisation of materials to build with and sites to build on, in our search for raw materials for industry, through the soils we cultivate, and in our never-ending quest for water for domestic use. A basic knowledge of the elements of geology seems therefore, to me at least, to be a pre-requisite for anyone who claims to be an intelligent and well-informed citizen of the country in which he lives. Only then can the misconceptions and lack of understanding about the subject, among laymen as well as scientists alike, be eradicated—and where better to start than with the child?

The second, and more serious, shortage is the absence of a Department of Geology from the University, though I understand that this, at long last, is soon to be remedied. The result has been not only a serious shortage of trained geologists in the country (such training having to be obtained abroad), but also of civil

engineers with a basic knowledge of the geological principles and processes so essential to some fields of their work. Such a scarcity has been, and still is, highly detrimental to the country in the present stage of its development. Further, although significant advances have been made on several fronts in understanding the general geology of the island, there are many specialised fields such as mineralogy, structural geology, palaeontology, geochemistry, and ore genesis, to name only a few, which are as yet untouched or only barely touched. These are fields in which those engaged primarily in geological mapping cannot spend much time, however much they would like to. We can only point the way—as I have shown in the course of this address—but detailed studies in such fields need to be carried out by those well trained in the methods of geological research, working in well equipped laboratories, and with time to carry out such research. They must be persons who are attracted to the many fascinating aspects of the rocks and minerals of the earth, and are not afraid of the inconveniences and hardships of a geologist's life in the field. The compensations, I can assure them from personal experience, are many and varied, and not to be found in any other science.

Acknowledgements

I am indebted to the Director of the Geological Survey Department for permission to draw on material that is still unpublished, and to my colleagues in the Survey for information freely given about the areas they have worked in and for many stimulating discussions throughout the years.

References

Adams, F.D. (1929). The Geology of Ceylon. *Canad. J. Res.*, **1**, 425.

Coates, J.S. (1913). High-level gravels of the lower Kelani Ganga valley. *Progress Report No.* 41, *Mineral Survey Ceylon* (unpublished).

————(1935). The geology of Ceylon. *Ceylon J. Sci.*, *(B)* **19**, 101.

Coomaraswamy, A.K. (1902). The Point-de-Galle Group (Ceylon): wollastonite-scapolite gneisses. *Quart. J. geol. Soc. Lond.*, **58**, 680.

Cooray, P.G. (1958). A preliminary account of the Tonigala Granite. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 42.

————(1959). in *Admin. Rep. Ceylon Govt. Mineralogist* for 1958.

————(1961a). The geology of the country around Rangala. *Ceylon Dept. Miner. Mem.* **2**.

————(1961b). On the sub-division of the Ceylon Pre-Cambrian. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 26.

————(1962). Charnockites and their associated gneisses in the Pre-Cambrian of Ceylon. *Quart. J. geol. Soc. Lond.*, **118**, 239.

————(1963). The Erunwala Gravel and the probable significance of its ferricrete cap. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 27.

————(1964*a*). The charnockites of Ceylon: a review. *Geol. Soc. India Mem.*, No. 3, in press.

————(1964*b*). The geology of the country around Alutgama. *Ceylon geol. Surv. Dep. Mem.* **3**, in press.

————and Money, N.J. (1963). Lithology and sedimentation of the (Jurassic) Tabbowa Beds. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 27.

Deraniyagala, P.E.P. (1939). A carbonaceous Jurassic Shale from Ceylon. *Spolia zelan.*, **21**, 193.

————(1959). *The Pleistocene of Ceylon.* Colombo: Government Printer.

————(1961). The amphitheatres of Minihagal Kanda, their possible origin and some of the fossils and stone artefacts collected from them. *Spolia zelan.*, **29**, 149.

Eames, E.E., (1950). On the ages of Certain Upper Tertiary Beds of Peninsular India and Ceylon. *Geological Mag.*, **81**, 233.

Fernando, L.J.D. (1949). Some aspects of the geology of Ceylon. *Proc. Ceylon Ass. Advanc. Sci.*, **2**.

————(1961). *Admin. Rep. Ceylon Govt. Mineralogist* for 1960.

————(1964). *Admin. Rep. Director Ceylon geol. Surv. Dept.* for 1962-63.

Gottfried, D., Senftle, F.E., and Waring, C.L. (1956). Age determination of zircon crystals from Ceylon. *Amer. Min.*, **41**, 157.

Hapuarachchi, D.J.A.C. (1961). A new approach to the problems of metamorphism in Ceylon. *Proc. Ceylon Ass. Advanc. Sci.*, **2**, 109.

Hapuarachchi, D.J.A.C., Herath, J.W. and Ranasinghe, V.V.C. (1964). The geological and geophysical investigations of the Sinharaja Forest area. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 23.

Hapuarachchi, D.J.A.C., Herath, J.W. and Pattiaratchi, D.B. (1964). Brief explanation to Geological Map of Gampaha, in press.

Herath, J.W., Pattiaratchi, D.B. and Fernando, L.J.D. (1961). Report on Cement Raw Material Investigations in the Puttalam North District. *Ceylon geol. Surv. Dept.* unpub. report.

Harpum, J.R. (1960). The concept of the geological cycle and its application to the problems of Pre-Cambrian geology. Rept. XXI. Int. geol. Congr., **9**, 201.

Holmes, A. (1955). Dating the Pre-Cambrian of Peninsular India and Ceylon. *Proc. Geol. Soc. Canada*, **7**, 81.

————and Cahen, L. (1955). African geochronology. *Colon. Geol. min. Resour.*, **5**, 3.

Krishnan, M.S. (1960). *The geology of India and Burma.* 4th Ed.

Kumarapeli, P.S. (1964). Method and results of investigations for iron ore at Panirendawa in Chilaw District. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 23.

Nier, A.O. (1939). The isotopic constitution of radiogenic leads and measurements of geologic time—II. *Physical. Rev.*, **55**, 153.

Oliver, R.L. (1957). The geological structure of Ceylon. *Ceylon Geographer*, **11**, 9.

————and Erb., D.K. (1957). Reconnaissance study of the geology of the Kirindi Oya basin. *Ceylon Geographer*, **11**, 25-29.

Pattiaratchi, D.B. (1959). In *Admin. Rep. Ceylon Govt. Mineralogist* for 1958.

Pattiaratchi, D.B. and Herath, J.W. (1962). The relations between the 'Vijayan' and 'Khondalite' Series in the country around Gampaha, W.P. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 25.

Rao, S.R.N., Tewari, B.S., Mohan, K., and Chatterji, A.K. (1957). The Miocene of Western India. *Geol. Mag.*, **94**, 81.

Searle, D.L. (1962). A geological report on the radioactive mineral occurrences of the Alutgama area. *Ceylon Geol. Surv. Dept.*, unpub. report.

Seneviratne, L.K., Kumarapeli, P.S., and Cooray, P.G. (1964). The Quaternary deposits of north-west Ceylon. *Proc. Ceylon Ass. Advanc. Sci.*, **i**, 22.

Shepard, F.P. (1964). Criteria in modern sediments useful in recognizing ancient sedimentary environments. *Developments in Sedimentology*—Vol. I. Amsterdam.

Sirimanne, C.H.L. (1957). The Puttalam Water Supply. *Ceylon Geol. Surv. Dept.*, unpub. report.

Sitholey, R.V. (1944). Jurassic plants from the Tabbowa Series in Ceylon. *Spolia zelan.*, **24**, 3.

Tewari, B.S. and Tandon, K.K. (1960). On microfauna and age of Jaffna Limestone, Ceylon. *Proc. 47th Sess. Ind. Sci. Congr.*, **5**, 281.

Tilton, G.R. and Aldrich, L.T. (1955). The reliability of zircons as age indicators. *Trans. Amer. geophys. Un.*, **36**, 531.

Vitanage, P.W. (1957). Studies in zircon types in Ceylon Pre-Cambrian Complex. *J. Geol.*, **65**, 117.

————(1959*a*). The geology of the country around Polonnaruwa. *Ceylon Dept. Miner. Mem.* 1.

————(1959*b*). In *Admin. Rep. Ceylon Govt. Mineralogist* for 1958.

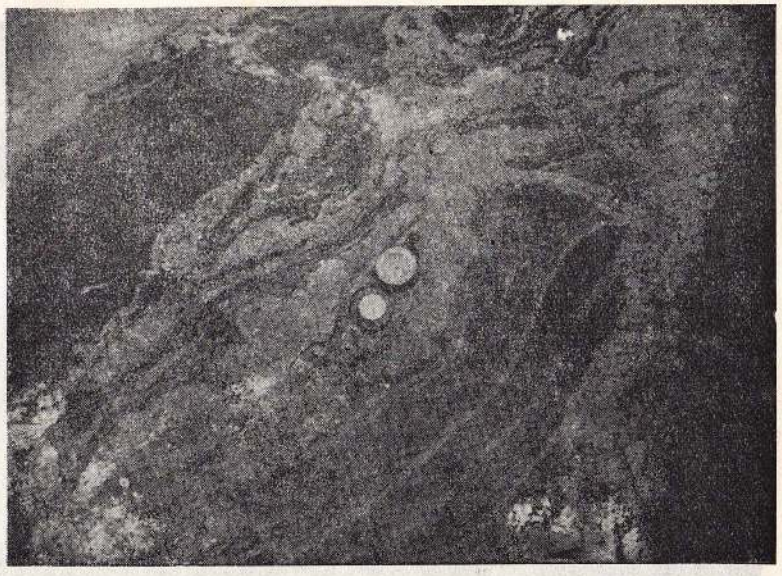
Wadia, D.N. (1941). The Geology of Colombo and its environs. *Spolia zelan.*, **23**, i.

————(1942). The making of India. *29th Sess. Ind. Sci. Cong.*

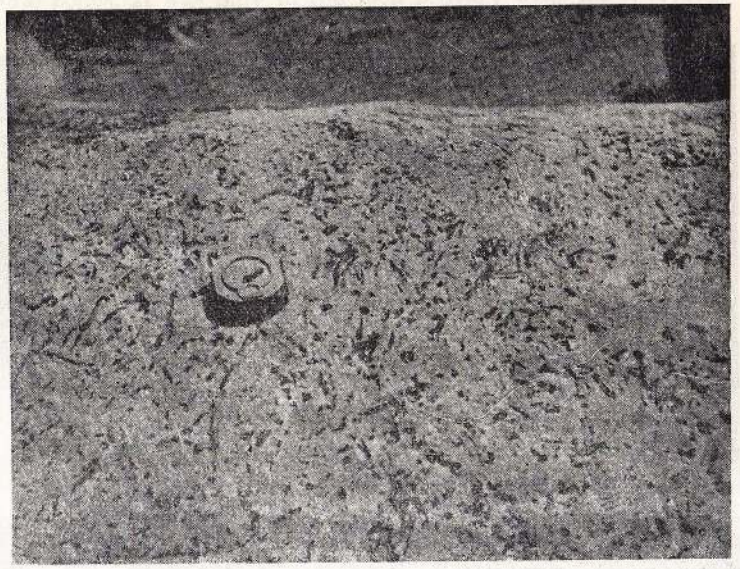
Wayland, E.J. (1919). Outline of the Stone Ages of Ceylon. *Spolia zelan.*, **11**, 85.

————(1925). The Jurassic rocks of Tabbowa. *Ceylon J. Sci.*, (*B*) **13**, 195.

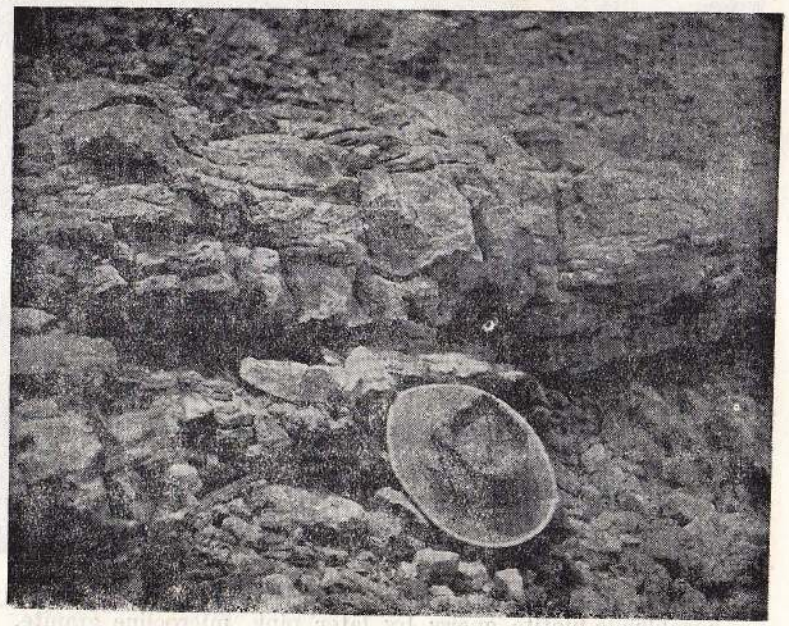
————and Davis, A.M. (1923). The Miocene of Ceylon. *Quart. J. geol. Soc. Lond.*, **79**, 577.



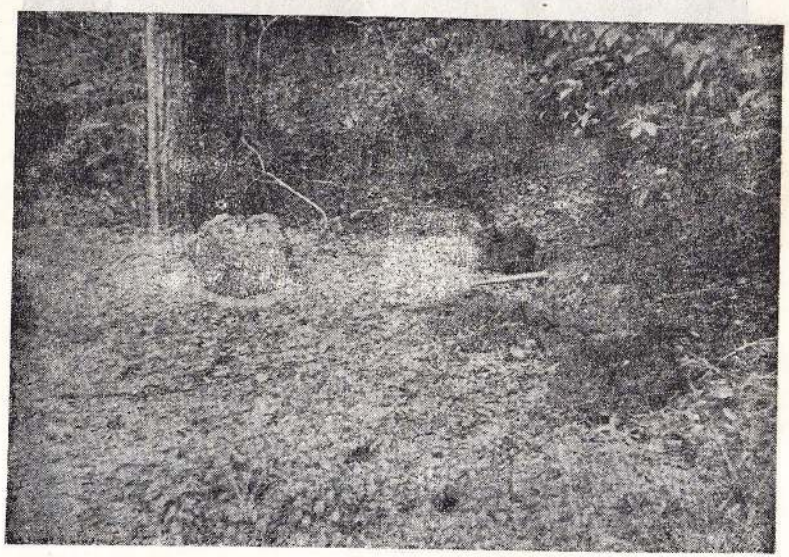
A. Migmatite in the Tonigala Complex, showing intrusion of early hornblende-biotite gneiss by later pink, microcline granite.



B. Porphyritic hornblende granite at Arangala. Note large hornblende crystals.



A. Arkosic bed at Tabbowa.



B. Boulders of ferricrete on surface, east of Battulu Oya.

SECTION E—PHYSICAL SCIENCES

Presidential Address
25th September, 1964

TRANSIENT VARIATIONS OF THE EARTH'S MAGNETIC FIELD

by

M. L. T. Kannangara

1. Introduction

The Earth's Magnetism was the subject of a Presidential Address to this section by Mr. S. P. Baliga a few years back. In that address, the Earth's main magnetic field, its origin and its secular variations were described and discussed. My address to you today is in many ways a continuation of the fascinating story unfolded to you on that occasion.

Transient variations of the Earth's Magnetic Field rarely exceed half per cent of the main field. The unit used to describe these variations is the gamma (γ) which equals 10^{-5} gauss. To quote from Sidney Chapman: "Throughout the aeons during which man evolved, these variations were entirely hidden from him and had no impact whatever on his life, except perhaps in the auroral zones. It was only in the 19th century with the invention of telegraphy, that they began to interfere, significantly with his affairs. In the present century they arouse still greater practical concern through their influence on radio propagation and very recently because of possible associated dangers to travellers in interplanetary space".

In this address I propose to give you a brief review of some of the important experimental facts that have emerged concerning transient variations of the Earth's magnetic field and a broad idea of the sort of picture of the Earth's environment that is crystallising out from an analysis of these facts.

2. The Sq, L and SD variations

If records produced by an instrument which makes a continuous recording of the Earth's magnetic field at a given station (a magnetogram) are examined, one finds that on an average sort of day, the trace shows a slight waviness with a period of the order of a day, and superposed on this one would see a few odd irregular transient variations of a shorter period. On some days, the trace would be smoother than usual, whereas on other days it would be much more irregular. On a qualitative basis any day could be classified as being quiet, average or disturbed and given

a character figure of 0, 1 or 2 respectively. Character figures collected from fifty magnetic laboratories spread out all over the world are collected and averaged to the first decimal place to give what is known as the international magnetic character Ci, which must by definition have a value between 0 and 2. The Ci figure is a rough qualitative measure of the magnetic activity. There is also a more quantitative measure of the magnetic activity called the "K index" which is calculated from the averaged amplitudes of the magnetic variation measured over three hourly periods.

When records pertaining to magnetically quiet days are separated and analysed, a definite pattern emerges. A regular variation having a period equal to a solar day and an amplitude ranging from 40 to about 100 γ depending on the latitude of the station is detectable. This variation is called the quiet Sun or Sq variation. It is found that the Sq variation is enhanced over a narrow strip, about 450 Km. wide, immediately above the magnetic dip equator. Sq variations larger than 100 γ have been reported by the Huancayo Station in Peru and also by Appapillai and Baliga in Colombo. Fig. 1. shows typical records obtained by Appapillai and Baliga.

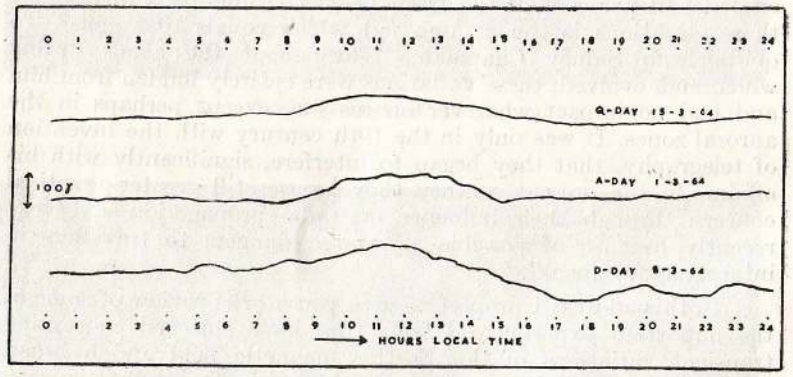


Fig. 1

Examples of magnetograms recorded in Colombo by Appapillai & Baliga.

A much smaller variation of amplitude 1 to 5 γ and having a period equal to the 25 hour lunar day has also been separated out from the quiet day records. This is called the lunar or L variation.

The daily variations that are obtained by averaging out over a suitable period of time, say a few months, magnetograms obtained for all days or for disturbed days alone are called Sa and Sd respectively. The difference variations Sa-Sq and Sd-Sq are found to be daily variations of a similar type, which is quite different from the Sq pattern. This difference variation is called the disturbed daily variation and is denoted by SD.

It is instructive to compare and contrast the Sq variation with the SD. Fig. 2 shows the pattern of the Sq and SD variation for different latitudes. It will be noticed that the Sq variation is greatest round about 12 noon with hardly any variation at night. In contrast, the SD variation at night is comparable with that by day. Also, unlike the Sq variation, the SD is a diurnal variation where the first harmonic is clearly predominant. Both variations

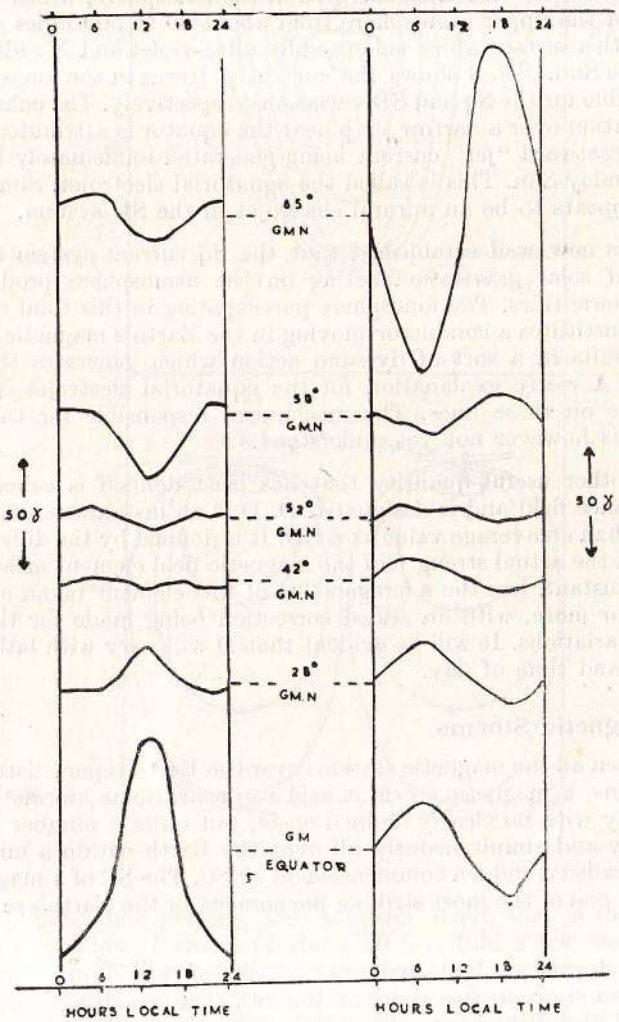


Fig. 2

Sq and SD patterns for different latitudes (Sugiaru and Chapman, 1961).

are more or less symmetrical with respect to the equator, but whereas the Sq is greatest near the equator, the SD is most prominent in auroral latitudes. Further both variations become reversed at certain latitudes; for the Sq this is between 30° and 40°, whereas it is at about 55° for the SD.

The Sq, L and SD variations have been traced to have their origin in electric currents generated in the ionosphere, which is the region of the upper atmosphere from about 50 to 500 miles above the Earth's surface which is ionised by ultra-violet and X radiation from the Sun. Fig. 3 shows the current patterns in the ionosphere responsible for the Sq and SD variations respectively. The enhanced Sq variation over a narrow strip near the equator is attributed to a sort of eastward "jet" current being generated immediately below the noonday Sun. This is called the equatorial electrojet. Similarly there appears to be an auroral electrojet in the SD system.

It is now well established that the Sq current system is the result of solar gravitation acting on the atmosphere producing atmospheric tides. The ionosphere participating in this tidal movement constitutes a conductor moving in the Earth's magnetic field. This results in a sort of dynamo action which generates the Sq system. A ready explanation for the equatorial electrojet is also available on these lines. The mechanism responsible for the SD system is however not yet understood.

Another useful quantity that has been defined is called the disturbance field and is designated D. D is an instantaneous value rather than an average value as is SD. It is defined by the difference between the actual strength of the magnetic field element measured at any instant, less the average value of this element taken over a month or more, with an added correction being made for the Sq and L variations. It will be evident that D will vary with latitude, station and time of day.

3. Magnetic Storms

When all the magnetic stations over the Earth report disturbed conditions, a magnetic storm is said to occur. Some storms begin gradually with no clearly defined onset, but quite a number begin suddenly and simultaneously all over the Earth within a minute. This is called a sudden commencement or SC. The SC of a magnetic storm is one of the most striking phenomena of the Earth's magnetism.

SC storms are best analysed in terms of the disturbance field D measured from the time of the SC. The resulting variation is called the disturbance storm time variation and is designated symbolically by Dst. The Dst variation of the horizontal component of the Earth's magnetic field has been found to follow a more

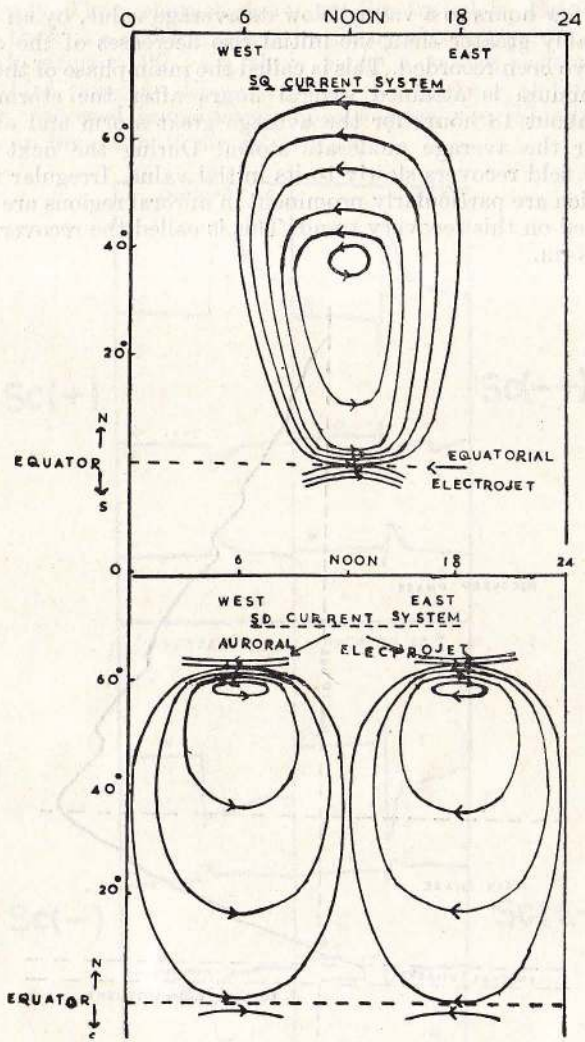


Fig. 3

or less definite time pattern. The inception itself, that is the SC, is usually a sudden increase of about 20 γ within a few seconds, of the magnetic field parallel to the geomagnetic axis (that is, the horizontal component H for middle and low latitudes, and the vertical component Z for the poles). The increased field stays elevated for one or two hours. This is called the initial phase of the storm. After this, the same field component decreases rapidly

within a few hours to a value below its average value, by an amount considerably greater than the initial rise decreases of the order of 400 γ have been recorded. This is called the main phase of the storm. The minimum is attained several hours after the storm onset, usually about 18 hours for the average great storm and about 26 hours for the average moderate storm. During the next several days the field recovers slowly to its initial value. Irregular fluctuations which are particularly prominent in auroral regions are usually superposed on this recovery trend. This is called the recovery phase of the storm.

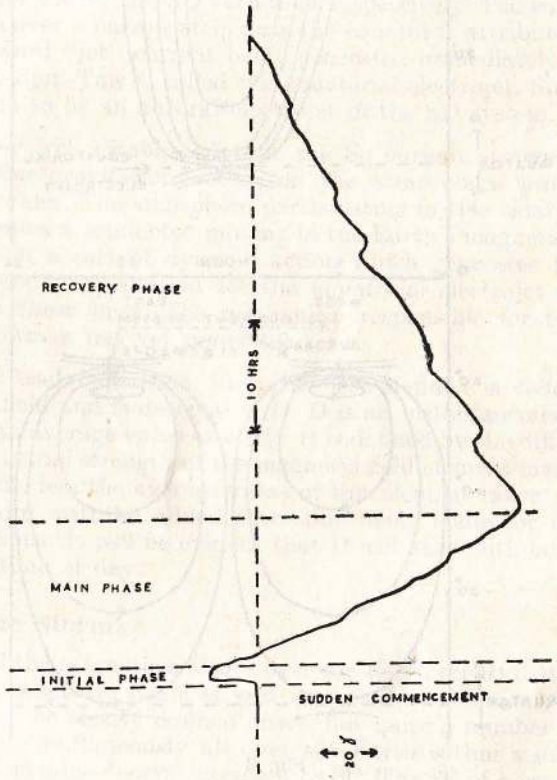


Fig. 4
Typical SC storm pattern.

The sudden commencement itself can take different forms. The commonest is the type just described which is denoted by SC (+) Ferraro, Parkinson and Unthank (1951) sorted out four types of SC's. They are now designated SC (+), SC (-), SC (-+) and SC (+-)

respectively. Their characteristics are illustrated in Fig. 5. In an SC (+) there is a sudden increase in H by about 20 to 30 γ. An SC (-) is an inverted SC (+), there being a sudden decrease in H. In the third category designated SC (-+), there is at first a small decrease, followed immediately by a sudden rise. Finally the SC (+-) is an inversion of the SC (-+).

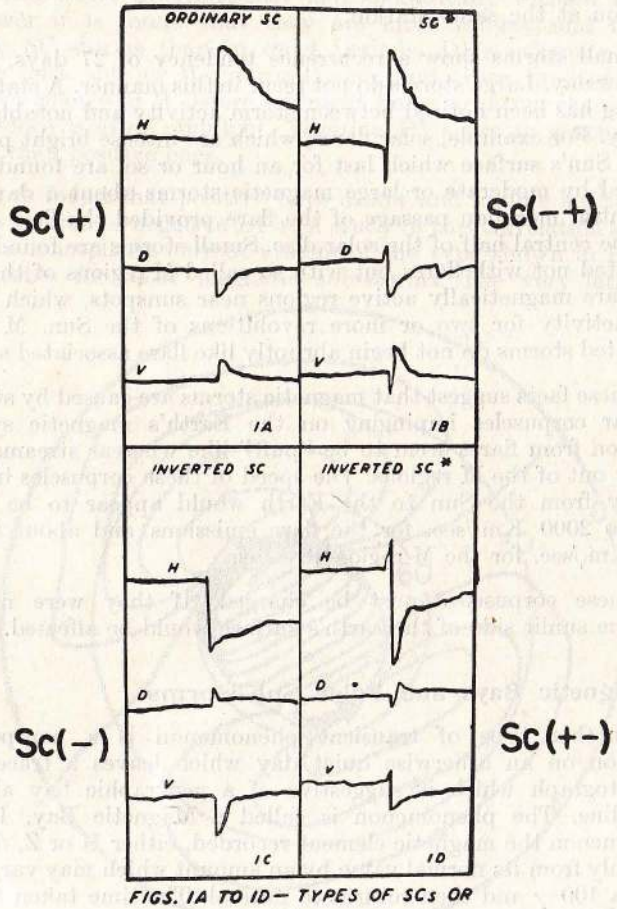


Fig. 5
Four typical SC's (Ferraro, Parkinson and Unthank, 1951).

The type of SC usually seen in middle and low latitude is the SC (+). For example, out of 800 SC's examined at a latitude of 9.5° GM North by Chakrabarthy in 1951, only 28 were SC (-), the rest were SC(+). The SC(-+) becomes more prominent in higher

latitudes. For example, out of 340 SC's examined at 62.5° GM North by Watson and McIntosh in 1950 only 65 were SC(+) and 162 were SC(−+).

Another interesting fact is that a marked diurnal variation in the amplitude of SC's has been discovered at the equatorial station Huyancayo in Peru. The variation is very similar to the Sq variation at the same station.

Small storms show a recurrence tendency of 27 days, which is a solar day. Large storms do not recur in this manner. A statistical time lag has been noticed between storm activity and notable solar activity. For example, solar flares, which are intense bright patches on the Sun's surface which last for an hour or so, are found to be followed by moderate or large magnetic storms about a day after the central meridian passage of the flare provided the flare occurs over the central half of the solar disc. Small storms are found to be associated not with flares but with so called M regions of the Sun. These are magnetically active regions near sunspots, which retain their activity for two or more revolutions of the Sun. M region associated storms do not begin abruptly like flare associated storms.

These facts suggest that magnetic storms are caused by streams of solar corpuscles impinging on the Earth's magnetic system. Emission from flares seem to be "puff" like whereas streams seem to flow out of the M regions. The speed of these corpuscles in their journey from the Sun to the Earth would appear to be about 1000 to 2000 Km./sec. for the flare emissions and about 500 to 1000 Km./sec. for the M region streams.

These corpuscles must be charged. If they were neutral, only the sunlit side of the earth's surface would be affected.

4. Magnetic Bays and Polar Sub-storms

Another type of transient phenomenon is a non-periodic variation on an otherwise quiet day which leaves a trace on a magnetograph which is suggestive of a geographic bay along a coast line. The phenomenon is called a Magnetic Bay. In this phenomenon the magnetic element recorded, either H or Z, departs smoothly from its normal value by an amount which may vary from 10 to a 100 γ and then returns to normal. The time taken for the entire phenomenon is about an hour or two.

For the same magnetic bay at a particular station, it has been observed that the variation of H is in a sense opposite to that of Z. A bay in which H increases is called positive and one which H decreases, negative. Positive bays are more common in low and middle latitudes and they occur mostly at night. Negative bays usually make their appearance in the early afternoon.

According to Chapman and Akosofu (1963) bays in middle and low latitudes are related to a type of severe polar disturbance called a polar sub-storm or a DP. In some of these DP's the magnetic field can change by an amount as large as 1500 γ, which is to be compared with a large Dst variation of—400 γ in low latitudes. DP's occur intermittently and sporadically without any SC. However it is found that they are more intense and frequent during SC storms than in quiet periods. DP's originate usually in the dark side of the auroral zone, and there they produce a positive change in H before midnight but a negative change if the occurrence is after midnight. Negative DP's are in general much larger than positive ones.

It is found that positive bays in low and middle latitudes are closely related to negative DP's. A study of this correlation suggests that an ionospheric current system of the type shown in Fig. 6 is responsible for polar magnetic sub-storms. The very large field

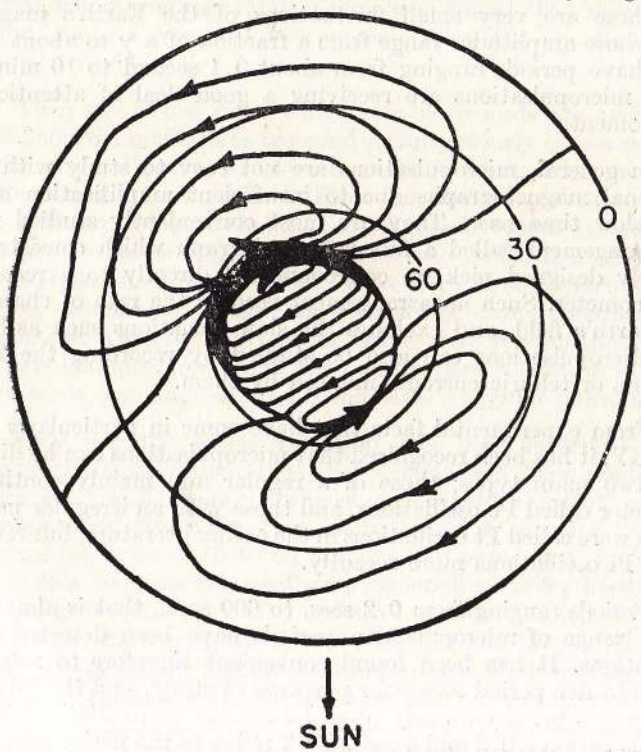


Fig. 6
The current system for a polar magnetic substorm (Silsbee and Vestine, 1942).

change in auroral latitudes is attributed to the generation of an intense westward "auroral electrojet" in the ionosphere. The mechanism responsible for such a DP current system is still obscure, although there is evidence to show that it is intimately tied up with the precipitation of streams of charged particles into the ionosphere along magnetic lines of force in the auroral regions; for example, DP's are often accompanied by spectacular auroral displays.

Another type of recognised magnetic disturbance is called a crochet. It is like a small scale bay. These crochets are often simultaneous with an observable solar flare and is then referred to as a solar flare effect or SFE. It would appear that direct ultraviolet or X radiation from the Sun acting on the ionosphere is responsible for this effect.

5. Magnetic Micropulsations

These are very small fluctuations of the Earth's magnetic field whose amplitudes range from a fraction of a γ to about 10 γ. They have periods ranging from about 0.1 second to 10 minutes. These micropulsations are receiving a good deal of attention at the moment.

In general, micropulsations are not easy to study with conventional magnetographs due to insufficient amplification and a very slow time base. They are most conveniently studied using an arrangement called a flux rate variograph which consists of a suitably designed pick up coil connected directly to a recording galvanometer. Such an arrangement records the rate of change of the Earth's field, and excludes the slow variations such as Sq, L etc. Micropulsations can also be studied by recording the Earth currents or telluric currents induced by them.

From experimental facts that have come in particularly since the IGY, it has been recognised that micropulsations can be divided into two main types; those of a regular and mainly continuous character called Pc oscillations, and those with an irregular pattern which were called Pt oscillations in the earlier literature, but referred to as Pi oscillations more recently.

Periods ranging from 0.2 secs. to 600 secs., that is almost the whole range of micropulsation periods have been detected in Pc oscillations. It has been found convenient therefore to subdivide Pc's into five period bands for purposes of study, and the following notation is being currently used. Pc 1 refers to oscillations having periods between 0.2 and 5 secs.; Pc 2 refers to the period range 5 to 10 secs.; Pc 3 from 10 to 45 secs.; Pc 4 from 45 to 150 secs. and Pc 5 from 150 to 600 secs. This division was recommended by Jacobs *et al.*, (1964).

Pc 3 and Pc 4 oscillations occur almost exclusively on the daylight side of the Earth and they tend to last for many hours. When Pc type oscillations occur at night they are almost always found to be of Pc 1 or Pc 2 type. There seems to be some indication that the maximum activity of Pc oscillations are at a period which changes from about 55 secs. in auroral latitudes to about 25 secs. in middle latitudes, and it was thought that this tendency towards smaller periods would continue up to equatorial regions. However two stations in the Phillipines, Baguio 16° N and Cobu 11° N have recently reported that they obtain a maximum activity at about 50 secs. period. At the last annual sessions of the CAAS, Dr. Fernando and I reported that the maximum activity in Colombo was at a period of about 55 secs.

There is general agreement that Pc 3 and Pc 4 oscillations are most active round about mid-day, peak activity being reached a little before mid-day. It has also been found that regardless of period, the occurrence of Pc's increase from the equator towards higher latitudes.

There are a few cases where reliable records are available of Pc 3 and Pc 4 oscillations recorded simultaneously over a wide area where phase differences between oscillations occurring at distant stations could be measured. From an analysis of these records it has been possible to deduce the overhead current system in the ionosphere which would induce the observed oscillations. Jacobs and Sinno (1960) in a review article show the current system deduced for two separate Pc's and Fig. 7 *a* and 7 *b* are reproduced from this article. They draw attention to the close resemblance between the two examples in spite of the comparative roughness in reading the records, and suggest that it indicates a regular phenomenon.

Very recently much attention has been focussed on Pc 1 type oscillations. Smith (1964) has summarised their predominant characteristics. They occur almost exclusively at night, generally towards dawn. Their duration rarely exceed 1 hour and often two or more intervals of 15 to 60 minutes will occur in a single night. When this happens the oscillations occurring earlier in the night have longer periods of about 1 or 2 secs., and the oscillation intervals that follow have progressively shorter periods till periods of the order of 0.5 sec. or less are seen near dawn. Pc 1 activity is most frequent during generally disturbed periods, but within these periods the oscillations most often come of a very quiet background.

Pt or Pi type micropulsations are characterised by their irregular form. They appear to be closely connected with other

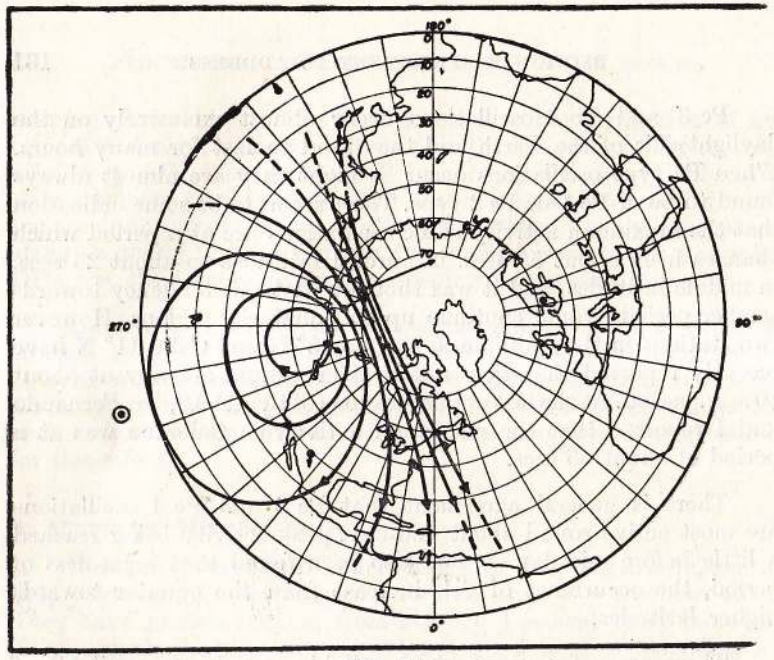


Fig. 7 *a*

Equivalent overhead current-system for Pc at 2345-2350 U.T. February 12th, 1958. (After Jacobs and Sinno, 1960).

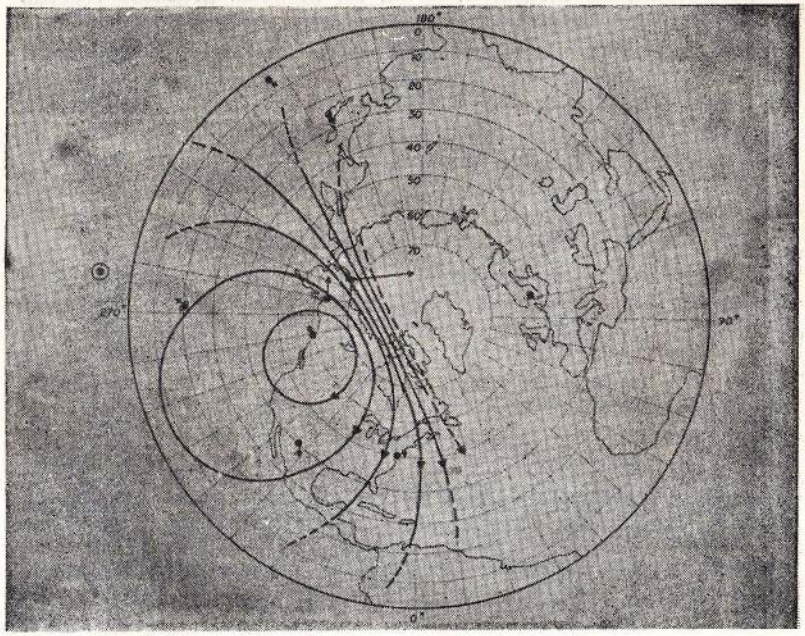


Fig. 7 *b*

Equivalent overhead current-system for Pc at 2220-2350 U.T. April 2nd, 1958. (After Jacobs and Sinno, 1960).

transient magnetic disturbances such as magnetic bays and also with some upper atmospheric phenomena such as the equatorial electrojet.

Until recently, these oscillations were supposed to have periods and amplitudes larger than Pc's. But now it is recognised that these irregular oscillations have a higher frequency component with periods less than 20 secs. They are thus classified into two groups called Pi.1 with periods between 1 and 20 secs. and Pi.2 with periods larger than 20 secs. and less than 150 secs. The Pi.2 oscillations correspond to the Pt oscillations referred to by earlier workers.

Most authors agree that the maximum activity of Pi.2's is round about local midnight. Magnetic bays are the most common phenomena associated with Pi.2 oscillations. Pi.2's also appear when there is a sudden commencement of a geomagnetic storm, but these Pi.2's seem to have significantly different characteristics from the bay associated oscillation. For example, bay Pi.2's have relatively long periods between 40 and 120 secs. whereas SC associated Pi.2's have periods nearer 20 secs. Bay Pi.2's very often have Pi.1's superposed on them, whereas SC Pi.2's are simple pulsations. They are also significantly different as regards waveform; bay Pi's are asymmetric whereas the SC Pi.2's are symmetric. Bay Pi.2's come in at night, SC Pi.2's are seen during the day time as well. Yanagihara 1959 has reported an 11 year cycle of Pi.2 activity which is inversely proportional to solar activity.

Pt. 1 oscillations appear to have the following characteristics. They start almost at the same time as Pi.2 oscillations. The Pi.1 periods seem to be independent of the period of the accompanying Pi.2's. The same Pi.1 trains are observed at different stations which confirm that they are not an artificially induced phenomenon. The predominant activity of Pi.1's is in the 5 to 10 second period band, which is the same as the predominant period band of night time Pc activity.

It has also been observed that if a Pc oscillation with period 5-10 secs. already exists before the beginning of a Pi.2 oscillation, then this Pc oscillation is enhanced when the Pi.2 appears.

This is only a brief summary of some of the main characteristics of magnetic micropulsation. Some examples of micropulsations recorded in Colombo by Dr. Fernando and myself are shown in Fig. 8.

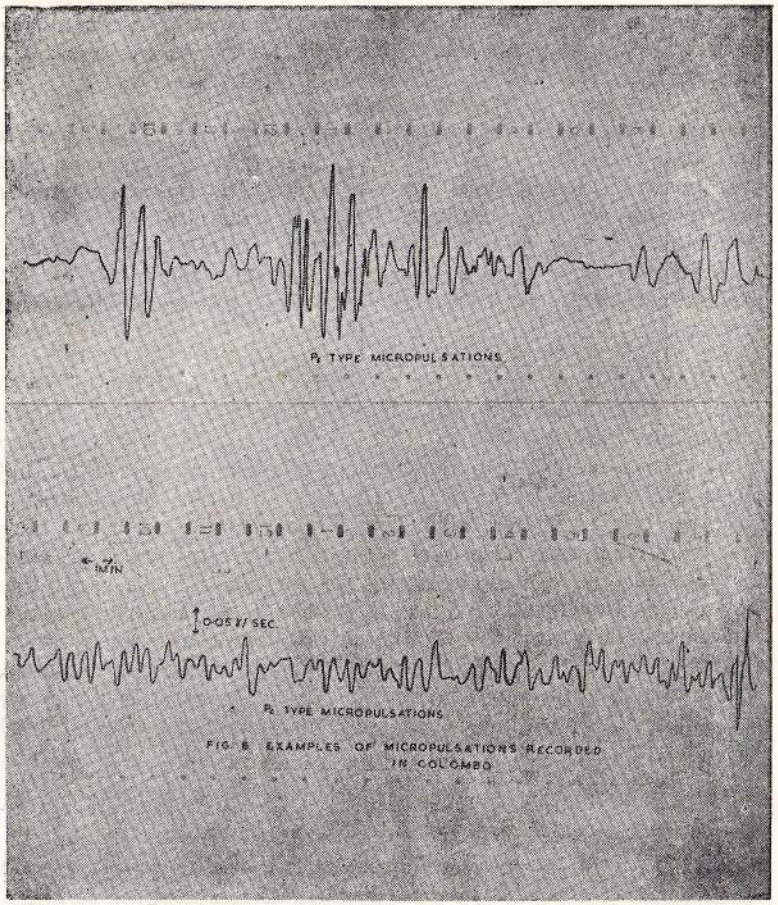


Fig. 8
Examples of micropulsations recorded in Colombo.

6. Theoretical Interpretation

No comprehensive theory exists which explain all the experimental facts. In a brief address of this sort I can only give you a glimpse of some of the lines on which theoreticians are working.

The "freezing in" of magnetic lines of force into a completely ionised gas or plasma is an important concept used in this type of work. It is merely an application to a particular situation of the familiar Lenz's Law, according to which if a rigid conductor is moved across a magnetic field or if a magnetic field is made to vary in a conductor, E.M.F's are generated which induce a current

system in the conductor which opposes the variation in the magnetic field. It is useful to recall how Maxwell extended this concept even to empty space and conceived of the idea that whenever there is a rate of change of magnetic field an appropriate electric field will always appear. It was this idea that led to Maxwell's famous equations which explained the propagation of electromagnetic waves through empty space. In empty space and in an insulating material, the induced electric field does not result in an electric current, so that magnetic lines of force can slip through these media easily. When the medium is a conductor, there is resistance to the movement of a line of force relative to the medium; a sort of magnetic viscosity develops in the medium. Going to the limit, we then find that in a completely conducting, unbound medium like a plasma, any relative motion of magnetic lines of force through the medium is completely inhibited. The lines of force are said to be "frozen in" to the medium. Any movement of the lines of force involves a movement of the medium itself, and conversely any motion of the medium must take with it the lines of force as well. Under such circumstances, a magnetic disturbance is transmitted by means of what are called "hydromagnetic waves" or Alfven waves after Alfven who first drew attention to them. Alfven waves are transverse waves travelling along the magnetic lines of force in the manner of transverse waves travelling along stretched strings and have a velocity, given by $u = \sqrt{\frac{\mu H^2}{4\pi P}}$ where μ is the magnetic permeability, H the strength of the field and P the density of the medium. Also, since there cannot be a movement of the medium without carrying with it the magnetic lines of force, ordinary sound or pressure waves through a plasma propagated at right angles to the lines of force are also affected, and one gets longitudinal waves known as "magneto-sonic waves" being propagated at right angle to the lines of force. The velocity of these waves is $(u^2 + C_S^2)^{\frac{1}{2}}$ where u is the Alfven wave velocity and C_S is the velocity of sound waves through the plasma in the absence of a magnetic field. For conditions prevailing in the upper atmosphere C_S is very much smaller than u and can be neglected.

Using this sort of hydromagnetic picture, the Earth's atmosphere can be divided into three regions. The first region, extending from the Earth's Surface to a height of about 100 Km is effectively non-ionised and is a non-conductor. In this region there is no coupling at all between the magnetic lines of force and the atmosphere. Magnetic disturbances would be propagated as normal electro magnetic waves with the speed of light.

The second region extending from a height of about 100 Km to 500 Km consists of the E and the lower F regions of the ionosphere. Here the atmosphere is ionised, but relatively weakly.

Numerous collisions with neutral molecules prevent the ions from drifting bodily. The region behaves more like a rigid conductor and when it moves bodily across the Earth's lines of force in the general tidal motion of the atmosphere currents are induced which result in the Sq and L systems. The important point is that problems connected with this region of the ionosphere could be called magneto-ionic and are not essentially hydromagnetic.

The third region which extends from a height of about 3000 Km right out to about 60,000 Km or ten Earth radii is called the magnetosphere. The outer boundary of this region merges with interplanetary matter. This is the region of the great radiation belts surrounding the Earth. Collisions between gas molecules are unimportant, and the gas in this region is completely ionised. The Earth's magnetic lines of force are frozen into this region, and the material contained therein is bound to the Earth by the magnetic lines of force. The Earth's magnetic field appears to be completely contained within the outer boundary of this region, because rocket observations as reported by Sonnet *et al.*, (1960) indicate that there is an abrupt cutoff in the distant geomagnetic field at a distance of about ten Earth radii.

In between the second and third regions just described here is a transition region, where although the lines of force are "frozen in" collisions cannot be neglected. If the Alfven wave velocity is plotted as a function of altitude for equatorial regions there is an inversion in the transition region as seen in Fig. 9 which is based on a curve given by Dessler 1961. It has been shown that this inversion effect causes the magnetosphere and the transition region to behave like two resonant cavities for hydromagnetic waves, each being able to sustain a different hydromagnetic oscillation. There is a school of thought, for example Yanagihara (1959) which is of the opinion that the continuous Pc micropulsations are due to such oscillations.

We shall now pass on to the space between the Earth and the Sun. Guided by the properties of the visible corona, Chapman (1959) conceived that the Earth might be embedded in an extended solar corona too tenuous to be observed spectroscopically. Chapman estimated a particle density of about 10^3 per cm^3 in the solar atmosphere in the vicinity of the Earth.

Bierman (1957) and Parker (1958) put forward a more dynamic picture. They based their ideas on the observation that comets, tails are always directed away from the Sun. They showed that radiation pressure alone was insufficient to cause this and inferred the existence of a "solar wind" with a particle density of 10 to 100

atoms per cm^3 blowing from the Sun at a speed of about 500 Km/sec. due to a continuous emission of fast particles. Superposed on this continuous "wind" there would be "guests" from solar flares and

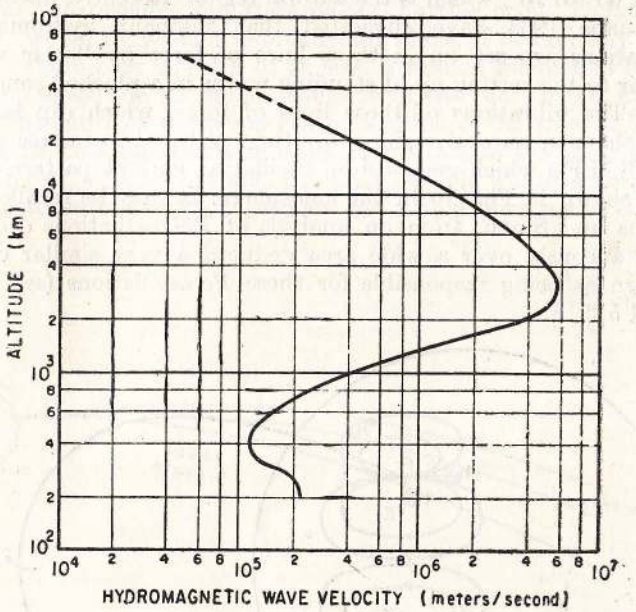


Fig. 9

Velocity of hydromagnetic waves in the geomagnetic equatorial plane versus altitude (Dessler, 1961).

"streams" from M-regions. These solar particles would consist largely of electrons and protons in equal numbers, because the streams as a whole cannot be charged. Electrostatic forces would cause any concentration of charge to dissipate in a fraction of a second.

One sees that all these theories lead to the idea that interplanetary space is filled with tenuous but highly conducting gas. Any magnetic lines of force existing in the region will be frozen into this gas. It is at once evident that the Earth's magnetic field cannot penetrate into this region, so that the Earth system held together by its magnetic field must be an entity separated from interplanetary matter by a fairly sharp boundary. The boundary region is estimated to be only about 1 Km thick.

A solar wind playing on the outer surface of the Earth system is thought to stimulate transverse hydromagnetic waves travelling along the outer lines of force of the geomagnetic field. If this boundary

region is at a distance of about seven to ten Earth radii in the equatorial plane of the Earth, then it can be shown that lines of force in this region will dip into the Earth's surface at a latitude of about 65° to 70°, which is the auroral region. Recently, Jacobs and Watanabe 1964 have suggested that resonant hydromagnetic oscillations are set up in these lines of force by "solar winds" similar to the setting up of standing waves in a plucked sonometer wire. The vibrations of these lines of force, which dip into the ionosphere in auroral regions, are then pictured to induce oscillating E.M.F's which generate an oscillating current pattern of the type shown in Fig. 10 in the ionosphere. It may be recalled that Jacobs and Sinno, from an analysis of Pc oscillations occurring simultaneously over a wide area deduced a very similar current pattern as being responsible for these Pc oscillations (see Fig. 7 *a* and *b*).

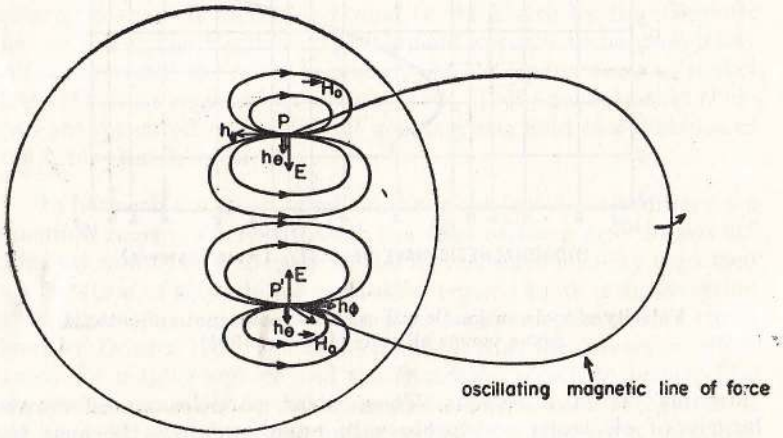


Fig. 10

Schematic representation of electric currents on the proposed model. (After Jacobs and Watanabe, 1964).

The mechanism put forward by Jacobs and Watanabe is interesting from the point of view that it requires Pc oscillation periods to decrease from auroral latitudes to middle latitudes and then to increase again until at the equator the periods are the same as that for auroral latitudes. This is in complete agreement with the observation made by Dr. Fernando and I in 1963, which I have referred to earlier, that the maximum activity of micropulsations in Colombo is at a period of about 55 secs. which is almost the same as that reported by Duffus and Shand (1958) for auroral latitudes.

Magnetic storms are due to "guests" and "streams" of particles from the Sun impinging on the Earth system. Let us picture a "gust" of particles thrown out by a solar flare spreading out from the Sun and reaching out towards the Earth's orbit. The stream "front" would by geometric broadening alone, if not for any other cause, extend over a width of at least a million kilometers by the time it crosses the 93 million miles from the Sun to the Earth's orbit. The entire earth system having a radius of the order of 50,000 Km would be insignificant in comparison. Due to the effect of the "freezing in" of lines of force, the solar particle will in general not be able to enter the Earth system but would pass round it. The

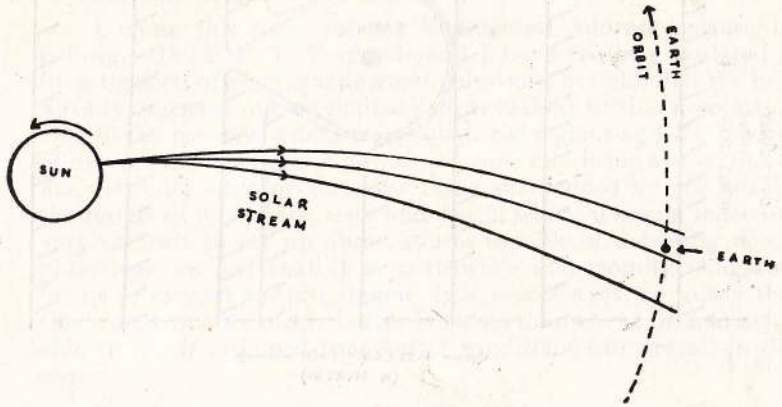


Fig. 11

Diagrammatic representation of a Solar Stream approaching the Earth's Orbit.

sudden commencement of geomagnetic storms within a matter of seconds all over the Earth's surface means that the stream front must arrive abruptly into the Earth's environment, in spite of the fact that the particles must have been emitted with a considerable velocity spread. Singer (1957) suggested that since interplanetary space is not empty, a solar stream must sweep up gas before it, giving rise to a sort of shock front. In Singer's model the velocity of the bulk of the solar gas behind the shock wave will travel at a speed of about three quarters that of the shock velocity, so that solar gas would arrive in the Earth's vicinity, seven to eight hours after the shock itself. The shock wave travelling at about 2000 Km per sec. is pictured to sweep round the magnetosphere of the Earth forming a sort of tear drop shaped cavity having a radius of about

ten Earth radii. This would take less than a minute, and would cause an initial sudden compression of the Earth's field over the entire surface of the magnetosphere. This compression would be transmitted to the Earth's surface via hydromagnetic waves, giving rise to the sudden commencement and the initial phase of the storm.

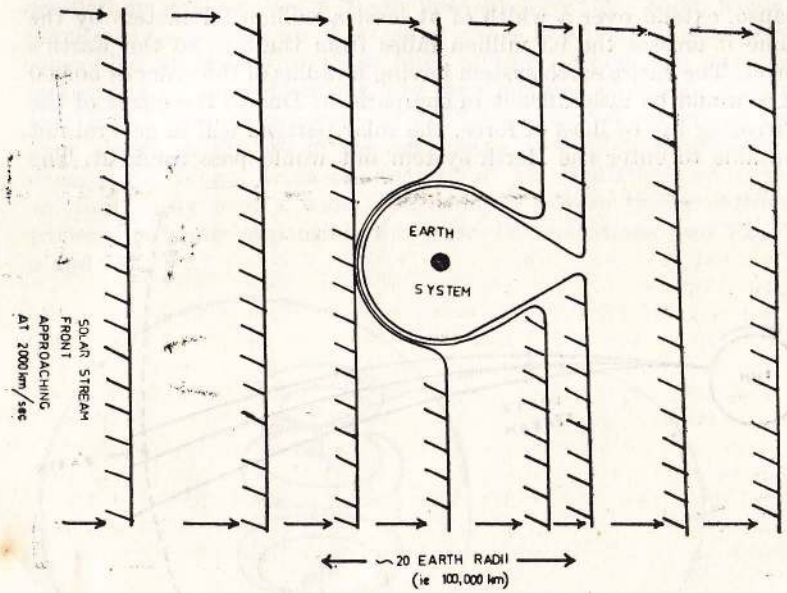


Fig. 12
Solar Stream "front" sweeping past the Earth System.

Singer considers that simultaneously with the shock front enveloping the Earth system, a transient radiation belt is also formed on the periphery of the magnetosphere. The mechanism by means of which the solar particles get injected into such a belt is however not too clear. In a radiation belt charged particles are trapped in closed orbits by the Earth's magnetic field. By Lenz's law the sum total magnetic effect of all the particles moving such a belt must be to diminish the Earth's field. We thus have two opposing effects, the shock front enveloping the Earth system increases the Earth's magnetic field, whereas the simultaneous formation of a radiation belt decreases the field. These competing processes are supposed to account for the different types of sudden commencement, SC (+), SC (−+) etc. seen in different locations.

The main bulk of the solar gas following after the shock wave is supposed to enrich the radiation belt causing the main phase decrease of the geomagnetic field. The belt is supposed to have a

short life time of the order of a day, unlike the radiation belts within the magnetosphere which are more or less permanent. The recovery phase of the storm would then be a manifestation of the decay of the transient belt. The entire picture is still very qualitative. The calculations involved are very complicated and probably much more time and effort will be required to put it on quantitative footing.

Conclusion

I hope I have conveyed to you some idea of the exciting concepts that are emerging out of the study of the transient variation of the Earth's Magnetic Field.

I chose this topic for my Presidential Address because my colleague Dr. P. C. B. Fernando and I have recently initiated an investigation of geomagnetic micropulsations in Colombo. We have already reported our preliminary observations to this Association. Due to the paucity of data from equatorial regions and also because of our peculiar position in Ceylon of, not only being almost on the magnetic dip equator but also being surrounded by the sea for thousands of miles east, west and south, where it would indeed be very difficult to set up observatories capable of detecting micropulsations we feel that it is worthwhile and morally obligatory for us to expand our programme. It is needless for me to say that this would require much better facilities than are at present available to us. It is hoped that better conditions will prevail in due course.

References

Akasofu, S. and Chapman, S. (1961). The Ring Current, Geomagnetic disturbances etc. *J. geophys. Res.*, **66**, 1321-1350.

Baliga, S.P. (1958). The Earth's Magnetism. *Proc. Ceylon Ass. Advanc. Sci.*, **2**, 151-180.

Chapman, S. (1960). Geomagnetic Storms and the Space Around the Earth. *Nature, Lond.*, **187**, 824-827.

Chapman, S. (1961). The Aurora Polaris and the Space Around the Earth. *Amer. Scient.*, **49**, 249-283.

Cowling, T.G. (1962). Magnetohydrodynamics. *Rep. Progr. Phys.*, 244-283.

Dessler, A.J. (1961). The stability of the interface between the Solar Wind and the Geomagnetic field. *J. geophys. Res.*, **66**, 3587-3590.

Duffus, H.J. and Shand, J.A. (1958). Some observations of Geomagnetic Micropulsations. *Canad. J. Phys.*, **36**, 508-529.

Ferraro, V.C.A., Parkinson, W.C. and Unthank, H.W. (1951). Sudden commencements and sudden impulses in Geomagnetism. *J. geophys. Res.*, **56**, 117-195.

Ferraro, V.C.A. (1961). Theory of Geomagnetic Storms. *Encyclopaedic Dictionary of Physics.* **3**, 448-451.

Jacobs, J.A. and Sinno K. (1960). World Wide Characteristics of Geomagnetic Micropulsations. *Univ. B.C. Geophys. Dept. Scientific Report. No.* 2.

Jacobs, J.A. Kato, Y., Matsushita, S. and Troitskaya, V.A. (1964). Classification of Geomagnetic Micropulsations, *J. geophys. Res.* **69**, 180-181.

Jacobs, J.A. and Watanabe, T. (1964). A model of Geomagnetic Micropulsations with long periods. *Canad. J. Phys.* **42**, 200-207.

Proceedings of the International Symposium of Equatorial Aeronomy, 1963. *J. geophys. Res.*, **68**.

Smith, W.H. (1964). Some observations and Characteristics of Type Pc. 1 Geomagnetic Micropulsations. *J. geophys. Res.*, **69**, 1875-1880.

Ward S.H. (1963). Dynamics of the Magnetosphere. *J. geophys. Res.*, **68**, 781-787.

Yanagihara, K. (1959). Some characteristics of Geomagnetic Pulsations. *J. Geomagn. Geoelect., Kyoto*, **10**, 172-176.

INVITATION LECTURE

22nd September, 1964

MAN'S REQUIREMENT FOR PROTEIN AND HIS SUPPLY

by

D. P. CUTHBERTSON,* F.R.S.E.

In my Presidential Address to Section I of the British Association for the Advancement of Science in 1953, I pointed out that the problem of quality and quantity of protein is ever present in our daily speculations about what we shall eat and wherewithal we shall be clothed, and that the scale of our agricultural production to meet our needs has always been a matter for concern. It is increasingly realised that food habits, particularly in respect of the sources of animal protein, have deep roots and sometimes obscure origins which are of great interest and which it is important to assess before any attempt is made to introduce new and better dietary practices. It is sometimes easier to modify the environment than to modify human ways of thinking and behaviour. Suitable and efficient methods of education can only be developed when complete knowledge is available of the reasons why men act in the way they do. But my theme is 'Man's Requirement of Protein and his Supply'. Let us first consider the metabolism of what has been held to contain the magic of life.

Protein Metabolism and its Regulation

The nutritive value to man of his dietary protein is influenced by many variables other than the character and the amount of the protein in his food. Its value is necessarily limited by the nature and the quantity of the amino acids or building blocks available from the proteins of a meal and it depends upon the completeness of digestion of these proteins, and the effectiveness of absorption of the mixture of amino acids into the blood. The completeness of digestion may be reduced by certain dietary factors, while the overall absorption of amino acids is diminished by a high proportion of non-utilisable carbohydrate in the food.

Since the amino acids from one protein may supplement those from another simultaneously present in the food, the mixture of amino acids yielded by a meal as a whole, as opposed to that coming

* Director, The Rowett Research Institute, Bucksburn, Aberdeen.

from the separate proteins it contains, is of importance. The needs of man vary with age, not only in respect of the total dietary protein, but also with regard to the proportion of the essential amino acids which must be yielded from his protein.

Utilisable carbohydrate exerts a protein sparing action, and the total calorie intake must be adequate if avoidable protein catabolism is to be prevented. The task of determining in practice what is the protein requirement of a group of people can be effectually undertaken only in relation to the diet as a whole.

Tissues are differently affected by changes in nutrition. Some gain or lose substance with much greater rapidity than others and this has led to the concept of a labile protein reserve in the body, available for rapid use in emergency. Various workers have referred to it as storage protein, circulating protein, reserve or deposit protein. It would seem desirable to regard labile body proteins as important in the metabolism of body protein and deserving of exploration in their own right as examples of tissue adaptation to amino acid supply. The evidence regarding their nutritional importance is still uncertain without a knowledge of their metabolic significance. It is not justifiable, meantime, to ignore them in assessing protein needs but no decision is possible about the amount of dietary protein which is needed to ensure an adequate store of labile body protein.

Requirements for Protein

The physiological requirement for protein is that which is sufficient to cover all the varied needs of healthy individuals and which enables them to meet both the common and the rare stresses of daily life. In the U.K. we attempted to fix a lower limit of requirement, which would be based as far as possible upon the results of laboratory investigations, and a higher limit based on observation of consumption of protein in situations where there was no reason to suppose a deficiency existed (Ministry of Health, 1964). It was felt that the physiological requirements must lie somewhere between these two limits and attempts can be made to narrow the gap. In defining requirements it must be assumed that the energy input is neither deficient nor excessive and thus to persons who are neither obese nor unusually lean.

We still cling to a rather crude method of expression, namely grammes of protein per kilogramme body weight in place of lean body mass, but this is not easily determined. As I have only but recently come to your shores, I am afraid that most of what I state relates to the mixed protein of the British diet. But in defence of this, the physiological principles involved are not different.

It has been well known that proteins differ in respect of their content and proportion of those amino acids which must be supplied in the diet for adequate growth, maintenance and reproduction—namely the essential amino acids. On the whole protoplasmic proteins in plants and animals do not differ markedly in respect of their content of an appropriate balance of amino acids. In calculating the requirement of protein it has recently become customary to assess this in relation to an ideal reference protein which has a biological value of 100, that is one which exactly matches the need. Requirements may then be converted into minimal intakes of protein in practical diets by using the known biological value of such dietary protein to make the conversion. Our assessment of the biological value of the British diet was of the order of 80.

In addition to fulfilling the needs for the specific essential amino acids there must also be a balance of nitrogenous substances capable of providing for the synthesis of the amino acids which can by synthesised in the body. A source of nitrogen for this is obvious. It would seem simplest if these were preformed in the diet but we have as yet no proper means of determining their optimal proportions to this end.

In our British diets the essential amino acids in the proteins of mixed diets of adults appear likely to be in excess of their requirements and in consequence a deficit, if any, will be in total protein.

A. The Lower Limit

Minimum requirement is used in this report to indicate the irreducible minimum.

Minimum requirements of adults:—For this we are largely dependent on the results of balance studies with a low protein intake. Unfortunately such balance trials are long, monotonous and unpopular, and, in consequence, can only be done on a comparatively few individuals, and individuals differ in requirement.

The FAO's Committee's figure of 0·35 g. reference protein per kg. bodyweight per day for average minimum requirement of adults was based upon the results of balance studies, largely those reviewed by Shermen in 1920. Yet Shermen confessed that the recorded balance studies which were in accord with his criteria were not sufficiently numerous, nor did they cover a sufficient range of foods, to form a satisfactory basis for any general deduction as to the amount of protein required in normal adult maintenance. Further, the balance experiments were made on a mixed diet at low levels of protein intake and generally gave a minimum requirement of protein for equilibrium of the order of 0·5 g. per kg. bodyweight per day. The FAO Committee arrived at the figure of

0·35 g. per kg. bodyweight per day for reference protein by applying a correction factor for biological value to Sherman's value, and to the value obtained by other workers of 0·4 g. per kg. bodyweight per day.

One of the problems is that the nitrogen excretion on a low diet is dependent on the previous nutritional history of the person for, if the protein intake is reduced from a higher to a lower level, breakdown of body protein predominates over synthesis, the "liable body proteins" are depleted, but the person may eventually come into nitrogen equilibrium again. This, however, does not mean that the lower intake of protein is sufficient for the person's needs, for his body now contains less protein than it did before. It is possible that the more labile components of tissue protein which he has lost are part of the functional cell cytoplasm and are essential for complete well-being.

A further method is based on the assumption that the minimum urinary nitrogen excretion is proportional to the basal metabolic rate, 2 mg. nitrogen being excreted per basal kilocalorie. On this basis the contribution of calories from protein to the basal metabolism would be the order of 5 per cent. To this must be added an allowance for faecal loss.

Assessment of minimum requirements does not usually include any allowance for cutaneous loss of nitrogen. This is influenced by the amount of sweating but even under minimal sweating nitrogen equivalent to 0·05 g. protein per kg. bodyweight per day may be lost.

Values for minimum requirements for groups or for populations must be big enough to provide for normal individuals with high requirements. If the mean minimum requirement is raised by something rather more than twice the coefficient of variation that has been observed experimentally the resulting increment is about 50 per cent—the incremental figure which Sherman recommended many years ago.

If these upward adjustments are made to the figure of 0·35 g. reference protein/kg. bodyweight/per day, the adult requirement becomes 0·6 g. reference protein/kg. bodyweight/day. For a protein of biological value 80, which corresponds to the value of the average British diet, the lower limit becomes 0·75 g./kg. bodyweight/day. It is possible that within this range of biological value the limiting factor for maintenance in adults is the total nitrogen and not the amount of pattern of the essential amino acids. If that is so the correction would not be valid.

It is of interest that the FAO Committee reduced the figures of Sherman for mixed diets to express protein in terms of reference protein. The two corrections almost cancel one another and the validity of either does not affect the end results.

Minimum requirements during growth

The protein intake of a breast-fed infant is 2-2·5 g./kg./day and this meets maintenance and allows for growth. This is accepted as being near to the physiological requirement. If every physiological function were completely adequate there would be no room for genetic improvement. In the case of older children it seems appropriate to estimate the requirement in relation to the basal energy expenditure, and for growth by calculating the probable increments of protein in the body. In our country from one year onwards the addition of protein to the body accounts for only a small fraction of the protein intake of children—and from the evidence available it seems that in well-nourished children after the first year the requirement for maintenance is considerably greater than the requirement for growth. The growth spurt in adolescence involves such a small increment as to be practically negligible as regards an extra rise in intake.

For practical purposes the protein requirement for the first year of life is 2 g. of reference protein (ideally balanced protein in terms of amino acids) per kg. bodyweight per day at 2 months, tapering off to 1·5 g. at 1 year. For older children it seems reasonable to diminish from 1·2 g. reference protein to 0·7 g. of such protein at 17 years. In terms of the mixed U.K. dietary this is a change from 1·5 g. to 0·9 g./kg. bodyweight per day.

Pregnancy

In the U.K. it is held that the average normal pregnant woman lays down 0·5, 3·0, 4·5 and 5·7 g. protein during the successive quarters of pregnancy. The efficiency with which protein is utilised in forming new tissue protein during pregnancy is not known. If it be assumed that it is of the order of 50 per cent, then the requirement for additional protein during the second half of pregnancy would be about 10 g. reference protein or 12·5 g. mixed protein of the U.K. pattern per day.

Lactation

The FAO Report estimates the average protein in milk secreted by women to be about 10 g. daily. If the efficiency is assumed to be 50 per cent, then the average additional requirement would be 25 g. daily of good mixed protein.

Old age

No special physiological demand which warrants a requirement over that of younger adults has emerged.

Serious infections and injury

The demand for extra protein as a result of disease and injury is three-fold:

(1) for regeneration, as far as is possible, of tissue lost, and for the healing of wounds;

(2) for the replacement of protein in body fluids lost either immediately or later;

(3) to make good protein loss through the catabolic response to infection and physical injury.

The last is the most important. Serious injury can impose on the patient a need to replace protein in amounts of one kilogramme or more and in the long run this may all have to be made good from the diet. The overriding consideration is the need to restore the patient to health as rapidly as possible. First blood or plasma loss should be made good. It is generally believed that the clinical and metabolic evidence indicates that during the first few days following serious injury in the previously adequately nourished person, it may be unwise to push the intake of a well-balanced diet beyond appetite and this may fail during the first few days. Thereafter the patient should be encouraged to take as much as he can of a well-balanced diet relatively rich in protein. However, when appetite fails or the unassisted intake is inadequate, there is a case for increasing the intake by fortified supplements and, if necessary, by tube feeding. It must be remembered that protein can only be used effectively for synthesis if the energy needs are covered. A high protein intake may be necessary for months after extensive injuries, such as a large burn. A daily intake of 3 g. or more protein per kg. bodyweight is suitable for burned children. For adults an intake of 2 or 3 g. per kg. bodyweight is appropriate.

As the evidence for a possible relationship between protein intake and resistance to infection is contradictory and generally complicated by a deficiency also in energy as well as protein, infection can be properly regarded as a frequency stress of daily life and probably taken care of by the apparent margin over apparent minimum requirements.

B. The Upper Limit

Intakes by adults:—The average daily intake of protein by adults in the U.K., excluding the elderly and the pregnant, is of the order of 90 g. daily for men and 70 g. for women, or 1·4 and 1·3 g./kg. bodywight/day respectively. No cases of primary protein deficiency are found there and it would seem probable that these mean figures are above the average physiological requirement but it is impossible to state by how much. But it is important to note the relative constancy of the proportion of energy in the U.K. diet represented by protein ranging from 11-12 per cent. In the elderly it may rise to 14 per cent. Actually

the mean adult level in other regions and subregions is not so very different as will be seen in Table I where the range is from 9-13·6 per cent.

TABLE I

The Contribution of Protein to Calories in National Food Supplies

(Calculated from "Second World Food Survey", FAO., Rome, 1952)

Region and Sub-region	*% of Calories from Protein*		*Animal Protein as % Total Protein*	
	Range	*Average*	*Range*	*Average*
Far East				
South Asia	7·9-10·3	9·5	13·6-21·2	16·7
East Asia	9·2-12·4	11·0	9·5-11·0	10·2
North-East Asia ..	—	10·2	—	15·1
Pacific Islands and Malay Peninsula	—	9·0	9·5-22·8	16·1
Near East	9·0-13·7	12·3	13·3-36·0	20·1
Africa				
Northern Africa ..	—	13·6	—	24·6
Central and Tropical Africa	8·7-11·4	10·2	11·9-24·2	17·1
Southern Africa ..	8·3-11·8	10·8	19·5-40·0	30·0
Latin America				
River Plate Countries ..	12·8-14·5	13·4	55·2-67·4	60·5
Rest of Latin America ..	9·8-12·2	10·8	27·0-46·5	36·1
Europe				
Western Europe ..	11·6-14·2	12·2	40·5-51·5	48·6
Northern Europe ..	11·9-14·0	12·8	46·4-66·5	56·8
Southern Europe ..	12·5-12·8	12·7	22·5-33·3	27·7
Germany and Austria ..	10·8-11·9	11·5	27·5-38·4	34·0
Eastern Europe ..	10·7-12·6	11·7	23·2-30·5	27·6
Union of Soviet Socialist Republics	—	12·9	—	25·8
North America and Oceania	11·5-12·0	11·8	62·0-68·5	66·7

The data indicate the range in the proportion of energy in the diet contributed by dietary protein that is encountered in various countries throughout the world, and may be contrasted with the relative constancy of this proportion in British diets.

This constancy in the proportion of protein is irrespective of level of energy expenditure, pregnancy or lactation. It is not generally accepted that urinary excretion of nitrogen is raised by

muscular exercise *per se*, but by the increase in muscular mass associated with athletic training and the maintenance of the increased musculature. The dermal loss of nitrogen rises when there is much sweating and this may be of practical importance with a low protein intake in a hot climate.

Other dietary constituents

Unavailable carbohydrate or other indigestible material reduces the apparent absorption of protein by causing an increase in the amount of the intestinal secretions, in the weight of the stools and in the rate of passage of material through the gut. The amount of introgen lost in the faeces depends much more upon the amount of unavailabe carbohydrate in the diet than upon the amount of protein.

Protein value

The "protein value" is thus related to the quantity and quality of the protein in the food consumed by the human being under defined conditions.

Judgement of protein quality must ultimately be based on biological tests, preferably on Man. In principle the protein quality should be measured by the amino acid composition: a score based on the amino acid composition of the important foodstuffs provides a simple and widely used means of assessing the quality of a diet. The use of such a measure requires a standard of comparison, hence the concept of the reference protein developed by the FAO Committee (1957). The reference protein is a hypothetical protein containing 16 per cent of nitrogen which can be completely used for body synthesis. The pattern was based on the knowledge available at that time, but later evidence suggests that the problem of defining a pattern of amino acid requirements is extremely complicated since the pattern may vary at different ages. Further, it cannot be taken for granted that the pattern of the unessential amino acids, to which the FAO Committee gave no attention, is unimportant.

If it were possible to reach agreement on an internationally accepted reference composition(s) of protein(s) such would have undoubted value. Currently no better suggestions have been made.

In general, the amino acids most likely to be limiting are lysine, the sulphur-containing amino acids, and tryptophan. It may be enough to base a system of evaluation on these alone the rest being automatically taken care of.

Effect of processing on availability of amino acids

Domestic cooking and commercial sterilisation in cans do not appear to result in significant damage except in a few special cases such as the "explosion" of breakfast cereals and in the roasting of meat. Even in roasting the excessive heat is applied to a very limited portion of the meat. Baking can cause a fall in nutritive value—essentially through loss of lysine—and is due to the Maillard reaction. Addition of reducing sugars to cereal products before baking enhances this reaction. Several reports suggest severe reduction in nutritive value in baked products which have been fortified with dried milk.

It has been shown repeatedly that most diets are limited by the sulphur amino acids so that loss of lysine in baking need have little effect on the nutritive value of the whole diet. There is little loss of nutritional value in the preparation of ordinary human foods other than that caused by baking.

The same is not true, however, of certain materials such as fish meal and oilseed cake, which are sometimes used as human food, though more frequently fed to animals. It is possible to process these foods without significant damage if sufficient care is taken. Heat and storage can, however, result in a loss of methionine and lysine of fish meal and certain oilseed cakes. It is not always possible to restore these by supplementation.

It is equally well recognised that certain foods, in particular soya beans and legumes, improve in nutritive value on heating. This in the case of soya beans is due to destruction of a heat-liable trypsin inhibitor(s). Excessive heating reduces nutritive value.

However good the quality of a protein, the requirement will not be met if the quantity is insufficient. Since intake is dictated by calorie requirement, the proportion of protein in the diet is important. Both aspects, quality and quantity, are combined in terms of utilisable protein calories per 100 total calories from the food. It remains to be seen whether this is preferable to expressing protein in terms of grams per day, since it is not accepted that physical exertion *per se* leads to an appreciable increase in nitrogen loss, provided that the energy remains sufficient. In the U.K. we feel that some reservation must be extended towards any proposal that protein requirement should be related to the total calorie requirement.

Requirements

The gap in our U.K. knowledge is bounded at the lower limit as we have seen by what can be derived—and derived with much uncertainty—from the results of balance experiments and kindred information, and at the upper limit by information about the average dietary intake in our country. Somewhere between these imprecise limits lie physiological requirements. Other countries may find this pattern of approach useful but in many of the developing countries the average dietary intake may be below the physiological require-

ment assessed by a developed country. This must obviously be so where there is chronic protein undernutrition in a substantial proportion of the population.

If we were more certain that labile body protein has a function and could assess the amount and the protein intake required to maintain it, then we could raise the lower limit and so narrow the gap. As this is not the case, the only means of narrowing the gap is by lowering the upper limit. But before this is done we, in Britain, must consider:

1. Protein intake might, in some measure, be based on instinct. Were this so, those with the lowest intake might be those with the least need. We know no evidence of this.
2. Those who eat least protein might be very small people. But in our country protein intake is largely bound to energy intake and the latter is not closely related to body size.
3. The lower intakes recorded during the period of survey (generally a week or less) may be offset by higher intakes at other times.

For such reasons we have not considered it wise to depart from the average figures based on protein intake and have therefore not lowered the upper limit to the gap. Thus our findings can be summarised in Table II.

TABLE II

Requirements in terms of the mixed protein of the British diet

Summary of Ministry of Health Working Party's calculations (1964)

Infancy	0-6 months	About 2 g. per kg. bodyweight per day (of human milk protein)	
	6 months-1 year	About 2 g. per kg. bodyweight per day (of protein of British diet)	
		Lower Limit of knowledge	*Upper Limit of knowledge*
Children	4 years	1·6 g. per kg. of bodyweight per day	2·9 g. per kg. of bodyweight per day
	12 years	1·2 g. per kg. of bodyweight per day	2·0 g. per kg. of bodyweight per day
Adults	Male	0·75 g. per kg. of bodyweight per day	1·4 g. per kg. of bodyweight per day
	Female	0·75 g. per kg. of bodyweight per day	1·3 g. per kg. of bodyweight per day
	Pregnancy (2nd half)	Add 12·5 g. per day	
	Lactation	Add 25·0 g. per day	

Recommended Allowances

In setting targets of nutrient supply for countries there are undoubtedly hazards in the use of an appropriately weighted average of the requirements of different sections of the community. People eat foods not nutrients and the amount of protein consumed by adults is determined to a very considerable degree by their calorie requirement. Is this constancy of the protein intake as a proportion of the total calories an indication that man needs protein to form an optimal metabolisable mixture or is it simply a reflection of the constant proportion of protein in all the cereals, or other staples, which bulk largely in the diet? Inspection shows that the latter can only be a partial answer. It rather looks as if man has tended, and still tends, to take out of the common pot according to his energy needs, and through the ages the composition of this in terms of nutrients has also been adjusted to meet his maximal demands for growth and reproduction. If this be so, the safety margin in developed countries is certainly wide for the adult man and non-pregnant woman receive relatively too much protein in relation to their needs but without harm, and may benefit in northerly latitudes from its specific dynamic action in cold weather. We realise that any other natural arrangement for the family would have been unlikely to survive, for no special thought or care in respect of the total amount of protein need be taken in view of such an apparently well-buffered system, provided the energy supply is sufficient and the protein is derived from several sources. But cultural and religious observances may upset such an apparently protective arrangement and the vulnerable groups denied the protective necessities.

To determine the protein requirement of a country or community it is generally first necessary to arrive at an appropriate weighted average of the energy requirement and to stipulate that the proportion of protein in the diet shall be such that when it is consumed by children in amounts necessary to satisfy their calorie requirements, the quantity of protein the children obtain from it shall be sufficient for their needs. However, even this might not protect children whose energy expenditure was lower than average, but targets so high that even these children were protected might well prove beyond the resources of countries whose protein supplies are low. It is in such countries that unfortunately education may not be in a state where food is likely to be distributed within the family in accord with physiological need for protein.

Food Technology

It is important in developing countries to consider as potential sources of dietary protein all products, both of animal and of vegetable origin, which can be locally produced. Today non-industrialised countries in the tropical and sub-tropical regions, because

of their climatic characteristics, soil composition, available facilities for preservation, transportation and processing of foods, economical structure and cultural factors, could more easily develop their agriculture for the production of the needed protein from vegetable sources rather than develop, to a sufficient extent, their animal husbandry, and the facilities required for an adequate utilisation of its products. In fact, in many regions where sources of animal protein for human composition are scarce, products of vegetable origin are already available and can, if properly utilised, go a long way in fulfilling protein needs. Many of these products are sub-products of industries already established, which use locally produced raw materials, oilseed cakes being an example. This has obvious economic advantages.

Can proteins of vegetable origin replace those of animal origin in human nutrition? Although most proteins of animal origin have a satisfactory amino acid pattern and therefore a high biological value, this is not true in every case. On the other hand, most proteins of vegetable origin have a low biological value because they are deficient in one or more essential amino acid and sometimes because of inhibitors to their digestion by the alimentary enzymes. If, however, such deficiencies are corrected either by combining different proteins in appropriate ways or by adding the missing amino acids, vegetable proteins cannot be distinguished nutritionally from those of animal origin. However, it must be noted that animal foods are more likely to be sources of other nutrients such as vitamin A and vitamin B_{12}. Another important difference between natural products of vegetable and animal origin is the amount of protein they contain. In general, the concentration of protein in most foods of vegetable origin is too low to make them, in their natural form, adequate sources of protein, particularly for young children. They have to be concentrated into the form of a high protein supplement with adequate biological value.

You have a new Bureau of Agricultural Economics in Ceylon and I trust that it will be able to help in showing how to lower the costs of cattle and poultry keeping so as to be able to sell meat, milk and eggs at prices within reach of the poorer people of your country. One of your dilemmas is that, whereas in recent years you have become apparently self-sufficient in egg production and table birds and no longer are forced to import millions of eggs from India, the price of eggs cannot be reduced to any extent because of the price of poultry feed. It would be tragic if this recent expansion should recede into a depressed state. It would seem especially necessary to lower the cost of poonac, bran and fish meal.

I understand also that you have still to import millions of pounds of dried or condensed milk. All these increases are needed to supply a more protective diet than apparently exists at present. Your daily per capita average energy consumption is 1790 Cal. compared to 1710 for India and 2165 for Japan (years 1952-53) and would seem thereby to preclude an intensive productive existence.

To balance inadequacy in protein supplies the resources of food science and technology have been investigated in many developing countries and work of special importance is being carried out or is under way for the development of low-cost protein-rich foods using mostly by-products of food processing industries as starting material. Efforts in this field have been supported and stimulated by FAO, WHO and UNICEF. The progress achieved so far is considerable, as indicated by the number of possible raw materials already investigated (e.g., fish, soybeans, groundnuts, cottonseed, seasame), and by the several food mixtures and products developed (e.g., Incaparina, fish flour, fish sausages, groundnut flour, cottonseed sausages, soyflour and soymilk) which have undergone successful laboratory and field trials. Certain of them, such as Incaparina and fish sausages, have already reached the stage of industrial production and commercial distribution. It is very necessary to avoid affecting the biological value of the processed protein to be preserved.

INVITATION LECTURE

23rd September, 1964

THE ACHIEVEMENT OF CHEMICAL ENGINEERING IN THE FIELD OF DISTILLATION AND RECTIFICATION

by

A. G. Evstafiev*

The importance of the chemical industry especially in more recent years is being realised more and more; and the role of this industry is being enlarged for the rapid development of other industries. At present it is almost impossible to find any industry which is not using chemical products. Nowadays the use of chemistry is implied everywhere : fertilizers, fungicides, insecticides in agriculture, tyres for the auto cars and nearly all the parts of electronic equipment, also for building materials and materials for the production of industrial equipment and machines.

It is a known fact that raw materials and intermediate products are mixtures of various components and that very often the final product is a mixture of its by-products. Any person who is in touch with chemistry even to a minimum extent also knows that each chemical complex has in it the means for the separation of the mixture. These means can be found in laboratories as well as in the chemical industry. It is especially important to get a chemically pure component from the mixture in the production of polymers with special properties, semi-conductors, isotopes etc.

Among various methods of the separation of liquids or gas mixtures for their individual components, the most widely being used is rectification. It is interesting to note that rectification for a long time was an art of wine producers and only in the second half of the 19th century the development of physical chemistry marked the beginning of scientific studies of this process of chemical technology. A great contribution to the study of the nature of solutions was made by the Russian scientists Konovalov and Vrewskiy. Konovalov's laws laid down the beginning of the quantitative study of the process of distillation (Konovalov, 1884). These laws served as a ground for the creation of the quantity equations of the theory of equilibrium of binary mixtures by Vrewskiy (1910), Duhem (1889), Margules (1895), Brawn (1897) and other

* UNESCO South Asia Science Cooperation Office, N-1, Ring Road, New Delhi.

investigators. The first approach to the calculation of the components in the column for the ethylalcohol water system was done by Sorel (1893). After that study of the rectification process was carried out in four directions:

1. summarising the practical data on the equilibrium of various mixtures and theoretical investigation of these practical data on the basis of classical thermodynamics and, in recent times, with the help of the irreversible thermodynamics (Evstafiev, 1963)

2. creation and perfection of the methods of calculation of rectification columns for binary and multi-component systems

3. investigation of the efficiency of columns of different designs which include :

(*a*) investigation of column hydraulics i.e. the study of factors which influence on the stable working of the column in different hydro-dynamic regimes for different constructions of the contact device for the vapour and liquid ; and

(*b*) investigation of the mechanism of the process of mass-transfer in the column apparatus and obtaining empirical equations for kinetics of the process.

4. Investigation of the problems of automatic control and regulation of rectification columns.

All this wealth of knowledge about the process of rectification helps investigators to design and to explore the rectification columns in a more rational way.

In view of the development of the chemical industry and its conversion into a mass production industry, the creation of economical equipment and its exploitation has assumed a place of major importance.

Further let us take separately the problems which confront the designers of new rectification equipment and those who are exploiting existing equipment. The designers are facing first of all the problems involved in the enlargement of output of new columns. This aim can be reached by (*a*) increasing of the speed of the substance (*b*) by adding to the number of units of similar apparatus or by enlargement of the size of the apparatus. If we take :

W as output of the plant, it will follow in general that

$$W = w F$$

where w — speed of the stream of the substance ; and

F — cross section of the apparatus.

It is clear from this equation that the enlargement of the output can be achieved either by increasing the speed of the stream of the substance (in the case of rectification, this is the speed of vapour in column); or by enlargement of the cross section of the apparatus. Considerable increase in the speed of vapour in the columns (10 to 15 times) can obviously be achieved only if one uses an entirely new construction or design of the internal packing of the columns. Such an enlargement of the speed in a usual film or plate column is impossible because it leads to considerable carrying away of the liquid phase by vapour and to the over flow of the column. This phenomenon leads us to the need for creation of an entirely new type of rectification column. Professor P. Simenov (USSR) has suggested parallel current in the contact elements of the mass-exchange apparatus instead of counter current or cross current which usually are used for this apparatus. In the new columns the vapour having high speed (30 meter M/sec.) rises up through the pipes and takes the liquid with it. It takes along with it the liquid phase by ejection. The liquid in the form of very thin film rises up on the internal sides of the walls of the tubes. In the upper part of the contact element the liquid is being separated from the vapour and the vapour goes up into the next contact element for the ejection of a new portion of the liquid coming from the upper element. These columns are quite efficient and their efficiency is about 10 times that of the bubble cups column. But because, of necessity, there is a large amount of vapour per unit of liquid phase, these columns can not always be used for rectification since the proportion between vapour and liquid in the rectification is being determined by conditions of masstransfer.

The second way to enlarge the output is to enlarge the cross section of the column which could be implemented either by enlargement of diameter of the column or by enlargement of the number of parallel working units of apparatus. It is easy to show that the simple increase in the number of columns will be more expensive than the enlargement of the size of a single unit.

The cost of a rectification column can be estimated approximately as being proportionate to the size of the column (its weight). If we take :

N — the cost of column

$$N = m\,G \qquad (2)$$

where m — proportional coefficient

G — weight of column

Approximately the weight of column G is proportionate to the two main dimensions of the column : its height and perimeter of cross section i.e.

$$G = K\,L\,P \qquad (3)$$

where P — summary perimeter of cross section parallel working columns.

L — height of the column

Then

$$P = n\,\pi\,D \qquad (4)$$

where D — diameter columns

n — number of parallel working columns for given value of F

D = f (n) i.e.

$$D = \sqrt{\frac{F}{\pi}\,\frac{4}{n}} \qquad (5)$$

if we apply (5), (4), (3) in equation (2) we will have

$$N = m\,K\,L\,n\sqrt{\frac{F}{n}\,\frac{4}{}}$$

or

$$N = m\sqrt{n} \qquad (6)$$

from this equation one can see that with the enlargement of the number of columns (n) the cost of the column will rise if it is assumed that the output of the column will remain the same.

From this equation, for instance, one can see that if it is necessary to enlarge the output up to 10 times in this case to have 9 additional columns similar in size to the original one will be 3 times more expensive than to have one bigger column. In addition to this it is necessary to note that if we have one bigger column instead of 10 small columns it will follow also that one bigger column will be less expensive in communication lines, auxiliary equipment and instruments and that it will be easier to handle. The reasoning above shows that if there is need in a mass production plant, it appears more advantageous to enlarge the diameter of column than to increase the number of columns. But the enlargement of diameter of the column cannot be implemented without considerable change of construction of the contact devices; because when one enlarges the diameter of the column, factors which will be invisible in the course of work of the normal column with a small diameter (1-2 meters) will start to play an important role. For instance, with the enlargement of the distance through which the liquid must pass on the plate bubble type, or sieve type, it

appears that a considerable hydraulic gradient may lead in its turn to unevenness of the work of the plate i.e. enlargement of diameter of the column leads, (in the same way as enlargement of the speed of vapour), to the need for creation of a new construction of the contact devices. In view of this it requires quite a serious investigation of the mechanism of the masstransfer in the columns; and to this task much work has been devoted by scientists during recent years.

Those who are exploiting existing equipment are facing the problem of intensifying in the most economical way the existing columns. In recent years scientists have been very seriously engaged in the problems involved in the choice and maintenance of the most appropriate regime in the work of columns. It is known that the purity of the product is determined by a ratio between the streams of the liquid and the vapour phases. Upon this ratio will depend the expenditure of heating vapour and the expenditure of the cooling agent. Taking into account the fact that the process of rectification requires quite a large amount of energy and that the main item of expenditure is the cost of the heating vapour, it is important to note that the choice of the ratio between the streams not only is a factor which determines the purity of the product but is also a factor which determines the cost of the main item of expenditure. Frequently the ratio between the streams of the vapour and the liquid phase have been characterised by the value known as the portion reflux :

$$R = \frac{L}{V}$$

R — portion of the reflux

L — quantity of the liquid flowing down the column.

V — the quantity of the vapour rising up on the column.

During the examination of the industrial benzol columns it was established that the dependence on the quality of the product on the quantity of the reflux in different quantities of output of the columns is also different (Evstafiev, Zikov and Karavaev, 1959). At the same time it was noticed that the positive influence of R on the quality of the product can be seen only up to the limit of the value of R. After this, further enriching of R does not give any notable improvement of the product and leads only to the overloading of the column and to additional expenditure on the heating vapour in the boiler column and to additional expenditure on the cooling water in the condenser. So an optimum R can be used during the exploitation of the column. This value of R will be such that its further enrichment will not give notable improvement of the product. It is also

established that with R less than optimum the column will be more sensitive to the variation of R and that the column will be less sensitive to the variation of R if the value of the R is more than optimum.

As a result of this the stable hand-regulation of the column is possible only if $R > R$ optimum i.e. with uneconomic regime and with over expenditure on the heating vapour. The stable regulation of the column, if $R < R$ optimum or $R = R$ optimum, will only be possible if one uses the automization of the column. So automization permits conduct of the column in the most economical way.

References

Brawn, F. (1897). *J. chem. Soc.*, **35**, 547.

Duhem. (1889). *Ann. de l'ecole Normale* Sup. **4** and **6**, 153.

Evastafiev, A.G. (1963). *Rektifikaticioni Oostanovki*. Moscow: MASHGIZ.

Evastafiev, A.G., Zikov, D.D. and Karavaev, N.M. (1959). *Trud.*, MIXM, **19**.

Konovalov, D.P. (1884). *Jurnal Rooskavo Physiko-khimicheskovo Obshectva*, **16**.

Margules. (1895). S.B. Akad. Wiss. Wien. **104**, 1243.

Sorel, E. (1893). *La restification de l'alcogole.*

Vrewskiy, M.S. (1910). *Jurnal Rooskavo Physiko-khimicheskovo Obshectvam.*

LECTURE BY CHIEF GUEST

22nd September, 1964

POPULATION GROWTH AND ECONOMIC DEVELOPMENT

by

FRANK W. NOTESTEIN*

Mr. President, I am particularly gratified with your invitation to address this distinguished Association. To me the invitation symbolizes the strides we are making in bringing, through science, the unified world of knowledge to the service of mankind. It does so in a number of ways. As a social scientist, I am pleased to be invited to participate in activities of an association that represents mainly the natural and physical sciences. I have been a Fellow of the American Association for the Advancement of Science for many years, and the recognition accorded by it to the social scientists is also growing rapidly. So is the recognition given to social science, and particularly to the field of population studies, by our National Academy of Sciences, which only last year published a pamphlet dealing with population growth. Two years ago, similarly, the Royal Society devoted a meeting to problems of population, and three weeks ago the Presidential Address to the British Association also dealt with the subject. My invitation to speak to your distinguished Association is, I trust, another indication that the time has passed in which we can limit the scope of subject matter to which scientific procedures may be fruitfully applied.

Also, I am particularly glad to be asked to discuss a subject fraught with ideological overtones and sensitive perhaps beyond others because it deals with the most intimate aspects of personal behavior, and because it is historically a vortex of religious disagreement and ideological controversy. I am delighted to accept this difficult assignment because it means that you and I can demonstrate our mutual belief that advances in scientific knowledge have minimized the relevance of the traditional disputes. But I hope your generous invitation also illustrates our mutual commitment as scientists to the service of mankind. If so, my topic, *Population Growth and Economic Development,* could scarcely be more appropriate because a human population that is healthy, educated, and prosperous, must be the goal for all of us—the reward of our search for truth in all of its manifestations.

* President, The Population Council, New York.

I wonder if those of you who have had little occasion to study the matter realize how bitter the disputes have been in the population field, and how far they have subsided during the past decade. This is not the time or place for a systematic discussion of the subject. A few observations may be useful, however, if only to illustrate the fact that deep disputes are seldom won by one side or another. Resolution comes eventually because advances in our knowledge make the issues substantially irrelevant to all save the fanatical fringe.

In recent centuries there have been three major sources of world conflict on demographic matters. One involved Malthus and Marx, or, more accurately, the formulations of the more doctrinaire adherents of the Malthusian and Marxism points of view. Another involved the competition for space and resources between the poor and the prosperous nations, centering on the pressure to migrate and the restrictions against migration. The third involved religious and moral questions related to sexual practices and the regulation of human fertility.

On all these matters, disagreement has been strong, tempers have been hot, and discourse amongst the various schools of thought has been at best strained. Today, there are still difficulties but these difficulties are minimal.

Let us look at the issues. First are those concerning Marxism and Malthusianism. Poverty to Malthus was mainly the result of man's insufficient restraint in reproduction, while to Marx it was mainly the result of exploitation and unfortunate economic institutions. Neither gave much attention to the problem that preoccupied the other. Through more than a century international discussions about population were blocked by the rigidity of these competing theories. Now the disputes have largely disappeared, so far as obstacles to action are concerned. The dramatic control of disease and the equally dramatic release of productive power by modern technology have clearly shown that both Marx and Malthus were pointing to real problems, and that neither could realize the full power of the forces involved. Today, our understanding of the determinants of both population and economic growth has progressed to a point where both Marxists and non-Marxists analyze the interrelations of population and economic change in essentially the same way.

Less progress has been made in the resolution of differences between the prosperous and the poor peoples of the world. Perhaps this is so because these differences continue to exist to an extent that shocks mankind's sense of justice. But even here there has been a change. The differences are no longer phrased in terms of the right to migrate. Even the poorest country realizes that migration

as a solution for population growth, although it may help the individuals who leave, will not help develop the economy for those who stay. Moreover, all of us realize that no manipulation of population alone will solve the problems of poverty. The problems of poverty ultimately will be solved only if there is also increased production.

Religious disputes similarly have been greatly reduced. If one is dealing with the intellectual leaders of the world's major religions, there is rather little differences in the interpretation of the economic meaning of modern population growth. Differences, where they occur, relate mainly to matters of emphasis and to the means by which it is deemed proper to regulate childbearing. Even here, advances in our knowledge of the physiology of reproduction give promise of resolving the problems that remain outstanding.

It is clear that the world's major disputes about population which have been viewed as differences of principle have in fact arisen mainly because of the inadequacy of our knowledge. Knowledge alone cannot solve all problems, but the advance of knowledge does narrow the range of disagreement. This is particularly true if we allow in the house of science the study of man and society as well as the study of the inanimate world.

Thus far, I have dealt in the briefest possible fashion with simple assertions about the ways in which advances in knowledge have served to make our traditional controversies obsolete. It is now time to start documenting the case. I shall sketch, firstly, this modern and generally-accepted view of the processes of population change and, secondly, the principal interrelations of population change and economic growth. Finally, against this background, I shall indicate some of the responses being made by various countries to their demographic situations. Let us turn immediately to a brief account of the principles of population change as they are generally understood today. The statement will, of course, be phrased in terms of my own orientation. I believe, however, that in its essentials it would be agreed to by major elements of all the world's principal ideological groups.

The modern era of rapid population growth, unique both in terms of speed and magnitude in mankind's history, has its source in the more rapid response of the death rate than of the birth rate to the modern development of science and technology. This difference in the responsiveness of death and birth to the events of the modern era in turn has its source in universal differences in the ways in which the normative orders of all societies impinge on human fertility and mortality. Moreover, it seems likely that differences thus universal have been essential for the survival of the human race.

This is the general statement. Let us now endeavor to make it concrete. Throughout most of mankind's history, indeed until very recently in most of the world, survival rather than over-rapid growth has been the problem. Man has always sought life and health by every means at his disposal, but until a moment ago in history he had limited success. It is hard now for us to realize this fact. Throughout the ages, up to a little more than a century ago, the new-born child anywhere probably had less than an even chance of surviving to age 30. In Ceylon, the fifty-fifty chance of survival was at about age 30 as recently as 1921. Throughout most of mankind's history this median length of life was probably less than 20 years. Moreover, such figures as 20 and 30 years were the values for the "normal" and "healthy" periods. In times of war, famine, and raging epidemic, the depletion of population by death was catastrophic.

The median length of life in Ceylon, which was about 30 in 1921, is now something like 70 years. The change came later and more rapidly here than in the Western world, and it has not gone quite so far. In the world's best experience, the median length of life is now about 75 years, and the arithmetic mean, or life expectancy at birth, is about 70 years. Loss through death, which throughout the ages has been inevitably high, is now controlled with an efficiency never dreamed of even a few short decades ago.

The source of this change in death rates is not hard to find. It is two-fold. On the one hand, it lies in the material prosperity made possible by the application of modern technology to agriculture, commerce, and industry, and on the other hand, it results from the spectacular effectiveness with which modern preventive and curative medicine control disease.

It should be noted that death rates have fallen responsively to advances in our knowledge because people always and everywhere prefer health to sickness and life to death. There have been no durable obstacles to the control of disease once the medical resources became available. Obviously, there has been great inefficiency, and on occasion there has been reluctance to move ahead on various grounds. But the obstacles have not lasted long in the presence of popular understanding of the effectiveness of particular courses of action. On the whole, death rates have fallen rapidly in response to new knowledge concerning the control of disease.

The case of human fertility stands in sharpest contrast with that of mortality. Fertility had to be sufficiently high to produce growth to people the world, in spite of the appalling mortality of the "normal" years and the periods of war, famine and epidemic. It is not surprising that the growth associated with peace and prosperity has through the ages been considered "good". "Bigger"

was not to be differentiated from "better" when the balance of survival was precarious. Inevitably, as a requirement of survival, the populations that inhabited the world's major land masses entered the modern era with both the physiological capacity and the social institutions required to elicit high rates of reproduction.

All major populations entered the modern era with high reproductive capacity, and do not, in some mysterious way, lose that capacity as they become more prosperous. J. de Castro (1952) has suggested that the rapid rate of reproduction in the world's poorest regions comes because dietary deprivations somehow lift human reproductive capacity, but this view simply does not square with the facts. There is ample evidence that the well-nourished, even over-nourished, women of the Western world's middle classes have very high reproductive capacity. Large numbers of studies have shown that when they are not practising contraception they become pregnant at rates that, if maintained, would yield birth rates quite as high as those found anywhere in the world past or present: rates substantially higher than those in Ceylon, India or China.

We may take it as proved that the high birth rates at the beginning of the modern era, and of the poor nations today, do not reflect enhanced reproductive capacity. If anything, the contrary is true. Poverty is not consonant with vigorous health. It is probable that in the poor nations the prevalence of anemia and fevers produces rather high rates of spontaneous abortion. Presumably the absence of these factors in the more prosperous nations has given their populations higher rather than lower capacity to reproduce. To sum it up simply, all populations entered the modern era with the capacity for high rates of reproduction and all have retained that capacity.

We must seek the source of the high birth rates of the newly developing countries, both two centuries ago and now, not just in their physiological conditions but in their social institutions. Without social institutions that fostered rapid reproduction populations could not have survived the ages of high mortality. It is sometimes supposed that the highest of all birth rates would occur in the virtual absence of social constraint, so that reproduction would be governed on almost an animal level of behavior. There is clear evidence that this is not true. The debauched society is too unstable, too insecure, and too unhealthy to yield maximum birth rates. The highest birth rates come characteristically from the traditional self-sufficient agrarian society that typified most of the world before the modern era.

The specific institutional arrangements by which high rates of reproduction are fostered vary greatly from society to society, but the functions served are the same. Marriage customs, family

organizations, property systems, the means of attaining status, the systems of community rewards and sanctions, educational processes and the religious doctrines are all organized in ways that promote nearly universal and fairly early marriage and high rates of marital reproduction. These institutions, customs, attitudes and beliefs are deeply rooted in long traditions. They represent the moral code, the normative order, which provides the non-rational cement of loyalty that binds individuals into groups and binds the past to the present. Virtually by the definition of a viable society they are slow to change.

The source of the present epoch of population growth is thus clear. The normative order of traditional societies favored both good health and high birth rates. Reduction of death rates required a change of means but no change of motivation. Reduction of birth rates requires changes of both means and motives, and the latter can be neither instantaneous nor frictionless, if life is to retain its coherence. Under the impact of the modern world, death rates fell but birth rates remained high pending the obsolescence of age-old customs and beliefs. The epoch of growth was under way.

Eventually the same science and technology that brought the reduction of mortality also transformed life in ways that were highly subversive to the institutions of the traditional society. Growing agricultural efficiency released manpower for non-agricultural employment. Industry, in the age of steam, built urban centers. In the anonymity of the burgeoning city, neighbors knew little of each other's antecedents. Men came to find their place less by their ancestry than by their accomplishments. Education became widespread and secularization proceeded apace. In the new technological age new occupations provided channels for vertical mobility, and new communities compelled spatial mobility. All of this served to weaken the extended family system and to give new importance to the nuclear family.

Parents eventually began to think less in terms of maintaining correct traditional behavior, and more in terms of providing opportunities in the modern world for their children's health, education and advancement. Under these circumstances married couples began to limit their childbearing by practising contraception. At first, this limitation was achieved mainly by folk methods that had been widely known but little used throughout the world for thousands of years. By the last quarter of the 19th Century declining birth rates were nearly universal in Western Europe. Later, as the demands for effective and acceptable contraception became more widespread, new methods were developed, and birth rates began to follow death rates in a rapid decline.

The epoch of transitional growth resulted because modern science and technology first permitted a decline of mortality, then gradually so transformed society as to motivate parents to limit their childbearing. Some delay in the reduction of the birth rate is probably almost inevitable in the nature of man and his institutions. The extent of the delay, the length and magnitude of the transitional growth are, however, by no means invarient. They depend on many things, including the efficiency of the contraceptive methods available, the nature of the educational stimulus, and the rapidity and nature of the social economic change.

Another and equally valid way of describing the transition in human fertility is in terms of the change from a situation in which the controlling factors, both positive and negative, are mainly institutional, to a situation in which the controlling factors lie mainly in the area of rational choice by the couples involved. In the traditional societies, the institutions govern the fact and timing of marriage and provide the rules governing sexual life within marriage. In sharpest contrast, under the modern system, the principal determinants of fertility are in the area of rational choice on the part of the couples.

The change from one system to another seems almost inevitably to involve growth, as we have seen. But this should not be taken to mean that growth may not also be quite rapid under the modern system. The postwar experience of the United States, Canada Australia and New Zealand has amply demonstrated that under conditions of economic prosperity, populations in which there is widespread and effective birth control may grow at higher rates than those experienced by Europe in its 19th Century transitional period. It should be clear that this generalized description of the demographic transition has related more to the nature of the processes than to the magnitude and tempo of the change.

Thus far, moreover, our description of the transitional process has run in terms of the Western world. Will these processes be at all similar in different cultures, races, climates and forms of organization? One cannot know with certainty, but the evidence strongly suggests that the major principles of the transition will be the same. Note that in our account of events in the Western world, we have invoked no unique characteristics such as culture, race, climate, or organization. The principles invoked were those of adaptive response to the requirements of an age of modern science and technology and the nature of society in which normative elements slow the processes of social change.

The principles are universal. Moreover, the transition has not been limited to the West. It came on very rapidly in the USSR, especially after the war. Japan after a late start now has moved as far as any nation on the transitional course. In Ceylon, too, there is some suggestion that the reduction of fertility has begun. It has certainly begun with the Chinese of Taiwan, Hong Kong, Singapore, and in the upper economic and educational levels of the Indian sub-continent.

Both theory and experience, therefore, suggest that the present period of population growth is mainly a by-product of a universal demographic transition arising from the nature of society and the modernizing process. Some countries have gone far along the path, while others, notably the newly-developing countries, have scarcely begun.

I have been stressing the universal nature of the principles governing human population change. It is time now to turn to an account of the diversity of actual experience. Perhaps two facts and their explanations will serve. Western Europe in its early period of modernization never experienced population growth of more than one percent per year. In today's newly-developing countries, the rates of growth are ranging from two to nearly four percent per year. If the principles are the same, why such differences in the orders of magnitude?

The reasons are not hard to find. In the first place, birth rates are higher in most of the newly-developing countries of Latin America, Asia and Africa, than they were in Western Europe before the modern era began. For the most part, the high birth rates in today's developing countries appear to come from more nearly universal and younger marriage than was the case in Western Europe. This is less true of Ceylon than of most developing countries. In the 19th Century, few countries of Western Europe had birth rates higher than 35, that is 35 births a year per thousand total population. Today, birth rates from 40 to 55 characterize most of the newly-developing countries. Their rate of accretion through birth is from 5 to 20 per thousand higher today than it was in 19th Century Europe.

But the contrast is even more striking when the rate of depletion by death is considered. Few of us realize the effectiveness with which disease is now controlled. Today, the death rates in the newly-developing countries are declining from three to five times the speed with which they moved through the similar levels in Europe's 19th Century. The major change has come in the last forty years with the development of vaccines, sulpha drugs, antibiotics, and the powerful new insecticides which control the vectors of infection.

In an epoch that has been characterized by man's most awful experiences in self-destruction, we too often forget our less dramatic achievements. War and catastrophe are news, but quiet progress gains little attention. Few people realize, for example, that is

spite of its two most destructive wars, Europe has more survivors today than it would have had without either war, but just with the peacetime death rates of 1910. Of course, the gains in your country, and indeed in most of Asia and Latin America, have been even more dramatic.

Incredible as it may seem, the lowest death rates of the world are no longer to be found in the most highly developed countries. Today the lowest death rates are in such areas, as Taiwan, Singapore, Hong Kong and Puerto Rico. Ceylon today has a lower death rate than France.

The reason is partly obvious. Many diseases are now prevented by controlling the agents of infection. Such control, although difficult in absolute terms, is relatively simple, because it is achieved mainly by rather inexpensive manipulations of the environment involving minimal alterations in individual behavior. Other diseases are controlled by treatment of individuals seeking help because they are ill. Thanks to the modern drugs, this treatment can often be rather effectively administered by personnel who are not very highly trained. Indeed, thanks to the modern control of disease, there is probably today less connection between economic conditions and mortality than there was in the earlier period when sound nutrition was the principal preventative of disease. Today, we are highly efficient in keeping people alive in poverty.

The other reason that death rates are low in countries such as yours is that birth rates are high. This seeming paradox is both important and easy to understand. The crude death rate is the number of deaths in the year per thousand total population. It is affected both by the risks of death at each age and by the age composition of the population. Even under the best of conditions, old people tend to have high risks of death, whereas young people survive under amazingly adverse circumstances. But the age composition of the population, in which migration in unimportant, is in turn almost entirely determined by the past trend of the birth rate. High birth rates give young populations and low birth rates give relatively old populations. The newly-developing areas with fairly well established health protection have the lowest crude death rates in the world, because their risks of death are relatively low and, thanks to their very high birth rates, the population is very young.

The high birth rates of today's newly-developing countries, therefore, contribute to speeding population growth in two ways; they swell the entering stream of life, and, by creating young populations, they diminish the withdrawal through death.

The combination of very high birth rates and very low death rates has meant that relatively poor countries in the early stages of modernization are experiencing rates of growth ranging from 2 to 4 or more times those experienced in Europe during its comparable stage of economic and technological development. Already often densely settled, with little opportunity for migration, these countries face rates of growth that if continued will double their numbers in from 18 to 35 years.

When one says that the age composition of a population closed to migration is determined by its past birth rate, this amounts to saying that death rates are relatively neutral in their effect on age composition. This seeming paradox is important, and readily explained. At first glance one tends to think that a reduction of the risk of death would yield an older population by permitting longer survival. For the individual, of course, it does. The thing we neglect to think about in this connection is that this effect of lengthening individual life is about cancelled by the more rapid growth coming from a larger number of births and reduced losses in each age. The decline in mortality thus far has, in fact, resulted in rather constant proportional increases in the probability of surviving between ages five and fifty. The effect therefore has been to multiply each age-class by a constant, resulting in no proportional change. Improvement in mortality has been a little more rapid in the ages below five, so that the reduction of mortality has tended, if anything, to make the population a little younger instead of a little older.

One consequence of this fact is that improvement of health will not reduce the proportion of the population to be found in the ages of youth dependency. Only a reduction in the birth rate will lighten this load. Indeed, because the newly-developing countries have very high birth rates they have more than 40 percent of the total population at ages under 15, and correspondingly smaller porportions of the total in the ages from which the effective labor force is drawn. In the developed countries with lower birth rates the burden of youth dependency is much lighter and the labor force constitutes from 10 to 15 percent more of the total population than in the newly-developing countries with high birth rates.

Let us turn briefly to the economic implications of the situation. Too often these are discussed in terms of the optimum population, that is, by asking what, in view of the advantages of specialization on the one hand, and the costs of diminishing returns on the other, would be the ideal size of population to maximize per capita income. Such discussions seem to me to have little meaning. They impose a static analysis on a dynamic problem.

The principles, it seems to me, are crystal clear. In the first place, it is evident that the traditional self-sufficient agrarian society cannot provide health, education, and material prosperity for the

present population of either the world as a whole, or that of its more densely settled parts. If the universally sought objectives of health, education, and escape from poverty are to be attained, reliance on inanimate energy and the sophisticated use of raw materials is essential. Only by economic modernization can the goals of all the world's major peoples be obtained.

Modernization requires heavy investment in capital equipment, of course, but also heavy investment in health, education, and social facilities. Everyone is fully aware of the need for material investment in factories and transportation, but too little attention has been given to the equally pressing need for investment in both health and education. We forget that, in the long run, skill is the most important resource of all.

The emerging nations must modernize as quickly as possible, and to do so they must invest in training and equipment. No one is quite sure what a given new investment in equipment will mean in terms of added income flow. A good guess in economies that have redundant labor and little capital is that the ratio is three to one. If so, an investment of, say, nine percent of the nation's annual income in productive equipment would be expected to add about three percent to the annual stream of income. It is here that the demographic factor comes into play. If the population were stationary, an annual investment of nine percent would result in a three percent increase in income and a doubled per capita income in twenty-three years. But if the population were growing at three percent, the investment would bring no net gain. If growth exceeded three percent, living conditions would deteriorate. The very poorest nations with the greatest need for investment face the requirement, unless fertility falls, of investing from six to twelve percent of their national incomes just to avoid a deterioration of living conditions. This means that they are finding it extremely difficult to save and invest sufficiently to lift their per capita income. The difficulty is that the very processes of modernization, which require heavy investment, also produce rapid population growth. Net gains can be made only after the costs of that population growth have been paid.

Obviously saving for investment of ten or even twenty percent of the national income is by no means impossible in prosperous economies. It is extremely difficult for populations that already live close to the margins of subsistence, such as India, where eighty percent of the total family income goes for food alone and the people are still malnourished. It is clear that economic modernization can proceed much more rapidly if the rate of population growth is slow than it can if it is fast. Since good health and low death rates are also essential, the slower growth can only come from a reduction in the birth rate.

A sharp reduction in the birth rate has another advantage. It quickly produces a population in which larger proportions of the total are in the productive years of life, which, given adequate capital, makes possible more rapid growth of per capita income. It was considerations of this sort that led Coale and Hoover to conclude that if India were to reduce its human fertility by about fifty percent it would yield in a generation a per capita income about forty percent above the level that would exist if fertility were not reduced at all (Coale and Hoover, 1958). Something of the same sort is true for all of the emerging countries.

Thus far, I have made essentially four points:

1. Rapid development of the economies in the emerging nations is essential for the health, education and prosperity of the existing population. It is still more critically important for the maintenance of the much larger populations that the future will almost certainly bring.
2. It is in the nature of society that this process of economic modernization would reduce the death rate before it reduces the birth rate, and under modern conditions the resulting population growth can be very rapid and sustained for decades unless positive measures are taken.
3. Such rapid growth greatly complicates the process of modernization because the investment required to meet the needs of added numbers reduces the investment available to bring per capita gains.
4. A rapid reduction of the birth rate in the early stages of modernization would speed the process of development in two ways (*a*) by providing more capital per person and (*b*) by building a population in which a larger proportion of the total is in the ages from which the labour force is drawn.

Although such a reduction in fertility could greatly help in the course of the next two decades, it is not clear that in the long run it would mean a smaller population. My guess is that the most densely settled of the developing countries will ultimately have larger populations if they reduce their birth rates quickly than if they do not.

The argument to this effect is straight forward. The carrying capacity of the agrarian economy is strictly limited. Carrying capacity can be advanced to entirely new dimensions only by reliance on a highly educated and skilled population well organized in a capital-rich system. The transformation to this kind of an economy will be greatly speeded if the rate of growth during the

years immediately ahead can be cut by a reduction of the birth rate. Once the modernization has been accomplished, population growth will not have stopped. Populations recruited on the basis of low death rates and planned families, still grow. For example, in the United States our growth if continued will double the population in forty years. Personally, I hope our population does not double, but it would be possible to sustain such growth without overriding economic difficulty. Indeed, it is very difficult to put limits to the amount of population nations can carry with fully modernized economies. These limits are certainly much higher than for simple extractive economies. A reduction of birth rates now would facilitate the economic transformation, making possible the support of larger populations than could otherwise be maintained. The alternative may well be a difficult one. Population growth continuing unchecked for considerable periods at present high levels may well block economic modernization and bring an increase in mortality that carries us back to the age-old equilibrium. Unending growth is of course impossible. We face now the choice of means by which it will be curtailed, whether through the reduction of fertility or the increase of mortality.

In principle, the foregoing analysis would be accepted by demographers drawn from most parts of the world's ideological spectrum. But note how far it departs from the stereotypes of the earlier extremes.

Our analysis has not run in terms of the optimum population and its derivative "over-population". It has run instead in terms of the dynamics of change. It has shown that the extremists who suggest that mortality should not be cut until birth rates are reduced are ignorant of the processes at work as well as immoral. But equally it has shown that those who counsel reliance on the productive powers of modern technology to avoid the necessity of curtailing population growth are giving very dangerous advice. Our analysis has presented the need for lower birth rates, not as a substitute for modernization, but as a means of hastening the process of modernization. The rapid extension of family planning will not carry the threat of race suicide. Indeed, by hastening development, it will in the long run make possible larger rather than smaller populations. In the immediate future, a fifty percent reduction of birth rates would still leave most developing countries with higher rates of population growth than those of Europe's 19th Century. The practical possibilities are only those of keeping rates of population increase from being very large.

In short, as we have learned more about the causes of population change, the issues have changed from the hitherto irreconcilable disputes to new, more dynamic views in which divergencies of sophisticated opinion lie mainly in matters of emphasis.

One example of a difference in emphasis rather than of basic principle may be given. To me it seems a dangerous one. It is occasionally argued that, since the declines in fertility come mainly as by-products of social-economic modernization, the way to reduce the birth rate is to set about the process of development, and not bother with the problem of birth control. This conclusion is possible only if one neglects magnitudes and time scales.

Social-economic development, education and prosperity, will reduce the birth rate eventually by stimulating the practice of contraception. The same developments will also reduce the death rate, but no responsible person for that reason suggests that we should simply wait for this to happen; that we should not make special efforts in the field of health. Similarly, unless all our generalizations about education as a means of spreading innovative behavior are wrong, the decline of the birth rate can be greatly speeded in a climate of social economic development by public education and a readily available supply of safe, efficient, cheap, and acceptable contraceptive supplies for married couples. Nor is the cost of such efforts and supplies sufficient to appreciably reduce the resources available for other aspects of development. To me it seems clear that the risks of failure in our efforts at modernization will be greatly cut if the birth rate can follow the death rate rather closely in its decline. And it is most unlikely that this situation will come about without strong efforts on the part of the governments involved.

Let us turn now to review briefly the positions being taken in the world in relation to population growth in the developing countries. They have changed remarkably in the past ten years.

Present attitudes toward family planning in the developing countries are perhaps best reflected in the vote of the General Assembly of the United Nations in 1962. A proposal generally interpreted as instructing the United Nations to give technical assistance in family planning if requested was defeated, partly for technical reasons. The United States abstained from voting, on the ground that the necessary authority already existed. All of the Communist block abstained except Yugoslavia, which voted in favor. In the new countries of Africa, south of the Sahara, the situation was very mixed, as it was in Roman Catholic Latin America. But Asia and the Arab world voted almost unanimously in favor of such assistance. The solid core of opposition to it came from the Catholic countries of Western Europe, all of which themselves have very low birth rates. In short, except for holding back by some Catholic countries, it is precisely the nations that face the difficulties of economic growth that are expressing major interest in attempting to reach a solution.

India was the first major government to adopt a national policy of limiting growth by reducing the birth rate, but even its actual program is less than ten years old. A great deal of work has gone into organization, training and experimentation. Now the program is expanding rapidly. The budget for the fiscal year 1962-1963, 253 lakhs of rupees, was larger than the entire family planning budget for India's Second Five Year Plan. It would be too much to claim that India is now on the point of reducing the national birth rate, but there are signs in the large cities that the birth rate is falling. It may be that the family planning program is now getting into a position to achieve its goal of a birth rate of 25 by 1973.

We know very little about mainland *China's* program, but if the reports of interviews with high officials are correct, the government is making a major effort to lift the age at marriage and to spread contraception and sterilization throughout the population. The two largest nations of the world are therefore attempting to reduce their birth rates. The third and fourth largest, which are the *USSR* and the *United States*, already have birth rates below the Indian goal. Neither has more than a permissive governmental policy, but in the U.S. family planning practices are well-nigh universal. In the USSR, legal abortion plays a substantial role as, indeed, it does in all the Communist countries of Europe except East Germany.

The world's fifth and sixth largest countries, *Indonesia* and *Pakistan*, have divergent policies. The former has no national program, but it has the beginning of experimental work under the auspices of its National Science Council. Pakistan, by contrast, had a strong national program since 1960. Its annual budget is 60 lakhs of rupees. Progress is slow, but gradually the institutions, personnel, and materials are being built up. The next Plan calls for an annual budget of 3 crores. One would expect it to take a few more years before the national birth rate is affected.

Others here are better qualified than I am to discuss the situation in *Ceylon*. As I understand it, the government gives a little help to the private birth control movement and has sponsored a successful pilot project by the Swedish Technical Assistance group. Clearly, major sections of the population are ready to practice birth control whenever suitable information and materials are made available to them. It is unlikely, however, that there will be substantial reductions in Ceylon's birth rate for a considerable period unless there is a much stronger program than now exists.

It is in the smaller countries of Asia, however, that efforts to spread the practice of birth control have been conspicuously successful. The Republic of Korea has the largest per capita budget for this work of any Asian country, and probably of the world.

It is moving ahead rapidly and if present plans are carried out, Korea should have a substantial reduction of the birth rate within five years. Birth rates of Chinese populations are already falling in Hong Kong, Singapore and in Taiwan. In each case, the governmental agencies or private agencies with government support are actively engaged in fostering family planning. Birth rates in these populations may well be cut by one-third, if not by one-half, in the next five years.

Thailand recently sponsored an official seminar to discuss the question of a national population policy. No final decisions have yet been made, but the government has asked its Ministry of Health to set up some pilot projects in family planning.

On the other side of Asia, *Turkey* has declared that a reduction of the rate of population growth is one of the major objectives of its first Plan. Turkey is now actively building a program.

Nor are such efforts confined to Asia. In Africa, *Egypt* has stated its intention to reduce the rate of growth, and is formulating an expanded birth control program. *Tunisia* has a firm national policy to foster the regulation of fertility and the beginnings of a national program. In a number of Latin American countries, governmental hospitals, pre-occupied with the medical problems of criminal abortion, are beginning to develop contraceptive services. Indeed, Catholic institutions are undertaking serious studies of the problem and are setting up a variety of experimental projects.

In short, although efforts to spread the voluntary regulation of fertility are far from universal or highly effective in the newly-developing countries, recognition of the problem and initial attempts to cope with it are widespread in Asia, Africa and Latin America. Well over half of the population of the newly-developing world lives in countries having strong policies and active programs designed to stimulate the practice of birth control.

What then of the future? Does it seem likely that efforts to spread the voluntary regulation of fertility can be successful with sufficient speed to help foster economic development and improved living conditions? My answer is clearly "yes" for two reasons.

In the first place, there is already both an understanding of the problem and the will to solve it on the part of many governments.

In the second place, we now appear to have available for the first time truly effective methods of contraception that meet the needs of the situation. This is not the place, and I am not the person, to enter into a medical discussion, but a few comments seem appropriate.

All of you have heard, I am sure, that certain steroid pills, if taken each day for twenty days of the female cycle, unfailingly prevent pregnancy, and appear to do so without serious dangers. Well over two million women in the United States are said to be using this method. Unfortunately, the drugs are still too expensive for general use in the developing countries. The costs, however, are becoming smaller.

Even more encouraging is the accumulating evidence that the new plastic intra-uterine devices are virtually ideal contraceptives for at least eighty percent of the women who have already born a child. They are ideal in the sense that once inserted they require in normal cases no attention until the patient wants another child, acting almost as a reversible sterilization. Furthermore, the devices are easily inserted, they are safe, and they afford a higher degree of protection than the most effective conventional methods. Finally, they are inexpensive, costing only a few annas. They remain experimental. However, the Medical Advisory Committee of the Population Council has come to the conclusion that there is no medical reason why the method should not be generally used in field trials that provide for adequate medical support. This conclusion was reached after reviewing the evidence from more than 55 research projects carried out in major medical centers of the U.S., Europe and Asia. The Ministry of Health of Pakistan, after observation of its own pilot projects, has decided to concentrate on the use of the method. The Health Services of Taiwan and Korea have already inserted thousands of the devices. In both countries, as the news spreads from patient to friend, it is scarcely possible to keep up with the demand. Plans looking to the large-scale manufacture of the devices are going forward in Pakistan, Hong Kong, Korea and Taiwan.

For the first time in history we appear to have methods that are appropriate to meet the needs of the weakly-motivated, illiterate and impoverished elements of the population. Given suitable organization for the dissemination of information, service, and supplies, there is no reason why all peoples everywhere cannot restrict their childbearing as they see fit.

Thanks to the new awareness of the problems by peoples and governments, and the new and extremely powerful methods of regulating fertility, I think we are entitled to new optimism. The problems will not be easily or automatically solved. Some countries may indeed proceed so inefficiently that their growth will ultimately be checked by rising mortality. But given solid planning and hard work, there can be success. Uncontrolled population growth need not defeat our best efforts to attain health, education and freedom from want.

References

Coale, A.J. and Hoover, E.M. (1958). *Population Growth and Economic Development in Low-Income countries*. Princeton.

de Castr J. (1952). *The Geography of Hunger*. London: Gollancz.

The Growth of World Population (1963). Washington: National Academy of Sciences.

POPULAR LECTURE

25th September, 1964

COMPARATIVE DIGESTION WITH PARTICULAR REFERENCE TO RUMINANTS

By

D. P. CUTHBERTSON, F.R.S.E.*

A casual comparison of the physical and chemical composition of the food supply of man and his domesticated animals with those of the living organism that is synthesized and maintained from them is amply sufficient to convince that profound stages take place in this transmutation. As agriculturalists we may be primarily interested in the meat, wool, eggs and milk produced, but, as scientists, we appreciate that we are wonderfully made. The digestive agents that bring about the first steps in this conversion must be as diverse in nature as the variety of substrates to be hydrolysed.

Digestion and absorption in non-ruminants

Starting with ptyalin in the mouth in the case of man, food is subjected later to the sequential attack of a series of digestive enzymes, the digestive carbohydrases, proteases and lipases. The primary function of the stomach is to act as a storehouse so that food need only be ingested at intervals—in the case of man, three or four times a day. By electing to follow such a time-table the inconvenience of eating many small meals is avoided. The removal of this storehouse by gastrectomy is always a considerable handicap and is inevitably followed by loss of weight. While movement is sluggish at the cardiac end of the body of the stomach and is chiefly concerned with this reservoir function, the mixing and churning movements are carried out by the pyloric antrum.

Proteases

Although the impressive array of enzymes which can attack bonds in the protein molecule suggests that few bonds can remain intact after traversing the alimentary trace, some stretches of peptide chain adjacent to sugar residues in mucoproteins have been found to be totally resistant to enzyme attack *in vitro* and consequently may not be well developed *in vivo*, but with normal mixed diets such proteins are not likely to contribute more than a minor proportion of the total dietary protein and in consequence resistant peptides are unlikely to be nutritionally significant.

*Director, The Rowett Research Institute, Bucksburn, Aberdeen.

In the reduction of proteins to their constituent free amino acids the enzymes are the pepsin of the gastric juice, trypsin and chymotrypsins of the pancreatic juice, and the various peptidases in the pancreatic juice, in the intestinal mucosa, and secreted into the intestinal juice.

Enzyme attack in the alimentary tract may also be inefficient because of inhibitors present in the diet as, for example, the soya-bean trypsin inhibitors. There are at least four. Other legumes also contain inhibitors and not all processes of heat treatment in the case of soya-bean inactivate these inhibitors. Indeed the factors determining digestibility in many vegetables, legumes, cereals and even partially purified proteins such as leaf proteins, need examination. Exploration of this field may well be worthwhile promoting, since in areas where the intake is marginal, the losses due to incomplete digestion may be critical. It has been shown that a large proportion of the population of eastern countries has an abnormal intestinal mucosa by western standards.

Carbohydrases

Carbohydrate digestion starts in the mouth with the action of the salivary amylase pytalin on starch, dextrins and glycogen to yield maltose, but pytalin is not found in the saliva of the dog, sheep, goat, ox or horse. Thereafter the carbohydrates are submitted to pancreatic and intestinal carbohydrates which reduce carbohydrates to monosaccharides. The only strange happening is the presence of lactase—in the absence of lactose—in the crop of chickens.

Lipases

The situation in ruminants which will be discussed later is in marked contrast to that obtaining in simple-stomached animals (including, of course, non-ruminant herbivores), in which the digestion of food lipids takes place almost entirely in the small intestine not long after the food has been ingested. This process involves lipolytic enzymes (especially pancreatic lipase) and bile salts to promote emulsification and hydrolysis of the lipids. Though the form in which lipids are absorbed from the small intestine of the non-ruminant animal has long been a contentious subject, it now seems that emulsification and lipolysis are but steps preliminary to the production of soluble micelles consisting largely of bile salts, free fatty acids and 2-monoglycerides (i.e. the remaining fatty acid of the triglyceride occupies the middle position of the glycerol molecule) from which the fatty acids and monoglycerides are taken up by the intestinal mucosal cells for conversion into the triglycerides of chylomicra. There is no reason to suppose that the absorption of long chain fatty acids from the small intestine of the ruminant follows a different pattern to that in other mammals.

Endogenous secretion of protein into the alimentary tract

The amount of protein secreted in the form of enzymes, mucin and as leakage of plasma proteins into the alimentary tract may amount in total to an appreciable amount of protein per day. This raises the question of whether the amino acid pattern absorbed from the gut may not sometimes be an improvement on that presented by the dietary proteins, because of their contribution of the amino acids from endogenous sources.

The nature of the absorbed products

The evidence, as we have seen, is now heavily in favour of the products of digestion being monosaccharides, free amino acids and, as far as lipides are concerned, free fatty acids and 2-monoglycerides and that they are absorbed as such. In the case of the products of lipolysis reconversion into triglycerides occurs in the mucosa.

The possibility of minor quantities of peptides and even whole protein molecules being absorbed into the blood seems equally certain, but from a nutritional point of view these are not quantitatively important.

The mechanism of absorption is not important nutritionally but it has been found in lower animals (sheep, rat and mouse), that the increased intake accompanying lactation is accompanied by hypertrophy and hyperplasia of the intestinal mucosa and this disappeared on weaning.

Microbial activity in the digestive tract of non-ruminants

The most important contribution of the gastrointestinal flora of non-ruminants is the reduction of cellulosic materials to water-soluble forms. No digestive enzymes are secreted by the tissues of the alimentary tract of higher forms of animal life. The advantages of cellulosic degradation to the host animal are, first, that it renders available a considerable part of the food energy locked up in these complex carbohydrates, and, second, that it liberates from plant cells enclosed by cellulosic membranes the nutrients that are digestible by animal enzymes. Where ingesta stagnate or are side-tracked temporarily from the main flow of digesta, microbial fermentation may set in. The caecum and colon are structures of this description.

The predominant micro-organisms in the large intestinal flora of the pig are streptococci and lactobacilli with smaller numbers of coliform bacteria, micrococci and moulds.

The comparative ability of different species of animals to digest cellulosic material can be assessed by measuring the production of volatile fatty acids or the digestibility of cellulose. Both man and the pig can digest non-lignified cellulose with an efficiency not far short of the ruminant. This action is brought about not by endogenous cellulases but through micro-organisms in the caecum and colon. The most striking difference between the ruminant and the pig is that the latter makes much poorer use of cellulose in the lignified form.

Uncooked tuber starches generally escape digestion in non-ruminants until they reach the caecum where they are attacked by bacteria. In the course of their propagation in the lower levels of the intestine, bacteria also synthesize microbial protein and other components present in normal cells, such as the B-vitamins, deriving energy for these syntheses from the breakdown of carbohydrates in the ingesta of their host. The B-vitamins and vitamin K are available to the host animal.

In the rabbit, both the stomach and caecum-colon are developed to accommodate bacterial proliferation but the activity in the latter is much greater. It would appear, however, that such bacterial cells are not subject to any considerable digestion. In man there appears to be an appreciable biosynthesis of thiamine, riboflavin snd nicotinic acid by micro organisms.

In the bird, the caeca are the sites of the greatest concentration of intestinal organisms and of B-vitamins. However, they are not essential for B-vitamin synthesis.

In the horse, fermentative digestion by bacteria is carried out after the food has been subjected to the action of the alimentary enzymes; those not hydrolysed like cellulose are subjected to considerable microbial attack in the large intestine. Indeed the cellulolytic enzymes found there appear to be more efficient than those in the rumen.

The adult horse has a stomach of relatively small capacity but in the sucking foal it is relatively large. A certain amount of bacterial activity takes place in the stomach.

Large amounts of B-vitamins are synthesized in the caecum and colon of the horse. Volatile fatty acids are produced and absorbed from there.

Digestion and absorption in the bird

The distinctive features of the digestive tract of the bird are: the mouth, which lacks teeth and has a limited supply of salivary glands; the crop, which performs the useful function of a storage organ and 'meter' for admitting food into the digestive apparatus; the small true stomach (proventriculus); the powerful muscular gizzard which performs the task of mastication; the large well-

defined duodenum and pancreas; the relatively short intestinal tract; the large well-defined caeca, and the short rectum ending in the common excretory organ for both faecal and urinary wastes.

Salivary secretion is probably of relatively little significance and although the crop undoubtedly plays a function in the normal pattern of food intake, and is a highly adaptive organ to conditions in which access to food is limited in time, the normal functioning of cropless birds shows that this organ is not necessary for survival if the removal is effected early in life. Excision in the mature bird has been reported to result in nutritional failure and death. Provided the bird is fed on a diet not requiring mastication the gizzard can be dispensed with barring a slight reduction in digestibility. Caecotomy can also be done without disturbing normal function.

The organ of central importance in the digestive tract of the bird is the pancreas. Its removal, or the ligation of the pancreatic ducts, leads to prompt impairment of digestion, particularly for proteins and fats. This work has also shown a marked loss of tonus of the intestine and atrophy of the gonads in the depancreatized bird; both effects were prevented by administering fresh pancreas as 10-20 per cent of the diet. Supplementation of the diet with fresh pancreas also restored digestion and absorption of protein, fat and carbohydrate to normal levels, even though the concentration of digestive enzymes in the gut contents was far below normal. There seems to be no work in which the true stomach of the bird has been removed surgically to determine its effect on digestion.

The overall efficiency of digestion and absorption is in general accord with those which have been obtained with other species. The bird is completely unable to use cellulose, though the possibility of partial degradation is not ruled out. In spite of a relatively short digestive tract the bird has a highly efficient apparatus.

The fat surplus in some of the developed countries has encouraged the utilization of inedible fats in animal feeding and has stimulated research. Beef tallow is poorly utilized, fluid vegetable fats are highly utilized. The higher the proportion of saturated fatty acids in the mixture, the poorer the absorption.

Digestion and absorption in the ruminant

Wild ruminants live on a variety of vegetation but domestic ruminants are faced with a wider range of dietary conditions, which, on the one hand, extend to free ranging cattle and sheep living close to the conditions affecting their wild relations; and, on the other, to intensive livestock husbandry on lines akin to pig and poultry production for the production of milk, beef or fat lambs.

The function of the reticulum-rumen

The outstanding difference between the ruminant and other herbivores is that the former pre-ferment their food through microbial action in the first two compartments of the complex stomach—for convenience termed the reticulo-rumen sac—before digestion by enzymes secreted by the digestive glands begins in the abomasum or fourth stomach. Fermentation of cellulose through microbial activity occurs in this sac and to a limited extent also in the caecum and colon. Digestion is not as efficient in these parts as in the rumen, possibly because the rate of passage is too rapid to permit bacterial action.

The reticulum and rumen thus retain food under suitable conditions for sufficiently long periods to allow extensive digestion through the metabolic activity of the bacteria and protozoa which inhabit these organs.

The reticulo-rumen sac forms a very special fermentation vat in which these micro-organisms are established under highly reducing conditions when the animal is quite young and the rumen relatively small. Removal of much of the cellulose and growth of micro-organisms both precede the normal alimentary digestion and absorption that we have discussed in non-ruminants. This enables the products of microbial synthesis and the residual food nutrients enclosed within the plant cell walls to become readily available for absorption.

In grass and especially in other root crops, other fermentable carbohydrates exceed cellulose in amount and, being more soluble and less structurally and mechanically resistant than cellulose, are more readily fermentable by a much wider variety of micro-organisms but not necessarily to the same ratio of desired products. The reticulo-rumen process is a functional field in which all activities are organised for utilization of microbial products and all factors therein make for the attainment of a steady state. Thus we find that the reticulo-rumen sac is normally never empty and this probably explains why the salivary, gastric and intestinal secretions are continuous. The activity of the salivary glands, the process of rumination involving re-mastication, the absorption of the products of microbial activity and the interchange of ions between the rumen contents and blood all contribute to the continuity of the microbial action during the whole life of the animal and produce conditions which one can hardly hope to reproduce *in vitro*. Nevertheless, by study of the micro-organisms—the bacteria and protozoa, the latter chiefly ciliates—at the site of their attack on food particles, and in isolation in culture, and by a study of the soluble products of fermentation and their absorption, much information about this microcosm has been obtained.

Of the total material digested by the animal, at least half or, with easily digested foodstuffs, a great deal more than half of the material digested, disappears in the reticulum and rumen.

The products of digestion in the rumen

The products of digestion are the soluble materials—products of microbial metabolism—mainly acetic, propionic and butyric acids with small quantities of higher acids; some ammonia is also formed and intermediate stages of the fermentation of the food substrates can be detected. Carbon dioxide and methane are also evolved. The bodies of the micro-organisms themselves are important products of this phase of digestion and in the course of their metabolism they also produce vitamins of the B-complex and, in addition, any lipids of the food are hydrolyzed and unsaturated long-chain acids are hydrogenated. The figures which follow shortly describe the various metabolic pathways in the rumen for protein, carbohydrate and lipid.

The quantities of each of the short-chain fatty acids produced and the quantity of bacterial and protozoan substance passing from the rumen to the remainder of the stomach and the intestines have been assessed. Microbial substance is important because it provides not only a source of protein to the animal which is nutritionally first class but also microbial polysaccharides and lipids. These microbial proteins are subsequently digested in the abomasum and intestines.

Movements of reticulum and rumen

These organs are designed so that solid food is retained until it is in a sufficiently small particulate state to leave the organ. The solid mass of food is continually flushed with the liquid phase of the contents which is largely derived from the salivary glands, and the material leaving the organ and entering the omasum or third compartment of the complex stomach system, has a more fluid consistency than that of the rumen. The half-life of rumen bacteria within the rumen is thought to be about 9 hours, while the half-life of roughage is about 48 hours.

When the ruminant eats highly digestible foodstuffs which are already in a fine state of division, the contents can pass through to the omasum without further disintegration. Under these circumstances the liquid and solid phases are likely to be indistinguishable. Animals fed such highly digestible material eat more than those fed roughage. Disappearance of food from the rumen is due to two causes: the first being the absorption of the soluble by-products of the microbial fermentation, and the second being due to the digesta passing to the omasum. The probability is that the quantity of food leaving the reticulum and rumen and passing to the omasum

in an unchanged or relatively unchanged state is greater with such readily digestible food than when coarse food is eaten. The quantity of food fermented is much greater than the quantity that passes unchanged from the organ. With highly digestible material fermentation can be so rapid that the animals tend to bloat, especially if no additional roughage is allowed other than that built into the cereal itself. It is for this reason that antibiotics are usually included in such rations.

Where animals are allowed roughage but also receive a meals supplement particles of the same size as those of meals leave the rumen more rapidly than roughage fed in a long state. The proportion of food digested in the rumen is of the order of 70 per cent of the total quantity digested in the alimentary tract, so that the quantity of meals arriving intact in the abomasum is not a large part of that eaten. The question of how much of the total digestion takes place in the reticulum and rumen has considerable nutritional implications as far as starches and proteins are concerned because we may expect these constituents of the food to be digested with less loss to the animal in the small intestine than in the rumen. On the other hand, cellulose and its associated constituents which make up the plant fibre, are digested only by microbial enzymes: consequently, the effective part of their digestion is confined for the most part to the reticulum and rumen and to a lesser extent to the large intestine, the next distal zone of microbial activity.

Rumen protozoa

The numbers of protozoa are sufficient to make their metabolism contribute to quite a substantial extent to the chemical changes in the rumen. Their numbers increase with the nutritional value of the food, and with rations with ample starch and protein support a population in the order of 1 to 2 millions per ml., provided the reaction in the rumen does not become too acid, for they do not tolerate such conditions. Hay diets are associated with populations of 200,000 to 500,000 per ml.

So far as can be judged by rather crude tests and digestibility studies, it does not seem as if they are essential to the animals nutrition. Nevertheless, the ease with which they are digested in the abomasum and intestines means that they can make a substantial contribution to the end of digestion in the rumen. These organisms ingest starch granules which are subsequently changed into a storage polysaccharide; many are known to ferment soluble sugars and also to store them as polysaccharide granules. They have also cellulolytic properties and ingest protein. In the course of their metabolism they produce short-chain fatty acids, carbon dioxide and ammonia.

Protozoa obtain their food by ingestion and their nitrogen requirements are presumably by the ingestion of protein fragments and bacteria. They cannot be sustained by ammonia nitrogen as can so many of the bacteria.

Rumen bacteria

Carbohydrate

A very large number of species and strains of bacteria have been isolated from the rumen and studied. Their culture presents difficulties to the orthodox bacteriologist for many of them are facultative anaerobes, but the more important species that ferment cellulose are strict anaerobes and rigid precautions to exclude oxygen from the medium and the gas phase over the media are necessary for their culture.

Most of the bacteria of the rumen that ferment cellulose and cellobiose are unable to ferment starch, but a considerable number of other species can ferment starch and many soluble sugars. Other species ferment xylan—a polysaccharide made up of pentose units of xylose—an important constituent of the hemicelluloses of plant fibres.

Bacteria are thus present which ferment all the carbohydrate of the food. Of the products formed, only carbon dioxide, acetic, propionic, butyric acids and sometimes lactic acid appear in any quantity in the rumen. Methane is probably formed by reduction of carbon dioxide.

The proportions of acetic, propionic and butyric acids present in the rumen contents vary with the diet. With pure cellulose and with starch and glucose *in vitro* fermentations yield approximately similar quantities of the first two but very little butyric acid. In the rumen, however, the proportions of acetic to propionic to butyric acids are of the order of 7 : 2 : 1 in sheep fed on a medium to poor hay; but when young, succulent, pasture grass is eaten, these proportions may change to about 5 : 3 : 2. Similarly, in animals that are fed on hay to which starch and protein-rich supplements are added, the proportions change in the same direction as the quantity of hay is reduced in the ration and the quantity of supplement increased.

The rate at which these acids are absorbed from the rumen is influenced by the acidity of the contents, so that the absorption is not necessarily proportional to the concentration present. Finally, certain bacteria are known to synthesize longer chain acids from shorter chain acids; for example, butyric acid may be synthesized from acetic acid and valeric from propionic acid.

The use of "artificial" rumen systems in which foodstuffs themselves are fermented by mixed cultures that include protozoa and which allow diffusion of the soluble products out of the system, show a fatty acid picture which is closer to that found in the rumen. Comparisons of artificial digests carried out in such systems in terms of the fibre digested, and with the results of digestibility trials with the same foodstuffs, have produced results which suggest that these *in vitro* methods may be used to give a reasonable accurate picture of the nutritional value of fodders.

Protein

The speed at which a protein is attacked in the rumen is influenced by its solubility. Thus, casein is rapidly broken down and only a negligible amount leaves the rumen, whereas with the insoluble protein such as zein about half appears in the abomasum intact. The rate of attack can be juded by the concentration of ammonia appearing in solution, providing the ration does not contain large quantities of starch which alter the picture. The addition of a readily available carbohydrate such as starch reduces markedly the concentrations of ammonia that occur.

The first stage of breakdown is presumably an hydrolysis and bacteria with proteolytic properties have been isolated from the rumen. The next stages are presumably peptides, and finally, amino acids. Amino acids can in fact be shown to accumulate *in vitro* when proteins are incubated in rumen liquor in the presence of toluene which inhibits deaminative activity of rumen bacteria. Deamination by mixed rumen bacteria is rapid. The products are ammonia, carbon dioxide, and a mixture of short-chain fatty acids which, besides acetic, propionic, and *n*-butyricacids, include branch chain isomers of butyric acid and straight and branched chain valeric acid and possibly of C_6 acids. These isomers of C_4, C_5 and C_6 acids stimulate fermentation of cellulose *in vitro* by rumen micro-organisms.

The nitrogen requirements of the rumen organisms come, for the most part, from the food, although urea and mucoprotein contained in the salivary secretion, and the urea that diffuses directly into the rumen from the blood, contribute to the nitrogen pool.

The probability is that much of the ammonia nitrogen is incorporated into bacterial protoplasm, but this does not preclude the utilization of amino acids as such by direct incorporation, for some of the rumen bacteria will not grow *in vitro* unless amino nitrogen is available. Many of the important strains are not so fastidious. Protein digestion in the rumen then is a process of disintegration followed by resynthesis of some of the nitrogen within bacterial cells, loss of ammonia, by direct absorption from the rumen and loss of energy during this complicated biological trans-

action mean that the nutritive value of high quality proteins may be reduced. On the other hand, low quality food proteins can give rise to bacterial and protozoan proteins of better value to the animal. When proteins are fed in a highly insoluble state, or in such quantity that a significant amount passes to the abomasum and small intestine intact, than their value will tend to revert to that which it it would have been if fed to pigs or poultry. The bacteria and protozoa are the principal products of protein metabolism in the rumen. This conversion of dietary to microbial protein is the reason why the so-called "biological values" of dietary proteins are less variable in ruminant nutrition than in the nutrition of pigs and poultry. While heat treatment generally damages protein for the non-ruminant the resulting slower microbial attack in the ruminant permits the elaboration of microbial protein which is useful to the animal.

Early development of the reticulo-rumen and the concentration of micro-organisms

The four chambers of the ruminant stomach are well defined at birth, the abomasum—that is the pepsin-HCl secreting compartment—being proportionately the largest. The reticulum and rumen are lined by an epithelium of oesophageal type, prolonged into low villous processes: in the case of the reticulum a network of ridges. The omasum also has a wall of oesophageal structure, folded into long and very muscular ridges, by which the mixture of digesta is further broken up. The abomasum, as has been mentioned, is the seat of gastric secretion and, apart from the large and curiously spiral colon, this and the rest of the alimentary tract are like that of the rat and man.

Normally, when the young ruminant sucks its mother the milk passes directly to the abomasum owing to reflex closure of the oesophageal groove. This avoids unnecessary bacterial degradation of the milk constituents. By 48 hours after birth the flow of gastric juice is well established. Up to this time the transfer of colostral proteins concerned with the transfer of immunity can take place. Salivation in the newborn animal is undeveloped.

By 6 weeks of age the rumen contents appear normal in so far as smell and consistency are concerned. The adult bacterial association is not built up at once and before rumination is established the rumen population is apparently quite different from that found in the adult. During the first three months of life the proportions of the volatile fatty acids differ from the general adult pattern in that more propionic and less acetic are present.

Protozoa do not appear in the reticulo-rumen until calves are about 1-2 months old and contact with normal animals harbouring these appears to be necessary as no one has yet proved the existence of free-living forms of these organisms. Provided

roughage is available in addition to cows' milk, ciliates will develop in animals under 2 weeks of age. However, if starch takes the place of the roughage the acid conditions (lactic acid) in the rumen cause the fauna to disappear.

The numbers of bacteria present are influenced by the diet and the factor that is important is the quantity of nitrogen in the food because the numbers increase with the nitrogen content of the ration. The addition of starch to a hay diet or food rich in starch with an insoluble kind of protein such as flaked maize, tends to depress the count.

The flow of food from the reticulum and rumen to the small intestine

In the sheep about 7 litres of fluid contents pass per 24 hr. to the omasum and because of the capacity of this organ to absorb water, only about 5·5 litres pass on to the abomasum. In cattle the flow to the omasum is of the order of 160 litres per 24 hr. and the absorption of water there presumably proportionately the same. Gastric juice secreted in the abomasum of the sheep increases the volume, and the flow to the duodenum is of the order of 9 litres per 24 hr.

The flow of food through the stomach is dependent on the amount of food eaten and presumably by the quantity of saliva secreted, and its rate of flow seems to be influenced by the time devoted to rumination although it does not automatically happen that flow is increased by rumination.

There is remarkably little variation in the 24 hours in the composition of the material entering the duodenum. The nitrogen passing to the duodenum per day when protein-rich supplements are fed is less than that consumed; but in hay-fed animals or hay with flaked maize, the nitrogen entering the duodenum is more than that consumed in the food—which is an indication that the endogenous secretion of nitrogen into the stomach is greater than any loss of nitrogen due to the absorption of ammonia that occurs, while the reverse situation seems to hold when protein-rich supplements are fed. Of the short-chain fatty acids, most have disappeared by the time the digesta are flowing to the duodenum: this is due to their absorption, mainly in the reticulo-rumen. The acid environment of the abomasum, the contents of which are usually at pH3, effectively kills protozoa and probably many of the bacteria and inhibits further activity by the remainder. There is no certain information on how much of the protein and starch-rich concentrates reaches the small intestine unchanged when these form a supplement to hay. Some starch, as such, undoubtedly passes on.

Nutritional value of the end products of fermentation

The three principal acids produced by fermentation do not have the same nutritional value. Their value has been assessed at the Hannah Diary Research Institute and in terms of their ability to spare energy losses in fasting sheep propionic acid has an efficiency of 86 per cent, butyric acid an efficiency of 76 per cent, but acetic acid has only an efficiency of 59 per cent. Glucose has an efficiency value of 94 per cent when given into the rumen. So these lower volatile fatty acids are less efficient sources of energy than glucose and this does not take into account losses that occur during their production. The efficiency of the three together was greater than the study of each acid suggested—about 85 per cent. For fattening, the efficiency for acetic, propionic and butyric acids in producing fat were 33, 56 and 62 per cent respectively but no synergistic effect was noted. Rations that favour the production of propionic and butyric acids will be better for fattening than those that favour the production of acetic acid. Starch—and protein-rich feeds produce higher concentrations of propionic and butyric acids in the rumen than roughages, which favour acetic acid.

This general principle, however, is not applicable to lactating cows; acetate is one of the components from which milk fat is synthesized in the udder, and when rations are fed which unduly depress the quantity of acetic acid produced in the rumen, then the percentage of fat in the milk is also depressed. Most systems of feeding produce a mixture of fatty acids in the rumen in which acetic forms about 60 per cent or more; but when large quantities of starch or soluble sugars and inadequate fibrous food are given, propionic acid production is encouraged and the proportion of acetic acid produced may be only 50 per cent or less of the total, thus adversely affecting the amount of fat in the milk.

Printed by Lakshman Perera at the Colombo Apothecaries' Co. Ltd.
125, Glennie Street, Colombo 2.