1

An introduction to ecological economics

The purpose of this short chapter is to introduce the subject matter and to explain the organisation of the book.

1.1 WHAT IS ECOLOGICAL ECONOMICS?

The Greek word 'oikos' is the origin of the 'eco' in both ecology and economics. Oikos means household. Ecology is the study of nature's housekeeping, and economics is the study of housekeeping in human societies. **Ecology** can be defined as the study of the relations of animals and plants to their organic and inorganic environments and **economics** as the study of how humans make their living, how they satisfy their needs and desires.

Ecological economics is the study of the relationships between human housekeeping and nature's housekeeping. Put another way, it is about the interactions between economic systems and ecological systems. Humans are a species of animal so that in a sense, on these definitions, the field of study for economics is a subset of that for ecology. However, humans are a special kind of animal, mainly distinguished by their capacity for social interaction between individuals, and their economic activity is now distinctly different from that of other animals. Rather than one being a subset of the other, economics and ecology are disciplines whose subject matters overlap, and, as shown in Figure 1.1, ecological economics is where they overlap. Figure 1.2 is a summary of the essentials of the interactions between economic and ecological systems. Whereas Figure 1.1 is about fields of study, Figure 1.2 concerns the systems of interest. In it the 'Economy' is the world's economies treated as a single system, and the 'Environment' is the whole natural environment, planet earth. The economy is located within the environment, and exchanges energy and matter with it. In making their living, humans extract various kinds of useful things – oil, iron ore, timber, etc., for example – from the environment. Humans also put back into the environment the various kinds of wastes that necessarily arise in the making of their living – sulphur dioxide and carbon dioxide from burning oil, for example. The environment for humans, planet Earth, itself has an environment, which is the rest of the universe. Our environment exchanges energy, but not matter, with its environment. Human economic activity has always involved the material and energy exchanges with the environment shown in Figure 1.2. It would be impossible for humans to satisfy their needs without interacting with nature. For most of human history, mainly because there were few humans, the

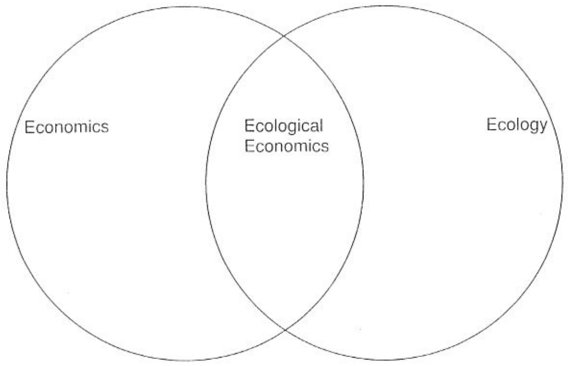


Figure 1.1 Locating ecological economics.

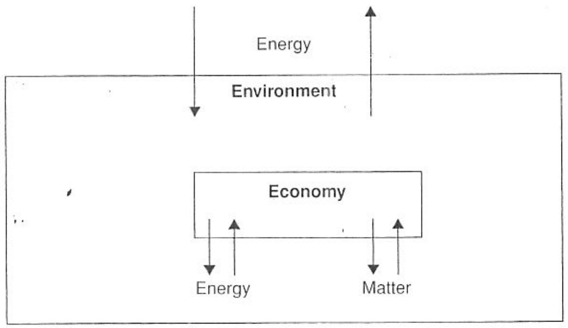


Figure 1.2 The economy in the environment.

level of interaction did not much affect the functioning of the environment, except locally. However, in the last three centuries the magnitude of the interactions has been increasing rapidly. The global scale of human economic activity is now such that the levels of its extractions from and insertions into the environment do affect the way that it works. Changes in the way that the environment works affect its ability to provide services to human economic activity. The economy and the environment are interdependent – what happens in the economy affects the environment which affects the economy. Another way that we shall sometimes put this is to say that the economy and the environment are a joint system.

One example of this is the role of carbon dioxide in climate change. Fossil fuels are extracted from the environment and burned in the economy, resulting in the release into the atmosphere of carbon dioxide. Carbon dioxide is one of several 'greenhouse gases'. The exchanges of energy between the environment and its environment shown in Figure 1.2 are affected by the amounts of these gases present in the atmosphere – higher concentrations of these gases mean that the environment, planet earth, gets warmer. As a result of the increasing use of fossil fuels in the last two hundred years, the amount of carbon dioxide in the atmosphere has increased. The expert consensus is that this has warmed the planet, and will warm it further. The amount of warming to be expected, by say 2100, is not known with any precision. But, the expert consensus is that it will be enough to have serious impacts on human economic activity and the satisfaction of needs and desires. Beyond 2100, the impacts may be catastrophic.

1.2 A BRIEF HISTORY OF THE ENVIRONMENT IN ECONOMICS

One way to introduce ecological economics is to look at the way that the natural environment has figured in economics through that subject's history.

Economics as a distinct field of study began in 1776 when Adam Smith (1723–1790) published *The Wealth of Nations*. This wide-ranging enquiry into the nature and causes of economic progress is now famous mainly for Smith's doctrine of the 'invisible hand'. This is the idea that, in the right circumstances, the social good will be best served by leaving individuals free to pursue their own selfish interests. Smith was one of a group now known as 'the classical economists', whose ideas dominated economics until the last quarter of the nineteenth century. **Classical economics** was widely known as 'the dismal science'. This was because it took the view, particularly associated with Thomas Malthus (1766–1834), that the long-run prospects for improving living standards were poor. This view was based on the assumed fixity of the supply of agricultural land, together with the propensity of the human population to grow in size. The environment, for the classical economists, set limits to the expansion of economic activity, so that the long-run tendency would be for the wages of workers to be driven down to subsistence level.

As a prediction, this has not fared well. In fact, to date, it has been wrong. For the economies of western Europe and their offshoots, the main features of experience since the beginning of the nineteenth century have been population growth and rising living standards. The standard explanation as to why Malthus got it wrong is that he overlooked technological progress. He, and the other classical economists, did assume an unchanging technology, when in fact it was changing very rapidly in the wake of the industrial revolution. However, it should also be noted that the economies of western Europe were not operating with a fixed supply of agricultural land during this period – increasingly food was being imported into those economies from 'new' land in the Americas and Australasia, to which those economies exported population.

This predictive failure was one factor leading to the demise of classical economics. Starting around 1870 mainstream economics began to evolve from classical economics towards what is now called 'neoclassical economics'. By 1950, the ideas of the classical economists were taught to students of economics only as part of the history of the subject. While the natural environment, in the particular form of the availability of land, had been a major concern of the classical economists, **neoclassical economics**, *circa* 1950, largely ignored the relationships between human housekeeping and nature's housekeeping. In the 1950s and 1960s, economists developed theories of economic growth in which the natural environment simply did not figure. These theories implied that given proper economic management, living standards could go on rising indefinitely. The pursuit of economic growth became a dominant objective of economic policy. One important reason for this was that economic growth seemed to offer the prospect of alleviating poverty in a relatively painless way. Neoclassical economics is not at all 'dismal'.

Starting in the early 1970s, neoclassical economics began to show renewed interest in the natural environment and it now includes the two important specialisations, or sub-disciplines, of **environmental economics** and **natural resource economics** (sometimes just resource economics). In terms of Figure 1.2, environmental economics (mainly) concerns itself with the economy's insertions into the environment, and with problems of environmental pollution. Natural resource economics concerns itself (mainly) with the economy's extractions from the environment, and with problems associated with the use of 'natural resources'. Many university economics programmes now offer higher-level optional courses in one or both of these specialisations. The compulsory courses in most economics programmes do not pay much attention to economy–environment interactions. It is possible to qualify as an economist and to know very little about environmental and resource economics. While neoclassical economists do not ignore the natural environment, they do not think that an understanding of the connections between the economy and the environment, as sketched in Figure 1.2, is an essential part of an economist's education.

Ecological economists do think that such an understanding is an essential part of an economist's education. Ecological economics is based on the idea that the proper study of 'how humans make their living' has to include the study of the relations of the human animal to its 'organic and inorganic environment'. Whereas neoclassical economics treats the study of economy–environment interdependence as an optional extra, for ecological economics it is foundational. It *starts* with the fact that economic activity takes place within the environment. Figure 1.2 – we shall look at a more detailed version of this in Chapter 4 – is the point of departure for ecological economics.

Ecological economics is a relatively new, transdisciplinary, field of study. In the last three decades of the twentieth century it became increasingly apparent to many scientists that human economic activity was having damaging impacts on the natural environment, and that this had economically harmful implications for future generations. The establishment, in 1989, of the International Society for Ecological Economics was motivated by the conviction, on the part of a number of scholars from several disciplines, that studying economy–environment interdependence and its implications requires a transdisciplinary approach, embracing parts of the traditional fields of study of the sciences of economics and ecology.

We need to explain our use of the term transdisciplinary here, and how it differs from terms such as interdisciplinary and multidisciplinary. For the prefixes here, the dictionary consulted gave the following meanings:

multi – many; more than two
inter – among; between; mutual, mutually
trans – across, over; beyond, on the far side of; through.

In connection with academic disciplines and research, the prefixes get used in slightly different ways by different people. However, the following captures what most people mean:

Multidisciplinary research tries to bring together knowledge from different disciplines – the problem is studied in several disciplines. Understanding of the problem is improved by the multidisciplinary approach, and the insights gained feed back into the development of the contributing disciplines.

Interdisciplinary research implies additionally that the disciplinary representatives are all involved in defining the problem, work to become familiar with the concepts and tools from the other disciplines, take on board results from the other disciplines, and that all are involved in presenting the results.

Transdisciplinary research is issue-oriented and interdisciplinary, and ideally involves stakeholders as well as scientists from relevant disciplines.

When we say that ecological economics is transdisciplinary, we do not simply mean that it is concerned with economic and ecological phenomena and draws on the disciplines of economics and ecology. It is and it does, but more is involved. The point of the 'trans' in relation to ecological economics is that there are phenomena and problems that cross, or are beyond, the disciplinary boundaries. Studying such phenomena and problems requires not just that an economist and an ecologist work on them together each using their own perspectives and tools. It requires a common perspective that 'transcends' those that are standard in the two disciplines. When working on economy–environment interdependence, the traditional perspective of economics needs to be modified to take on board the material basis for economic activity and the fact that humans are, whatever else as well, a species of animal. The traditional perspective of ecology needs to recognise the role of humanity as a species in the functioning of all ecosystems. With these shifts of perspective go the recognition of the usefulness of tools and methods of analysis historically seen as going with the other discipline.

Two more points. First, the proper study of economy–environment interdependence involves more than ecological economics as we have described it – many disciplines are highly relevant. However, we do consider that ecological economics is a useful starting point. Second, there are many phenomena and problems to do with economies and ecosystems that can be handled within the traditional disciplinary boundaries. If you only want to study the way the stock market works, you do not really need to take much from ecology: if you are concerned with only the food chains in a remote lake, you do not need to think much about economics. However, if you want to understand the global economy as a system for satisfying human needs and desires, or the operation of the global ecosystem in terms of the distribution and abundance of species, then you do need to cross boundaries.

Throughout the history of economics, as well as studying how humans actually do make their living, economists have offered advice on how they should make their living. One of the reasons that many are attracted to the study of economics is its prescriptive role. In the beginning, Adam Smith urged more reliance on markets and less state intervention in economic affairs than was actually the case at the time that he wrote. Since his time, the views of economists on many issues of public policy have always been an important input to political debate. Notoriously, economists do not, and have never, spoken with a single voice on any given policy issue. There are differences within the ranks of neoclassical economists, as well as between neoclassical and ecological economists. In order to prepare the ground for an introduction to the relationship between ecological and neoclassical economics, we need to look at the origins of differences on policy.

We will do that in section 1.5. First we need to explain the way we will use the terms 'economist(s)', 'neoclassical economist(s)' and 'ecological economist(s)' there, and throughout the rest of this text. There is much that the majority of neoclassical and the majority of ecological economists agree about. Where we are discussing something of this nature, we will refer to 'economists' or to 'economics' without any qualification. Where we are discussing something where there are significant differences we will refer to 'neoclassical economists/economics' or to 'ecological economists/economics' as appropriate.

1.3 SCIENCE AND ETHICS

In considering modes of study, a distinction is made between the 'positive' and the 'normative'. A positive study is purely descriptive, whereas a normative study includes prescriptive elements. A report on a positive study would consist entirely of statements about what is, or might be – it would be about facts and explanations. A report on a normative study would likely include such positive statements, but would also include normative statements about what ought to be – it would involve recommendations. A positive statement takes the form 'event *A* always follows action *B*'. A related normative statement would be 'event *A* is bad, and therefore action *B* should be avoided'. The recommendation here requires two elements – the factual link from *B* to *A*, and the classification of the outcome *A* as something bad. All recommendations, all policy advice, involve both positive and normative elements.

In principle, it is possible to establish the truth or falsity of positive statements in a way that would satisfy all interested parties. Suppose that Jack and Jill are the interested parties. Jack believes that *A* always follows *B*, but Jill does not. The disagreement can be resolved. Jack and Jill could, for example, observe many repetitions of action *B* and record the subsequent occurrence, or non-occurrence, of event *A*. If ever *A* did not occur, Jack would have to agree that the statement 'event *A* always follows action *B*' is incorrect. The situation is different with normative statements – they cannot be classified as true or false on a factual basis. If Jack and Jill disagree about whether *A* is a bad outcome, there is no experiment that can resolve that difference.

One definition of science is that it is the business of sorting positive statements into the categories of true and false. Some people would argue that any field of study that involves making recommendations is not a science. However, many people working in fields generally regarded as branches of science do make recommendations. There need not be a contradiction here. Many recommendations are really conditional advice. Thus, if it were established knowledge in some field that *A* does always follow *B*, a recommendation from a scientist working in that field could take the form: 'if you want *A* to happen, make *B* happen'. This is the sort of thing that medical scientists, for example, spend a lot of time doing – 'if you want to feel less pain, then take this medication'. Where, as in this case, the objective that is the basis for the recommendation – pain reduction – would be generally regarded as self-evidently desirable, this kind of statement by a scientist does not give rise to any problems. Often, the conditionality is so obvious and so uncontroversial, that it is not explicitly stated.

The recommendations that economists make can be regarded as conditional advice-type statements of this sort – 'if you want a healthy economy, then repeal the minimum wage legislation'. Although, the economist's and the doctor's statements both have an 'if . . . then . . .' structure, there are important differences between them. Whereas pain is experienced directly via the senses of an individual, 'economic health' is an abstraction defined with reference to many individuals. Exactly what a 'healthy economy' might be is itself something to be enquired into, and any definition must involve normative elements.

There are two sorts of reason why different economists come up with different recommendations – some disagreements have positive origins, some normative origins. Not all positive statements in economics have been definitively classified as true or false. Economists disagree as to how the economy actually works – some consider that minimum wage legislation increases unemployment, others that it does not. However, even if all economists agreed on the true/false classification of all possible positive statements about the workings of the economy, different recommendations could still follow from different appreciations of what 'economic health' is – economist Jack could consider it to require an unemployment below 3 per cent, while Jill could consider any level of unemployment below 10 per cent to be consistent with a healthy economy.

In so far as economists agree about recommendations, it is because they agree about both positive descriptions of how things work and normative criteria for assessing performance. At the level of studying individuals choosing between alternatives, we refer to the normative criteria that they use as 'preferences' or 'tastes'. Given that Jack could buy oranges or lemons, we say that what he actually buys is determined by his preferences as between oranges and lemons. In the context of analysing policy choices, we look at the normative criteria involved in terms of their basis in some ethical position. **Ethics**, or moral philosophy, is the study of the principles that ought to govern human conduct. One of its fundamental questions is: how do we decide whether or not an action is morally correct? There are two broad schools of thought.

According to deontological theories, moral correctness is a matter of fulfilling obligations, a matter of duty. According to consequentialist theories, moral correctness is to be judged in terms of the consequences that follow from an action. To illustrate the difference, consider the question: can it ever be right to tell a lie? The answer is 'no' on deontological criteria, 'yes' on consequential criteria. In the former case, it is argued that there is a universal duty to tell the truth. In the latter case, that there may be circumstances such that telling a lie produces a better outcome than telling the truth.

Utilitarianism is a particular variety of consequentialism. According to utilitarianism, the moral correctness of an action depends on the balance of pleasure and pain that it produces. Actions that increase the totality of pleasure or reduce the totality of pain are morally correct; actions that reduce the totality of pleasure or increase the totality of pain are morally incorrect. The term 'utility' refers to the situation of an individual in regard to the balance of pleasure and pain – pleasure is that which increases an individual's utility; pain is that which reduces an individual's utility. The term 'welfare' is used for the totality of utility across individuals, and according to utilitarianism morally correct actions are those that increase welfare. Utilitarianism is the ethical basis for economics.

There are three main questions for utilitarianism. First, whose utility counts? Second, how is utility assessed? Third, how is utility across individuals added up to get welfare? There are different varieties of utilitarianism according to the answers to these three questions. We will look at differences, and commonalities, between neoclassical and ecological economics in terms of these questions later in this chapter.

1.4 SUSTAINABILITY AND SUSTAINABLE DEVELOPMENT

The ideas of sustainability and sustainable development will figure very large in this book, as they are very important central ideas in ecological economics. **Sustainability** is:

> maintaining the capacity of the joint economy–environment system to continue to satisfy the needs and desires of humans for a long time into the future

If the joint economy–environment system is operating as required for sustainability, it is in a sustainable mode of operation, otherwise it is unsustainable. As subsequent chapters will explain, the difference between sustainable and unsustainable configurations for the economy involves questions about both the scale and the composition, in terms of the sorts of extractions from and insertions into the environment, of economic activity. The scholars who set up the International Society for Ecological Economics in 1989 were largely motivated by the judgement that the way the world economy was operating was unsustainable. They were concerned by what they judged to be threats to sustainability, features of current economic activity that could undermine the capacity of the joint economy–environment system to continue to satisfy human needs and desires. Climate change is an example of a threat to sustainability.

The idea that it is important to 'maintain' a capacity implies that it is sufficient. In fact, in the second half of the twentieth century many scholars argued that the capacity of the joint economy–environment system to deliver human satisfactions needed to be increased rather than maintained. A major feature of the current human condition is the existence of mass poverty. The generally accepted remedy for poverty is economic growth, increasing the scale of economic activity. Here is a major problem. On the one hand, many judge that the current scale of global economic activity threatens sustainability: threatens to reduce the future capacity to satisfy human needs and desires. On the other hand, many argue that it is necessary to increase the scale of economic activity to alleviate poverty. Dealing with poverty now, it seems, is going to create future economic problems, via the environmental impacts arising from increasing the scale of current economic activity.

One of the most important and influential publications of the last part of the twentieth century was *Our common future*. This report by the World Commission on Environment and Development, WCED, was published in 1987, two years before the formation of the International Society for Ecological Economics. It is sometimes referred to as the 'Brundtland Report', Ms Brundtland having been the commission's chair. *Our common future* described both the extent of poverty and the various threats to sustainability. It argued that the circle could be squared, that the economic growth required to deal with poverty need not, via its environmental impacts, create future economic problems. What was needed, the **Brundtland Report** argued, was a new kind of economic growth that had much less environmental impact and which, rather than threatening sustainability, actually increased the joint economy–environment system's capacity to deliver human satisfactions. It argued that what was needed could be done, and called it **sustainable development**. It is:

> a form of economic growth that would meet the needs and desires of the present without compromising the economy–environment system's capacity to meet them in the future.

1.5 THE RELATIONSHIP BETWEEN ECOLOGICAL AND NEOCLASSICAL ECONOMICS

In this section we want to look at the broad relationship between ecological and neoclassical economics in terms of the normative and positive elements of both.

The first question about utilitarianism that we noted was: whose utility counts? In economics, ecological and neoclassical, the answer is: all of the humans who are affected by the action. There is no reason, in principle, why utilitarianism could not take account of the pleasure/pain of all affected animals. Some moral philosophers belonging to the utilitarian school argue that in working out the balance as between pleasure and pain, all affected beings capable of feeling pain and pleasure should be accounted for. If this argument were accepted, welfare would depend on the utilities of all 'sentient' beings, not just on the utilities of humans. The suggested candidates for consideration along with humans have mainly been the higher mammals. Normative economics does not take account of the utilities of non-human beings. It is **anthropocentric** in that the effects of an action on non-human beings are taken into account only in so far as they produce pain or pleasure for human beings. If no humans feel (mental) pain on account of animal suffering caused by an action, then that suffering does not figure in the calculation of the pleasure/pain balance to be used to judge the action. If any human does feel pain, that pain, not the animal suffering, does figure in the pleasure/pain balance. Also, if any human feels pain on account of the damage to a non-sentiment entity, such as a building for example, then that should be accounted for in evaluating the action responsible for the damage and the pain.

In terms of the answer to this first question, there is no difference at all between ecological economics and neoclassical economics. Both are anthropocentric, as well as utilitarian. In regard to the second question – how is human pleasure/pain to be measured? – there are some differences. In neoclassical economics, each affected human individual is the sole judge of whether her utility has increased or

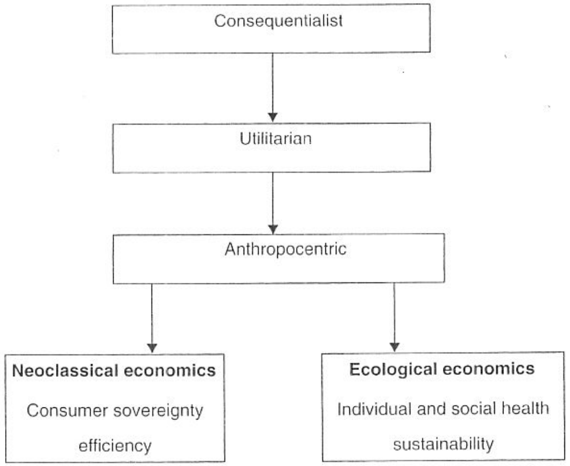


Figure 1.3 Ethical positions of neoclassical and ecological economics.

decreased. The change in an individual's utility is measured solely in terms of the preferences of that individual. Individual preferences are taken as given, and are not subject to any moral evaluation. This is sometimes referred to as the doctrine of 'consumer sovereignty'. Ecological economics does not ignore individual preferences, but it treats them neither as sovereign, nor as the only source of normative criteria.

In neoclassical economics, provided it can be assumed that an individual is in possession of all relevant information, there can be no ethical basis for seeking to change his preferences. There can be no basis for saying that a taste for cycling should be encouraged, while a taste for driving motor cars should be discouraged. In ecological economics, there can be an ethical basis for comparing, evaluating and seeking to change tastes. Ecological economists would be sympathetic to the argument that tastes should be educated in the direction of cycling and away from motoring on the grounds that more cycling and less motoring promotes individual and social health. They consider sustainability to be a requirement of social health. In ecological economics, sustainability requirements are a source of normative criteria. Figure 1.3 summarises the discussion thus far of the ethical underpinnings of neoclassical and ecological economics.

We now look at the third question about utilitarianism – how to add up increases and decreases in utility across affected human individuals so as to get welfare. To make things simple, assume that there are just two individuals, identified as A and B, and use U^A and U^B to represent their utility levels, and W to represent welfare. Then simple addition for welfare would be

$$W = U^A + U^B$$

The problem that some see here is that this way of getting from utilities to welfare takes no account of the relative positions of A and B. Suppose that A's utility is much higher than B's, and that the action being considered would increase A's utility by more than it decreased B's. According to simple addition this would increase welfare, and the action would be morally correct, though it makes the better-off even better-off and the worse-off even worse-off. This, to many utilitarians, does not seem fair.

They would argue that welfare should be defined as a weighted sum of individual utilities with more weight being given to the utility of those whose utility is low. Instead of simple addition, this argument is, proposed actions should be assessed using

$$W = (w_A \times U^A) + (w_B \times U^B)$$

where w_A and w_B are the weights to be assigned to the utilities of A and B respectively. This becomes simple addition if $w_A = w_B = 1$. For B with lower utility than A, the argument would be that w_B should be larger than w_A. Suppose that $w_A = 1$ and $w_B = 5$, for example. An action that increased A's utility and decreased B's would have to increase A's by five times as much as it decreased B's in order to be considered morally correct. The choice of the weights is itself an ethical issue. Ecological economists tend to be more inclined to argue for the use of weights that favour the less well-off than do neoclassical economists. They tend, that is, in judging alternative policies to be more concerned with the equity dimensions of the choice than neoclassical economists are. While neoclassical economists do not ignore **equity** issues, they focus more on policies to promote **efficiency**, a situation where it is not possible to increase one person's utility without reducing that of one or more other persons.

Sustainability and sustainable development are central concerns of ecological economics, which has been defined as the science of sustainability, but not of neoclassical economics. In part this is because of the differences in ethical positions just described, normative differences. But it is also because of differences about positive matters, questions of fact. Ecological economists judge that serious threats to sustainability exist, and they are somewhat sceptical about the feasibility of sustainable development. Neoclassical economists do not claim that there are no threats to sustainability, but they judge them to be less serious than do ecological economists, and they tend to believe that sustainable development will come about given some relatively minor policy changes. They have confidence in the ability of markets to drive technological and behavioural changes that will enable the capacity of the economy–environment system to satisfy humans to go on increasing. Ecological economists have less confidence in markets and technology. They tend to believe that solving the problem of poverty cannot be left to economic growth alone, but will require the redistribution of income and wealth from the better- to the worse-off.

Earlier, we said that positive statements are the business of science, and that differences over their validity can be resolved by appeal to evidence, as a matter of principle. This is a useful way to distinguish positive from normative statements, because differences over the latter cannot be so resolved, even as a matter of principle. But, without keeping firmly in mind the qualification 'as a matter of principle', the statement about positive statements can mislead. Science has not yet sorted all positive statements into true and false classes, and it never will. It has been very successful where controlled experiments are possible, much less where they are not. Many of the positive issues that divide ecological and neoclassical economists are not amenable to definitive resolution by controlled experiment.

Again, the example of climate change can be cited. Most scientists working in the field consider that the global climate is changing, and that this is due, mainly, to the release into the atmosphere of greenhouse gases by human activities such as burning coal, oil and gas. There are some scientists who dispute that the global climate really is changing. There are others who accept that the global climate is changing, but dispute that the cause is human activity. All agree that the atmospheric concentration of these gases is one of the things that influences climate, and that humans have been releasing increasing amounts of these gases into the atmosphere since 1750. The problem is that in the historical record all of the things that influence global climate have been changing, so even if everyone accepted that climate had been changing, it would not be possible to definitively establish whether or not that was due to human activities. Doing that would require a controlled experiment where human releases of the gases were held constant at the 1750 level, while all the other influences on climate behaved as they did in history since 1750. That is not possible.

The construction of a model is a response to this kind of problem. A **model** is a simplified version of the set of relationships which are thought to determine some phenomenon. In principle, a model can be stated in several ways – using language, constructing a physical system, drawing graphs, as a set of equations. Most usually, and most effectively, models are stated mathematically, as sets of equations. A model is a substitute for a controlled experiment. The investigator can turn relationships in the model on and off to see what difference it makes to the model outcome. This is exactly how climate scientists investigate the role of the various influences on the global climate – they run their model of that phenomenon with and without, for example, the history of human greenhouse gas emissions since 1750, so as to see what difference those emissions make.

The problem is, of course, that the model is a model. Ideally, it incorporates accurately all of the relationships that actually do have a role in determining the phenomenon being investigated. In practice, as in the climate change case, different investigators have different models because it is not definitively known which those relationships are. What happens is that an investigator reviews previous work in the field from which she selects the relationships that she judges to be the ones that a model needs to incorporate. The resulting model is then tested by seeing whether it can replicate to a reasonable degree of accuracy the behaviour of the phenomenon of interest as observed in the historical data. If it is judged that it does replicate history satisfactorily, then it is used to conduct 'what if?' investigations, experiments, by modifying the relationships that it includes. One type of 'what if?' experiment is forecasting – using the model to predict the behaviour of the phenomenon of interest conditional on assumptions about how the things at the other end – to it – of the included relationships behave.

In the last few paragraphs we have often used the word 'judgement', and sometimes 'belief'. Many of the positive issues that neoclassical and ecological economists disagree about are matters to be investigated by modelling rather than controlled experiment. While there are certain agreed conventions about how to decide whether or not a relationship has a role in determining some phenomenon – these are the rules of statistical inference – their application necessarily involves judgement. Two equally honest and skilled investigators can quite reasonably come up with different models for the same phenomenon. Similarly, the application of the conventions for deciding whether a model explains the phenomenon satisfactorily is a matter of judgement.

Many of the differences between neoclassical and ecological economics are differences about the models judged to be useful in explaining various economic and environmental phenomena, and, therefore, predicting what will happen to those phenomena. For example, a fundamental judgement of ecological economics is that a useful explanation – model – of the rapid growth in the average level of consumption of goods and services in the industrial economies in the last 200 years must include relationships describing economy–environment interdependence. Figure 1.2 presents a very simple version of such a model as a picture. Figure 4.1 will present a less simple version as a picture. Some such model of economy–environment interdependence is the starting point for ecological economics. The judgement in neoclassical economics is that these relationships are not an essential part of a useful model of economic growth. Their existence is not denied. It just does not figure in the core models by means of which students are introduced to the study of economics. As ecological economists, the authors of this book judge that to be a major failing on the part of neoclassical economics, which is why we have written an introductory ecological economics textbook.

That said, it also needs to be stated, and emphasised, that there are very many, important, positive questions where ecological and neoclassical economics are in agreement.

1.6 A GUIDED TOUR

There are four parts to this book. Part I, 'Interdependent Systems', explains properly the necessary interdependence of the economy and the environment sketched in outline terms in Figure 1.2. Chapter 2, 'The environment', reviews the basic environmental science necessary for an understanding of ecological economics. Chapter 3, 'Humans in the environment – some history', looks at the evolution of economy–environment interdependence in human history. Chapter 4, 'The economy in the environment – a conceptual framework', sets out our basic model of the current relationships between economic activity and the natural environment.

Part II, 'Economic Activity', focuses on the modern industrial economy and the means by which it is mainly organised, the market system. Chapter 5, 'Economic accounting', sets out the framework used for economic analysis, and explains how GDP and the like are measured and what they mean. GDP growth has been the dominant feature of the economic history of the last few hundred years. Chapter 6, 'Economic growth and human well-being', looks at explanations for the phenomenon and at the relationship between it and human well-being. Economic growth is widely seen as the only way to eliminate poverty. However, the facts of economy–environment interdependence have led many to ask whether the environment can accommodate further growth of global GDP. Chapter 7, 'Economic growth and the environment', uses the model of economic growth, and some other models, to look at whether sustainable development is feasible.

The market system, is now the dominant mode of economic organisation. Chapter 8, 'Exchange and markets', explains how markets work and how they make possible the realisation of the benefits – in terms of efficiency – that specialisation and exchange offer. Chapter 9, 'Limits to markets', explains why economic organisation cannot be left entirely to markets, why there is a role for government. As explained there, markets are often absent, or function badly, in relation to the regulation of economy–environment interdependencies as is required for sustainable development. If this is going to happen, it requires government to guide market forces in the necessary directions.

This is what Part III, 'Governance', is about. In thinking about what government does it is useful to distinguish between questions about ends and means. Chapter 10, 'Determining policy objectives', is about setting the ends at which policies should be directed so as promote sustainability – how much pollution should be allowed, for example. Chapter 11, 'Environmental policy instruments', is about the means by which the ends decided on should be pursued – how to control the activities of polluters, for example.

The sustainable development problem is a global problem in both its economic and its environmental dimensions, but there is no world government. Human society is organised around the institution of the nation state. Part IV, 'The International Dimension', is concerned with this mis-match and some of its implications. Chapter 12, 'A world of nation states', looks at trade between nations, at the ways in which some environmental problems cross national borders, and at the institutions that have been developed to address the, many, problems that require coordination and cooperation between nation states. The last two chapters – 13, 'Climate change' and 14, 'Biodiversity loss' – draw on the look at two major, and related, problems of this kind, which are major threats to the prospects of realising sustainable development.

Nowhere in the book is any prior knowledge assumed – it is an introductory text. You should be able to use this book successfully even if you have not previously studied either economics or ecology. No familiarity with environmental science is assumed. Nor is any mathematical ability beyond arithmetic assumed, a matter to which we return in a moment. Those who come to the book with some previous knowledge of some of the fields covered can be selective in their use of the various chapters. At the start of each chapter there is a statement of what it is going to cover, and at the end there is a summary and a list of keywords, with page references, and their meanings. These should help you to use the book effectively. At the end of each chapter there is a section on Further Reading, which is intended to guide those who want to take their study of ecological economics further.

You may well have flicked through the book by now and formed the impression that what looks like mathematics appears quite a lot. Your impression is quite right, but, even if you consider yourself somewhat weak as far as mathematics goes, you have no cause for concern. There is in this book quite extensive use of arithmetic and simple algebra, where that is the simplest and most efficient way of getting across the basic ideas at an introductory level – as it often is. But, be assured, there is nothing beyond arithmetic and simple algebra, and every time either is used it is explained very carefully. Most of the time, it is just arithmetic. The most advanced algebra used is the solving of (easy) pairs of simultaneous equations. In a few places, the algebra is simple but tedious and it has been put in an appendix.

In the text we often make use of simulations, and some of the exercises invite you to deepen your understanding by doing your own simulations. A **simulation**, as we will explain in detail when we get to the first one in the next chapter, is just doing repetitive arithmetic to study the time paths for variables determined by a model. This is a simple way to explore the properties of a model. It may sound hard and/or tedious, but it is not hard and need not be tedious. To do the arithmetic easily and accurately, all you need is a calculating machine that will – as well as add, subtract, multiply and divide – raise numbers to powers, and give logarithms and anti-logarithms. Doing the arithmetic this way is easy, but can be tedious. The way to avoid the tedium is by automating the arithmetic using the copy-and-paste formula facility of a spreadsheet, such as Excel^{TM}, for a PC. We will not go into the details of this – if you are not already familiar with such facilities, you will need a course or a book about Excel^{TM}, or whatever spreadsheet you are going to use. We will, however, for every simulation that we introduce, spell out the arithmetic that needs to be done. Once you get the hang of simulating models it is a very powerful way to learn about the properties of different kinds of systems.

Finally, we need to come back to the question of the relationship between ecological and neoclassical economics, and how we deal with it in this book. In the section of this chapter we looked, in general terms, at the relationship in its normative and its positive dimensions. We noted that in both dimensions there is much common ground, as well as areas where they diverge. Much of what you will learn from this book carries over into neoclassical economics. If, that is, you go on to study more advanced economics of a basically neoclassical kind you will not have to unlearn what you have learned from this book. What you have learned here should, however, give you a different, and often more critical, perspective there than would be the case had you not been introduced to economics via ecological economics.

The purpose of this book is to introduce you to ecological economics, not to develop a critique of neoclassical economics. On the other hand, exposure to different ideas, and the origins of the differences, is part of learning about economics as an active field of enquiry and debate, rather than just a repository of established truth. There is a choice to be made here, as economics teaches that there is almost everywhere. In this case it is between a very long, but comprehensive book, and a fairly long book that concentrates very much on telling the ecological economics story and largely neglects differentiating that story from the neoclassical story. We have chosen the latter option. We explicitly compare and contrast only when that is necessary for understanding ecological economics. For those who are interested we will provide references to works that do more of the compare and contrast sort of thing.

KEYWORDS

Anthropocentric (p. 9): centred on human beings.

Brundtland Report (p. 9): *Our Common Future* (1987) put the idea of sustainable development on the political agenda.

Classical economics (p. 3): the economic thinking of the first half of the nineteenth century.
Ecology (p. 1): the study of the relations of animals and plants to their organic and inorganic environments.
Ecological economics (p. 1): the study of the human economy as part of nature's economy.
Economics (p. 1): the study of how humans satisfy their needs and desires.
Efficiency (p. 11): a situation where nobody can be made to feel better-off except by making somebody else feel worse-off.
Environmental economics (p. 4): the specialisation within neoclassical economics that is concerned with the economy's insertions into the natural environment.
Equity (p. 11): the question of fairness.
Ethics (p. 7): the study of the principles that ought to govern human conduct.
Model (p. 12): a simplified version of the set of relationships which are thought to determine some phenomenon.
Neoclassical economics (p. 3): the currently dominant school of economics.
Natural resource economics (p. 4): the specialisation within neoclassical economics that is concerned with the economy's extractions from the natural environment.
Simulation (p. 15): numerical analysis of the properties of a model.
Sustainability (p. 8): maintaining the capacity of the joint economy–environment system to continue to satisfy the needs and desires of humans for a long time into the future.
Sustainable development (p. 9): economic growth that would meet the needs and aspirations of the present without compromising the ability to meet those of the future.
Utilitarianism (p. 7): the school of ethics according to which the moral correctness of an action depends on the balance of pleasure and pain that it produces.

FURTHER READING

This book is written as an introductory text, and it is assumed that many, but not all, readers will be going on to do, or concurrently be doing, other courses in economics, ecology, environmental science and management. The suggestions for further reading reflect these assumptions. Where a reading is marked with an asterisk: *, this indicates material that should prove useful to all readers, that could be regarded as a source of desirable supplementary reading in an introductory course centred on this book, that is at a similar level to this text. Otherwise, the suggestions take things further and/or move up a level in technical difficulty, and are there primarily for those not going to get exposed to more advanced material in other courses.

The standard text on the history of economics is Blaug (1985). It is really for specialists, and Barber (1967) and, especially, Heilbronner* (1991) are more accessible and cover the essentials. Crocker (1999) covers the rediscovery of the environment by neoclassical economics, while the first part of Costanza *et al.* (1997a) deals with the emergence of ecological economics. Important journals covering neoclassical economics work on the natural environment are: *Journal of Environmental Economics and Management*, *Environment and Resource Economics*, *Land Economics*, and *Environment and Development Economics*. *Ecological Economics* is the journal of the International Society for Ecological Economics.

Singer (1993) is a good introduction to ethics which considers environmental applications. Brennan* (2003) is a recent survey of philosophical writing on matters environmental. Glasser (1999) is a survey article on ethics and environmental policy. Sen (1987) looks at ethics in relation to economics. The journal *Environmental Values* publishes articles by people from a variety of academic disciplines on ethics and the environment. *Ecological Economics* often carries papers about ethics and philosophy, and the February/March 1998 issue (vol. 24, nos. 2 and 3) was a special issue on 'Economics, ethics and the environment'.

WEBSITES

The address of the website of the International Society for Ecological Economics, ISEE, is http://www.ecoeco.org. It has links to a number of other relevant sites. One of the features of the ISEE site is the ongoing assembly of an online encyclopedia of ecological economics, in which there is an entry on 'The early history of ecological economics and ISEE' written by Robert Costanza, one of the founders of ISEE.

DISCUSSION QUESTIONS

1. Should sustainability be an objective of government policy?
2. Is mathematics a science? Is history?
3. Is utilitarianism that takes account of all sentient beings feasible?

PART I
INTERDEPENDENT SYSTEMS

Ecological economics starts from the fact that human and natural systems are interdependent. The environment is the material base for economic activity.

Chapter 2 will explain those aspects of the functioning of environmental systems that are particularly relevant to an understanding of economy–environment interdependence. The nature of that interdependence has changed a great deal in the course of human history, as is explained in Chapter 3. Chapter 4 develops a conceptual framework, a model, for the study of the way a modern economy interacts with its environment.

2

The environment

In this chapter you will:

- Learn about the ways in which the natural environment functions and sustains life;
- Look at the first and second laws of thermodynamics;
- Learn about energy and nutrient flows in ecosystems;
- See how the fossil fuels came into existence;
- Study population dynamics;
- Consider the concept of ecosystem resilience;
- Learn about global nutrient cycles;
- Look at evolutionary processes.

In the previous chapter we introduced the idea that the economy and the natural environment are interdependent systems, with the economy located within the environment. That idea is to be developed in the following chapters of Part I. This chapter looks at the functioning of the natural environment itself, largely ignoring the role of humanity. It is a simple, and brief, overview of the material from environmental science that is necessary for an understanding of ecological economics. Readers who are familiar with environmental science will find that they can get through the chapter quickly, though they probably should not skip it completely. For other readers, the Further Reading section at the end of the chapter offers some guidance on how to go further into the environmental science topics introduced here.

This chapter is organised as follows. First, in section 2.1, we look at the planet in terms of four interacting systems. Section 2.2 is about thermodynamics, the science of energy. Some appreciation of the essentials of thermodynamics is essential for understanding the way that the planet works, and particularly the nature of life on earth, which is dependent on energy. In section 2.3, we shall explore various aspects of the organisation of life on earth by considering ecosystems, which are systems of interaction among living organisms. Life requires matter as well as energy, and in section 2.4 we will look at some of the important cycles of matter through the planetary systems. The ways in which planetary systems, especially the living

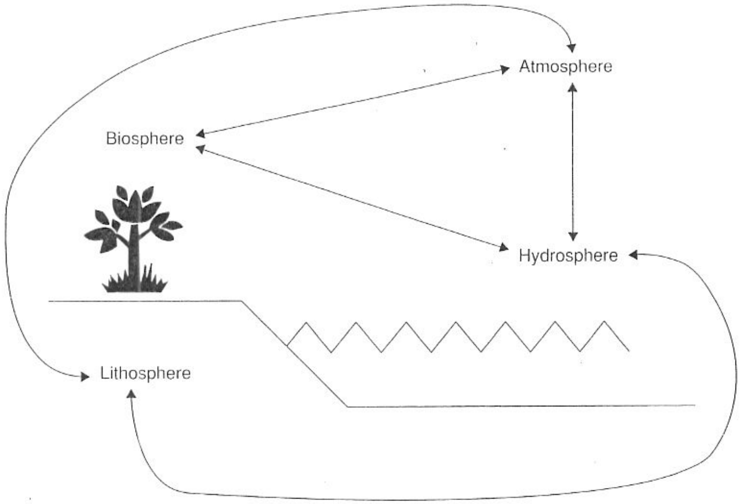


Figure 2.1 Four interacting environmental systems.

systems, work have changed through the history of the planet, and the chapter finishes, in section 2.5, by looking briefly at some aspects of that coevolutionary history.

2.1 PLANET EARTH

By 'the natural environment', or just 'the environment', we mean planet earth. It is one of nine planets in the solar system, and is, as far as we know, the only one that supports life. The system that is planet earth can itself be seen as comprising four main systems:

(1) **Lithosphere** – the solid outer shell of the earth;
(2) **Hydrosphere** – the water on or near the surface of the earth;
(3) **Atmosphere** – the gases surrounding the earth's surface;
(4) **Biosphere** – living organisms and their immediate environment.

As indicated in Figure 2.1, these systems all interact with one another. We will discuss some aspects of the interactions later in the chapter. Here, after saying something about the idea of a system, we will concentrate mainly on simple descriptions of the four systems considered separately.

2.1.1 Systems

A **system** is a set of components that interact with each other. The idea of a system necessarily entails the idea of an environment within which the system exists, and of a boundary between the system and its environment, as illustrated in Figure 2.2.

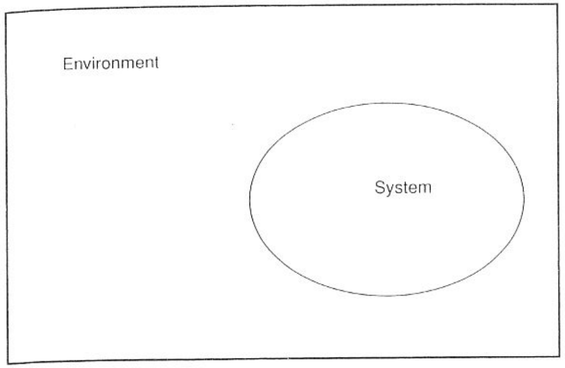


Figure 2.2 System and environment.

In this context, the term 'environment' has nothing necessarily to do with the natural environment. It simply means what is outside the boundary of whatever system is under consideration.

A system must be distinguishable from its environment. There is no general and precise rule about what establishes distinguishability. The definition of a system and its environment cannot be reduced to a set of precise rules that apply in all circumstances. What is distinguished as a system will depend on what the purpose of the exercise is. For some purposes, it will make sense to treat planet earth as a system and consider it in relation to its environment. For others, it will make sense to treat, say, the hydrosphere as a system with its environment comprising the other three planetary systems.

One way of thinking about some standard academic disciplines is in terms of systems definitions and boundaries. Biology is concerned with things in the biosphere, hydrology with things in the hydrosphere, geology with things in the lithosphere, and so on. Systems analysis is a discipline which takes the position that there are insights to be gained from focussing on systems as such, rather than the particular natures of their component parts. A distinguishing characteristic of systems analysis is that it is as concerned with the nature of the interactions as it is with the nature of the components. One of the ways in which it is useful is that it turns out that there are patterns of interaction arising across quite different sorts of assemblages of components. The fact that there are such patterns means that knowledge gained about the behaviour of one system can be applied to the behaviour of a system with quite different components, if it can be established that the two systems have the same pattern of interactions. This sort of transfer can occur across the boundaries of conventionally defined disciplines. Lessons learned about system characteristics in, say, ecology, can be applicable in, say, economics.

2.1.2 The lithosphere

The lithosphere comprises the upper part of the mantle and the earth's crust. While the lithosphere as such is geologically important, especially in regard to volcanic activity, it is really the crust that is of interest here as this is the part of the lithosphere that interacts with other environmental systems. The crust comprises less than 1 per cent of the earth's mass, and about 0.5 per cent of its radius. Its thickness varies from 35 km to 5 km. The crust is made up of rocks which are composed of minerals. Over 2,000 minerals are known to exist. Just 8 of the 100 plus chemical elements known to exist account for over 99 per cent of the mass of the earth's crust. Oxygen accounts for 47 per cent, silicon 28 per cent, aluminium 8 per cent, iron 5 per cent.

Rocks are classified according to the way in which they were formed. Igneous rocks (e.g. granite) arise from the solidification of molten material, magma, originating in the earth's mantle. Sedimentary rocks (e.g. sandstone, limestone) come into being as the result of erosion, or as the result of dissolved material precipitating from water, or as the result of biological activity. Metamorphic rocks are the result of the alteration of some parent rock (e.g. marble from limestone) by extreme heat and/or pressure.

Rocks are, over geological time, created, modified and destroyed in cyclical processes driven by energy which comes from the cooling of the interior of the planet, radioactive processes, and the sun. The processes involved are, relative to the processes involved in the other three environmental systems distinguished here, very slow – they operate on timescales of millions of years. From the human perspective this is so slow as to be imperceptible, and for many purposes features of the lithosphere are taken as unchanging. From the human perspective, the lithosphere system is of direct economic interest mainly on account of the formation of exploitable mineral deposits, by geological processes, and soils, by climatic and biological processes.

2.1.3 The hydrosphere

The hydrosphere includes oceans, lakes, rivers and water vapour in the atmosphere. Approximately 70 per cent of the earth's surface is covered with water, and 10 per cent of the land is covered with ice. Of the total amount of water, about 97 per cent is stored in the oceans and 2 per cent in ice caps and glaciers. Water vapour in the atmosphere accounts for 0.0001 per cent, and lakes and rivers 0.009 per cent, of total water.

The basic general process involving water is the hydrological cycle. Driven by the energy of solar radiation, water evaporates from the oceans, lakes and rivers, and from soil, to become water vapour in the atmosphere. Precipitation returns water to the oceans directly when it falls on them, and indirectly when it falls upon land from which it reaches the oceans via rivers. The processes involved in the hydrosphere are much faster than those of the lithosphere. The average length of time that a water molecule remains in one of its stores varies from days, in the case of residence in the atmosphere as water vapour, to thousands of years, in the case of residence as salt water in the oceans.

Water is important for several, related, reasons. It is directly necessary for life: see the discussion of plants and animals in section 2.2 below. Many elements dissolve in water, and are thereby dispersed through the lithosphere and the atmosphere by the operation of the hydrological cycle. It therefore plays a key role in the major bio-element, or nutrient, cycles to be discussed in section 2.4 below. Water is strongly involved in most of the climatic and biological processes by means of which soils are produced from rocks.

2.1.4 The atmosphere

The atmosphere is predominantly a mixture of gases, though it also contains particulate matter. The most abundant gases are nitrogen, approximately 78 per cent of the total volume, and oxygen, 21 per cent. All the other gases, then, together comprise only 1 per cent of the atmosphere. This does not mean that these other gases are unimportant. For example, carbon dioxide (0.04 per cent of the total volume) and methane (0.0002 per cent) are 'greenhouse gases', variations in the amounts of which in the atmosphere affect the global climate system. It should also be noted that, given the size of the atmosphere, a small concentration of a gas goes with a large absolute amount in the atmosphere – the total amount of carbon dioxide in the earth's atmosphere is approximately 2,800 Gigatonnes. Box 2.1 explains that the prefix 'Giga' means thousands of millions.

The boundary between the earth's atmosphere and space is not a sharp one. But, in effect all of the atmosphere lies within 80,000 km of the surface of the earth, and 99 per cent within 50 km of the surface. The troposphere extends up to between 8 (at the poles) and 16 km (at the equator) above the surface of the earth, and contains about 75 per cent of the mass of the atmosphere, including about 90 per cent of the particulate matter and water vapour. It is the only part of the atmosphere where the temperature is above 0°C, and is where weather patterns are mainly determined. The stratosphere is the region of the atmosphere immediately above the troposphere, and extends up to about 60 km above the surface of the earth. Although processes in the stratosphere are not so directly and closely related to circumstances at the earth's surface, they are still important. For example, the stratosphere is where the ozone that screens the earth's surface from the sun's ultraviolet radiation resides. That ultraviolet radiation is harmful to organisms, and life on earth would be impossible without the presence of the stratospheric ozone layer.

Atmospheric processes operate on timescales that are more similar to hydrological processes than they are to geological processes. The atmospheric residence times of the principal greenhouse gases, for example, vary from the order of 10 years for methane to that of the order of 100 years for carbon dioxide and nitrous oxide.

2.1.5 The biosphere

The biosphere is that part of the earth in which living things, i.e. biota, exist. It includes parts of the lithosphere, the hydrosphere and the atmosphere. The biosphere extends from the top of the troposphere to about 10 km below sea level. That is a maximum vertical extent of about 25 km. The radius of the earth is about 6,400 km, so the biosphere is a very thin layer of the earth – less than 0.4 per cent of the radius. In fact, only pollen grains and spores and a few species of insects and birds can exist at more than 6 km above sea level, so that most life on earth exists within a layer which is about 16 km deep.

The conditions that enable the biosphere to support life are: a supply of water; a supply of usable energy; a supply of air; a suitable temperature range; the presence of essential nutrients and trace elements. That these conditions exist is due to the fact that the biosphere is located where the systems which are the lithosphere, the hydrosphere and the atmosphere interact. These conditions, and hence life as we now know it, have not always existed on planet earth, as will be discussed in section 2.5.

The functioning of the biosphere will be considered in section 2.3. In order to be able to do that, we need to look at thermodynamics.

2.2 THERMODYNAMICS

Thermodynamics is the study of energy transformations. The laws of thermodynamics are fundamental to an understanding of the operation of environmental systems. It follows that they are also fundamental to an understanding of the operation of economic systems, and although this chapter is basically about the natural environment, in this section we will say something about the economic implications of the laws of thermodynamics.

2.2.1 Energy, heat and work

In order to state and explain the laws of thermodynamics, we need to begin with some definitions. Energy transformations involve work, heat and energy. **Energy** is the potential to do work or supply heat. Work is what is done when something is moved, and the amount done is the product of the force applied and the distance moved. The possible effects of heat on a substance are an increase in temperature, expansion, a change of state (melting of a solid/vaporisation of a liquid), or an increase in pressure. Energy, work and heat are all measured in the same units. In the SI system, the basic unit is the joule. In work terms one joule is the work done when one kilogram is moved one metre. It is also the heat required to raise the temperature of one cubic centimetre of water by 0.239°C. For many purposes the joule is an inconveniently small unit. Box 2.1 gives standard prefixes used with joules, and other small measurement units, to specify larger units which are often more convenient. Box 2.1 also gives some conversion factors for SI and other systems.

Power is work per unit of time. The unit of power corresponding to the joule is the watt, which is one joule per second. As the potential to do work or supply heat, energy can take a variety of forms. Potential energy exists by virtue of position, as in the case of water in an elevated lake. Kinetic energy exists by virtue of motion, as in the case of flowing water. Radiant energy is given off by hot objects, as with the solar energy given off by the sun. Electrical energy is carried by a flow of charged particles in a conductor. Chemical energy is that given off in chemical reactions such as the combustion of coal.

2.2.2 First law of thermodynamics

The first law of thermodynamics says that energy can be converted from one form to another, but can be neither created nor destroyed. Consider a coal-fired electricity

Box 2.1 Energy measurement

The basic SI unit of measurement for energy (and heat and work) is the joule. It is a very small quantity – the work done when 1 kilogram (kg) is moved 1 metre (m). Energy, heat and work are therefore usually measured in units which are multiples of the joule. The same multiples can be used with other basic SI units, such as the gramme (g) for mass, the metre (m) for distance and the litre (l) for volume.

A simple standard mathematical notation is frequently used in defining and using these multiples. Consider the number 5 million, i.e. 5,000,000. One million is 1,000,000 which is equal to 100 multiplied by 100 multiplied by 100, i.e. 100 × 100 × 100. One hundred is 10 multiplied by 10, i.e. 10 × 10, which is 10 squared or 10 raised to the power 2, written 10^2. So one million is equal to 10 multiplied by 10 six times, i.e. 1,000,000 equals 10 × 10 × 10 × 10 × 10 × 10, which is 10 raised to the power 6, written as 10^6. The number 5 million can be, and frequently would be, written as 5×10^6. In the same way, one thousand is 10 to the power 3, written as 10^3, and 5,000 could be written as 5×10^3, while one billion (one thousand million throughout this book) is 10 to the power 9 and 5 billion (5,000,000,000) could be written as 5×10^9.

The following table lists the word prefixes used for the standard multiples, the corresponding symbols or abbreviations, the size of the multiple in power of ten notation, and – to indicate the usefulness and economy of the power notation – the corresponding number in standard arithmetic form.

Prefix	*Symbol*	*Multiple as power of 10*	*Multiple*
hecto	h	10^2	100
kilo	k	10^3	1,000
mega	M	10^6	1,000,000
giga	G	10^9	1,000,000,000
tera	T	10^{12}	1,000,000,000,000
peta	P	10^{15}	1,000,000,000,000,000
exa	E	10^{18}	1,000,000,000,000,000,000

To give some sense of the orders of magnitude here, consider electricity supply from a coal-burning power station. The size of such a plant is usually discussed in terms of the maximum amount of power that it could send out. Recall that the basic unit for power is the watt, which is one joule per second. The size of a typical modern coal-fired electric power plant is 1,000 megawatts, or 1,000 Mw. If the plant ran at maximum power for 1 hour it would send out 1,000 megawatt hours, 1,000 Mwh, of electrical energy. For a thermal efficiency of 33 per cent, that would mean burning an amount of coal with chemical energy content 10800000 (= 3,000 × 60 minutes × 60 seconds) Mj, or 10,800 Gj, or 10.8 Tj. From the definition of a joule as the heat required to raise the temperature of one cubic centimetre of water by 0.239 C, and assuming an ambient temperature of 15 C, it would require 180,000 joules, or 0.18 Mj, to bring a 0.5 litre kettle of water to the boil. In one hour the power plant could boil 20 million such kettles (1,000 × 60 × 60 is 3,600,000 Mj, which divided by 0.18 is 20×10^6 kettles).

The use of measurement units based on the joule is now widespread in energy analysis, but it is by no means universal, and conversions as between the units used in different sources can be tedious.

In the SI system there is another basic unit for energy/heat/work which is the calorie. One calorie is the heat required to raise the temperature of one gramme of water by one degree centigrade. It is approximately equal to 4.2 joules. The use of the calorie as the basic unit is particularly widespread in analysis of the chemical energy of food, and in discussions of weight-loss programmes. Again, as the calorie is a small amount, such analysis is usually reported in terms of kilocalories, often written as kcals, or sometimes as Cals. The amount of food energy required by a human adult varies with her size, the ambient temperature, and the activities engaged in. A widely used figure for an average human adult leading a moderately active live is 2,500 kcals per day. In terms of joules this is 2,500 × 1,000 × 4.2 = 10,5000,000, which is often stated as 10 Mj per day, or 10 × 365 = 3,650 Mj per year.

Not all sources use SI units. Particularly, but not exclusively, in economic analysis involving energy originating in the USA the basic unit is the British Thermal Unit, BTU. One BTU is the amount of heat required to raise the temperature of one pound of water by one degree Fahrenheit. It is equal to 1,055 joules, and for many purposes the approximation of 1 BTU equals 1,000 j will suffice. The BTU is quite a small amount, and a widely used unit based on it is the Therm, which is 100,000 BTU, and, therefore, approximately equal to 100,000,000 joules, i.e. 100 Mj. A very large unit in this system of measurement is the Quad, which is 10^{15} BTU, approximately 10^{18} joules or 1 Ej (exajoule).

As will be discussed in later chapters, in the modern economy the fossil fuels – coal, oil and gas – are the dominant source of energy. In many sources, data on the fossil fuels are reported in the mass and volume units, rather than in energy units. Thus, coal is typically measured in metric tonnes, where one tonne is 1,000 kilograms, which is approximately 0.98 Imperial tons. Oil is frequently measured in units called 'barrels', which refers to the size of the barrels used to transport oil away from the world's first oil well, opened in the 1860s, in Pennsylvania. One barrel is 42 US gallons, approximately 35 Imperial gallons. Oil varies in weight. The number of barrels to a tonne of oil varies from approximately 6.5 to 8, with an average of about 7. Gas is often measured in cubic feet or cubic metres, and multiples thereof.

For many purposes it is useful to be able to compare, or add, across quantities of the different fossil fuels, and alternative energy sources such as nuclear power or wind power, in common, energy, units. This is sometimes done by expressing everything in terms of tonnes of coal, or oil, equivalent. More usually, and more usefully, nowadays the more common practice is to express everything in SI energy terms based on the joule. The following are conversion factors that can be used for this purpose:

Fuel	*Quantity*	*Gj*
Coal	1 tonne	29
Oil	1 tonne	42
Gas	1 tonne	55

These are approximate averages. Just as the weight of a barrel of crude oil varies a little according to where it comes from, so does the exact amount of heat released when it is burned. The same goes for coal and gas – the heat content of one tonne of, say, east-coast US coal is not exactly the same as the heat content of one tonne of coal mined in, say, Queensland in Australia. Whereas the average heat content of 1 tonne of UK coal is 26 Gj, the figure of 29 Gj given above is widely used for compiling data for international comparisons.

All this means that some caution is appropriate when working with energy data from different sources, as it is not necessarily the case that the same conversion factors have been used in all of the sources. On the other hand, where the data for the different fossil fuels comes in tonnes, converting it all into energy units using averages such as those given above will involve errors in any particular case. Given that the raw numbers are usually large, small differences in the conversion factors can give rise to non-trivial differences in the energy data produced.

To see what can be involved here, go back to the 1,000 Mw coal-fired electricity generating station considered above. We saw that operated at capacity for one hour it would burn 10,800 Gj of coal energy. If we use the UK average figure of 26 Gj per tonne, this is 415 tonnes of coal. If we use the international average of 29 Gj per tonne, this is 372 tonnes of coal.

generating plant. With combustion, all of the chemical energy in the coal is converted to other forms of energy – electrical in the desired output from the plant sent out over the grid, heat energy as waste heat carried away in cooling water or vented to the atmosphere, and chemical energy in the residual matter such as ash. Note that the electrical energy sent out is later transformed to work or heat in homes and factories. Although all of the chemical energy in the coal is conserved, from a human point of view some of the energy transformations are more useful than others. Seen as a source of electrical energy, the plant has a thermal efficiency of (considerably) less than one – the thermal efficiency is the ratio of the electrical energy to the chemical energy content of the burned coal. For modern large generating plants, thermal efficiency is of the order of 35 per cent. What are known as 'combined heat and power' plants use some of the heat that is wasted in a pure electricity generating plant to warm buildings or run production processes. In this way, more of the input chemical energy is converted to energy forms useful to humans, and in that sense the 'efficiency' of the plant is increased.

The first law of thermodynamics is a conservation law. It says that energy is conserved. There is a corresponding conservation law for matter. Matter can neither be created nor destroyed. This law of conservation of matter is sometimes known as the **materials balance principle**. We shall discuss it further at various points in this chapter and in Chapter 4.

Many of those who are concerned about the environment want to encourage people to go in for 'energy conservation'. But, the first law says that there is always 100 per cent energy conservation whatever people do. There is no real contradiction here, just an imprecise use of language on the part of those seeking to promote 'energy conservation'. What they actually want to encourage is people doing the things that they do now but in ways that require less heat and/or less work, and therefore less energy conversion.

There is another widespread use of words in regard to energy that is strictly inaccurate. Often, and especially in economics, people talk about energy 'consumption'. The first law says that energy cannot be consumed in the sense of being used up so that there is less of it than there was previously. What is meant by energy consumption is the conversion of energy from one form to another, and into work and heat. This strictly incorrect usage will often be followed in this book, as it is so widespread, and does not cause any real problems in the contexts where we shall follow it.

The first law of thermodynamics is about energy quantity. The other thermodynamic law that we need to consider, the second law, is about energy quality. Before looking at the second law, we need to look at the way that thermodynamics classifies systems.

2.2.3 Thermodynamic systems classification

Based on a differentiation between flows of energy and flows of matter across the system boundary, thermodynamics distinguishes three types of system:

(1) An open system exchanges matter and energy with its environment;
(2) A closed system exchanges only energy with its environment;
(3) An isolated system exchanges neither matter nor energy with its environment.

If you refer back to Figure 1.1, you will see an example of a thermodynamically open system and an example of a thermodynamically closed system. Thermodynamically, the economy is an open system. It takes from and returns to its environment – which is 'the natural environment' or often just 'the environment' in this book – both matter and energy. The environment is a thermodynamically closed system. It receives from and returns to its environment – the rest of the universe – only energy.

Energy goes from the environment to the economy in many forms – radiation (sunshine), kinetic (flowing water, wind, waves), potential (water reservoirs) and chemical (plant and animal tissue, fossil fuels), for example. Energy goes from the economy to the environment mainly as waste heat and chemical energy in residues. Material flows across the economy–environment boundary take many forms, in both directions. Note that the law of conservation of matter means that the mass of flows across the boundary in each direction will be equal – in terms of total mass, extractions by the economy from the environment equal insertions by the economy into the environment. The composition of the extraction stream is, of course, different from that of the insertion stream. We shall return to this in Chapter 4.

It is not strictly true that 'the environment', i.e. planet earth, is a closed system in a thermodynamic sense. However, it exchanges much energy and very little matter with its environment, and is generally treated as a closed system. As regards matter, meteorites regularly and frequently (thousands each year) enter the environment, and have done so throughout the history of the planet. Meteorites vary in size, but most are very small and burn up in the atmosphere. Of those that have reached the surface, the largest that can still be seen weighs 60 tonnes. Exceptionally large meteorites may have had major impacts on the history of planet earth. A favoured explanation for the extinction of the dinosaurs, and hence the rise of the mammals, is the climatic change that followed the impact, 65 million years ago, of a meteorite 6 miles across. For most of the planet's history there has been no outgoing matter. In the last fifty years human beings have developed the capacity to send matter (as space vehicles of various kinds) out across the environment/universe boundary, but the amount involved has been very small. This is likely to remain the case for some time.

The energy flows crossing the boundary between our environment and its environment are very large, and have been so throughout the history of the planet. The incoming flow is solar radiation, of which approximately $2{,}500 \times 10^3$ Ej reaches the surface of the earth each year. As shown in Box 2.1, an Exajoule or Ej (approximately equal to a Quad), is a very big energy unit. Later we will compare this number for incident solar radiation with some other 'big' energy numbers of economic relevance.

All living organisms are open systems, which exchange energy and matter with their environments. We shall look at plants and animals as open systems in a little detail after considering the second law of thermodynamics.

Strictly, the only isolated system that exists is the entire universe. All other systems that could be delineated must be, at least, closed systems. However, thermodynamicists often use the idea of an isolated system for analytical purposes, and actual systems can be constructed in the laboratory that approximate to isolated systems in the same way as planet earth approximates to a closed system.

We can now state the first law in a slightly different way: the energy content of an isolated system is constant. This is a more precise way of stating the first law. It avoids a possible misunderstanding of it based on the way it was stated above. To say that energy can be neither created nor destroyed is not to say that the energy content of a system cannot change. It is only the energy content of an isolated system that cannot change. Open and closed systems can exchange energy with their environments, and it follows that their energy content can change. Consider again a coal-fired electricity generating plant, and let it with a given stock of coal on its premises be the, open, system. As the coal is burned and electricity (and waste) sent out, so the energy content of this system decreases, reaching a minimum when all the coal is burnt. The energy content of this system's environment increases by the same amount as the system's decreases. Once all of the initial stock of coal is burnt, bringing in more coal will increase the energy content of the system, and decrease that of its environment.

2.2.4 Second law of thermodynamics

It has been said that whereas the first law of thermodynamics is that you cannot get anything for nothing, the second law is that you will always pay over the odds anyway. According to the first law, that is, energy cannot be created, only converted from one form to another. As regards the second law, the point being made is that all conversions involve losses. This seems to contradict the first law, but does not. The loss is not in terms of energy quantity, but in terms of energy quality. All energy conversion processes involve some downgrading of the quality of energy. Quality here refers to the proportion of energy that is available for conversion.

The second law is known as 'the entropy law' because its most basic statement is: the **entropy** of an isolated system cannot decrease. What is entropy? One answer is that it is energy that is not available for conversion. Another is that it is a measure of disorder.

In order to explain how these answers are related, and hence the meaning and implications of the basic statement of the second law, we need to go back to the idea of energy as the potential to do work or supply heat. This implies that heat and work are related. Recall that they are measured in the same units. In the middle of the nineteenth century an engineer called Carnot formulated the relationship that governs the conversion of heat to work – the maximum amount of work that can be obtained from a quantity of heat depends on only the temperature of the heat source relative to its surroundings. The maximum proportion of the heat that can be turned into work is given by

$$E = (T - T_0) \div T$$

where T is the temperature of the heat source and T_0 is the temperature of its surroundings, where temperature is measured in degrees Absolute. This is

$$E = (T \div T) - (T_0 \div T) = 1 - (T_0 \div T)$$

which is 1 only if $T_0 = 0$. But a temperature of zero degrees Absolute (which is minus 273 degrees Centigrade) is impossible, so E must be less than 1. The Carnot efficiency of conversion must be less than one.

Matter in its various forms comprises assemblies of molecules. The molecules that comprise a lump of matter do not completely fill the space that the lump occupies. In the air that surrounds us, the average distance between molecules is about ten times the size of a molecule. In solids, the molecules are more tightly packed together – which is why one can sit on a chair but not on air. In all forms of matter, the molecules are constantly in random motion. The speed of the motion increases with temperature. Faster random motion means less order. Think of heating a suitable solid so that it passes from a solid to a liquid and then to a gaseous state. What is happening is that the amount of random motion is increasing, and goes through critical values so as to produce the transition from one state to another. More random motion is more disorder – a solid is more ordered than a liquid, which is more ordered than a gas. Entropy is a measure of disorder.

To see the connection between the two meanings of entropy, think about a given mass of gas expanding to fill the volume of its container which is increased by the gas pushing back a piston. Suppose that there is no temperature change involved. Looked at from the disorder point of view, the number of molecules is constant, so the distances between them increase, so the amount of random motion increases, so the entropy of the gas increases. Looked at from the energy point of view, the expansion of the gas must be accompanied by an influx of heat to compensate for the energy converted to work to push back the piston, so the entropy of the gas increases. From both points of view, the gas has higher entropy after it has expanded to fill the larger volume.

The second law can, then, be stated in two equivalent ways. It says that the unavailable energy in an isolated system cannot decrease. It says that disorder cannot decrease in an isolated system. Looked at either way, this seems like very bad news. It seems to be saying, and has been interpreted as saying, that things necessarily run down, becoming more disordered, less structured. However, it must be kept in mind that in this version, the entropy law *applies only to isolated systems*. The entropy of a closed or an open system does not necessarily increase, as such systems can import available energy and thereby reduce disorder. The entropy law does have implications for closed and open systems, but they do not include the implication that entropy always increases. One of the scientists who developed thermodynamics, Clausius, said that 'The entropy of the world grows to a maximum'. If by the 'world' he meant planet earth, and if he meant continuously and inevitably, he was wrong. So long as the sun continues to deliver solar radiation, the entropy of the system which is planet earth need not increase. What is true for any system is that in the absence of some input of energy, the system becomes more disorganised.

One implication of the entropy law for non-isolated systems, such as planet earth, is that all conversions of energy from one form to another are, in terms of available energy, less than 100 per cent efficient. It follows from this that all conversions of energy from one form to another are irreversible.

When these implications are put together with those of the first law of thermodynamics, they are extremely important for the study of economics. Were it not for the laws of thermodynamics, material economic production could be expanded indefinitely. That production involves doing work, moving and transforming materials. Doing work requires energy. If energy conversions were 100 per cent efficient and reversible, limited energy availability would not imply a limited capacity to do work. We will come back to various particular aspects of this in the other chapters in Part I. We now use what we have learned about thermodynamics to look at life, in the form of plants and animals.

2.2.5 Plants as open systems

Looked at from the point of view of thermodynamics, a living plant is an open system. It exchanges energy and matter with its environment. A living plant is a highly ordered system, in which disorder is not increasing because the plant is taking energy from its environment to maintain order, i.e. life. At death, the plant ceases to take energy from its environment, and a process of increasing disorder, decay, starts. Eventually, when the decay process is complete, the system that was the plant has become so disordered that it is indistinguishable from its environment.

Plants are a subset of the class of organisms known as 'autotrophs' or 'producers'. The distinguishing characteristic of autotrophs is that they use chlorophyll to make organic matter from inorganic matter, using energy. Most autotrophs are 'phototrophs' in that the energy used is solar radiation, and the process by which organic matter is made is **photosynthesis**. The word 'photosynthesis' means 'building by light'. As well as land and water plants, the class of phototrophs includes algae, plankton and bacteria – plankton, or phytoplankton, are actually plants without roots. We look at things in terms of land plants, but what is said about them also goes for the other phototrophs. Autotrophs that are not phototrophs are not very important in the big picture and will be ignored here (they include some bacteria and some algae).

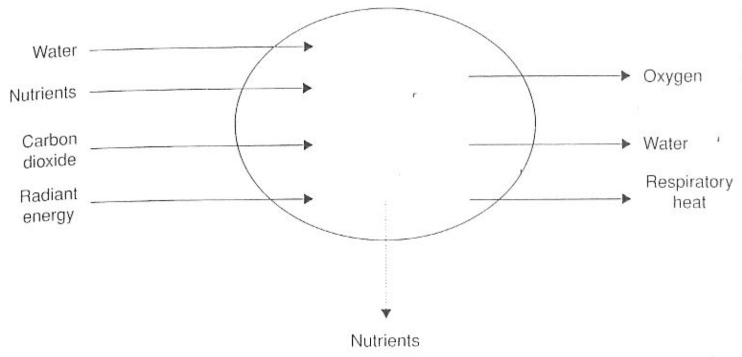


Figure 2.3 Living plant as an open system.

Figure 2.3 shows the important features of a living plant as an open system. The inflows are water, carbon dioxide and radiant energy, i.e. solar radiation or sunlight. The photosynthetic process converts some of the radiant energy to chemical energy stored in the plant tissue, and some of the input energy is returned to the environment as heat. This return of heat is known as respiration, and reflects the energy required to run the photosynthetic process and for the maintenance of the plant system. Oxygen and water also cross the system boundary from the plant to the environment. The operation of the process of photosynthesis requires the presence of certain mineral elements, known as nutrients. These are taken up by the growing plant, from the soil and water, and incorporated into its tissue. When the plant dies and decays these nutrients are returned to the environment. Included in the nutrients necessary for the operation of the process of photosynthesis are: nitrogen, phosphorous, potassium, sulphur, copper, iron, zinc.

The rate at which plants produce plant tissue is known as **primary productivity**, and is usually measured in terms of energy per unit area per unit time – calories per square metre per year, say. Individual plant species, and assemblages of different species of plants, can be compared in terms of their primary productivity. Gross primary productivity is the total amount of solar energy that is fixed by photosynthesis, whereas net primary productivity is that less the amount of energy lost to the environment as respiration, and so the amount that is actually stored in the plant tissue. Net primary productivity is the measure of the energy that is potentially available to the animals that eat the plants in question.

Plant species vary in their primary productivity. For a given species, the primary productivity of a particular population will vary with the environment. The environmental conditions most relevant are: the amount of light (solar radiation), the amount of water, availability of carbon dioxide, temperature, and nutrient availability. Scarcity relative to the plant's requirements in respect of any one of these factors will inhibit plant growth, and reduce primary productivity, and cannot be compensated for by the uptake of more of some other factor for which there is no scarcity. If, for example, the availability of water is inhibiting growth, the inhibition can only be overcome by making more water available – providing more nutrients or carbon dioxide will not solve a water supply problem. In the language of economics (see Chapter 8 on this), in terms of primary production by plants, the various inputs are complements rather than substitutes. Biologists talk in terms of limiting factors – if some input is scarce relative to requirements, it is a limiting factor on growth, no matter how abundant the others may be. At the extreme, the inhibition is so great that the plant cannot grow at all. Arid deserts are the most obvious examples of such extremes, where water is the limiting factor.

The efficiency with which plants convert incident solar energy into tissue varies from 2 per cent to 6 per cent. Much of the solar energy that reaches the surface of the earth does not fall upon plants, or upon places where plants might grow. It was noted above that each year approximately 2,500 Ej of solar energy arrives at the surface of the earth. Photosynthesis annually produces approximately 1.2 Ej of plant tissue, which is 0.05 per cent of the solar energy arriving at the earth's surface.

2.2.6 Animals as open systems

With the substitution of 'animal' for 'plant', the first paragraph of the last section serves as the first for this, as follows.

Looked at from the point of view of thermodynamics, a living animal is an open system. It exchanges energy and matter with its environment. A living animal is a highly ordered system, in which disorder is not increasing because the animal is taking energy from its environment to maintain order, i.e. life. At death, the animal ceases to take energy from its environment, and a process of increasing disorder, decay, starts. Eventually, when the decay process is complete, the system that was the animal has become so disordered that it is indistinguishable from its environment.

Viewed as open systems, there are two main differences between plants and animals. The first is in terms of the source and form of the input energy. For plants it is solar radiation. For animals, input energy is the chemical energy in the food that is taken in. The second is that, whereas plants take in carbon dioxide and give out oxygen, animals take in oxygen and give out carbon dioxide. An animal as an open system is depicted in Figure 2.4. Like plants, animals need inputs of water and nutrients. The latter are obtained, along with energy, from food input. Energy/heat goes from the animal system to its environment in two ways. The animal's faeces contain stored chemical energy in the undigested food. There is also, as with plants, respiratory heat. The animal's faeces also contain nutrients. These are also returned to the environment when the animal dies, and decomposition releases those that were stored in the animal tissue. Whereas plants are known as producers, animals are known as consumers, or 'heterotrophs'. Animals are classified according to the source of the food that they consume. Animals that consume plants as food are known as 'herbivores', or 'primary consumers', while animals that consume animals are 'carnivores', or 'secondary consumers'.

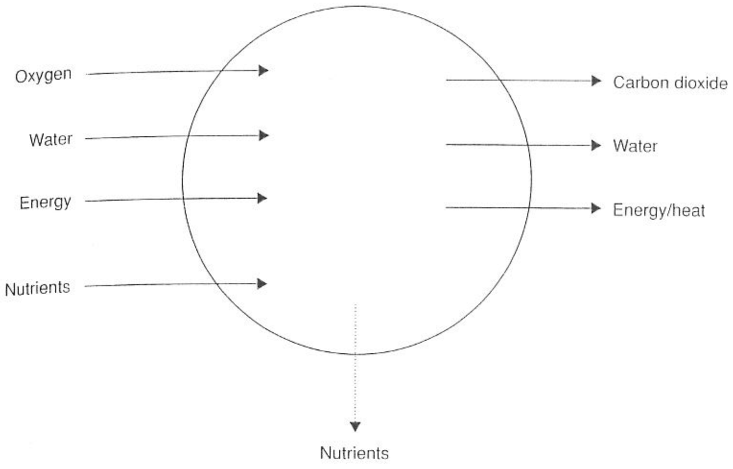


Figure 2.4 Living animal a an open system.

We saw that plants convert only a small proportion of the incident solar energy into chemical energy stored in plant tissue. Herbivores likewise convert only a small proportion of the chemical energy of plant tissue into animal tissue. Consumption efficiency is

$$CE = \frac{I}{P}$$

where P is the net primary productivity of the plant system and I is the amount of productivity ingested by the animal system. Assimilation efficiency is

$$AE = \frac{A}{I}$$

where A is the productivity actually assimilated by the herbivore system, the difference between A and I being accounted for by energy expelled with the faeces. Production efficiency is

$$PE = \frac{T}{A}$$

where T is the net productivity of the herbivore system, i.e. the chemical energy incorporated into animal tissue. The difference between T and A is due to heat respiration, which in the case of an animal system additionally arises when energy is converted to the work involved in the animal moving around. The overall efficiency, E, for the conversion of energy from storage in plant to storage in animal tissue is

Box 2.2 Animal food-gathering strategies

Of the food energy taken in, eaten or consumed, by an animal, some is assimilated and some leaves the animal in faeces and urine. Of the assimilated food energy, some is stored as new tissue and some is used in respiration and eventually dissipated as heat. Respiration is the work done to maintain the animal's structure. That work includes the gathering of the food that is the source of the energy taken in. Clearly, if the system that is an animal is to be viable, it must be the case that the energy taken in from food per unit time is greater than the energy expended in acquiring the food. In fact, the ratio:

$$\frac{C}{E} = \frac{\text{Energy consumed}}{\text{Energy expended}}$$

must be greater than some number greater than 1, because some of the energy consumed is not assimilated. If, say, 50 per cent of food eaten were assimilated, then a viable design for an animal system would require that C/E were greater than 2.

The table below gives data for six animals on the energy expended in feeding and the C/E ratio. Given that these are animals that exist, the fact that the minimum value for C/E is greater than one is not the point about these data – matters could not be otherwise. Rather, the point is that the range for the ratio is much narrower than the range for the rates of energy expenditure, E.

In the next chapter we will look at human food-provision methods in this way.

Animal	E *calories per minute*	$\frac{C}{E}$
Hummingbird	32.9	7–70
3 Hummingbird species	16.1–21.5	3.8–22.2
Finch	15.6	12.8
Bumblebee	0.32–0.46	4.4–20.2
Damselfly larva*	5×10^{-6}–5×10^{-5}	1.1–3.6
Black bass	2.2–3.0	3.8–10.3

* Note the use here, in $5 \times 10^{-6} - 5 \times 10^{-5}$, of an extension of the notation introduced in Box 2.1. 10^{-6} is $1/10^6 = 1/(10 \times 10 \times 10 \times 10 \times 10 \times 10) = 1 \div 1{,}000{,}000 = 0.000001$, so that $5 \times 10^{-6} = 0.000005$. Similarly, $5 \times 10^{-5} = 0.00005$.

Source: Lawton (1973).

then given by

$$E = CE \times AE \times PE$$

According to the plant species and the herbivore species, it is estimated that E varies between 5 per cent and 20 per cent.

Carnivores could be looked at in the same way, with P in the first, consumption efficiency, ratio being the productivity of the animal that gets eaten rather than of the plant that gets eaten.

The production efficiency of an animal, herbivore or carnivore, varies with the level of activity. Respiration is at its minimum when the animal is at rest, and increases with the level of activity. The continuance of life for the animal requires that, at least, as much energy is acquired from food as is required for respiration on account of movement and the maintenance of basic metabolic functions. More than that is required for non-adult animals so that growth in mass can occur. Reproduction also requires additional energy inputs. Different animal species have different strategies for food acquisition, which differ in their implications for energy input requirements. However, any viable strategy must have the characteristic that, on average and over a suitably defined period of time, energy acquired as food must be at least as great as energy expended in acquisition. Otherwise, the animal will die.

2.3 ECOSYSTEMS

An **ecosystem** is a system comprising living organisms, known as biota, and their non-living, or abiotic, environment, and all of the interactions between all of the biotic and abiotic components of the system. The delineation of the boundary of an ecosystem is a matter of judgement, and depends to some extent on the purpose at hand. Very detailed studies can be conducted of ecosystems of small spatial extent, such as, for example, a pond or a small woodland area. At the other extreme, the entire biosphere can be treated as a single ecosystem, and studied at a less detailed level. Both extremes have their uses in trying to understand how 'the environment' works. For some purposes, the world is divided into large areas of similar climate and plant life, which large ecosystems are referred to as 'biomes'.

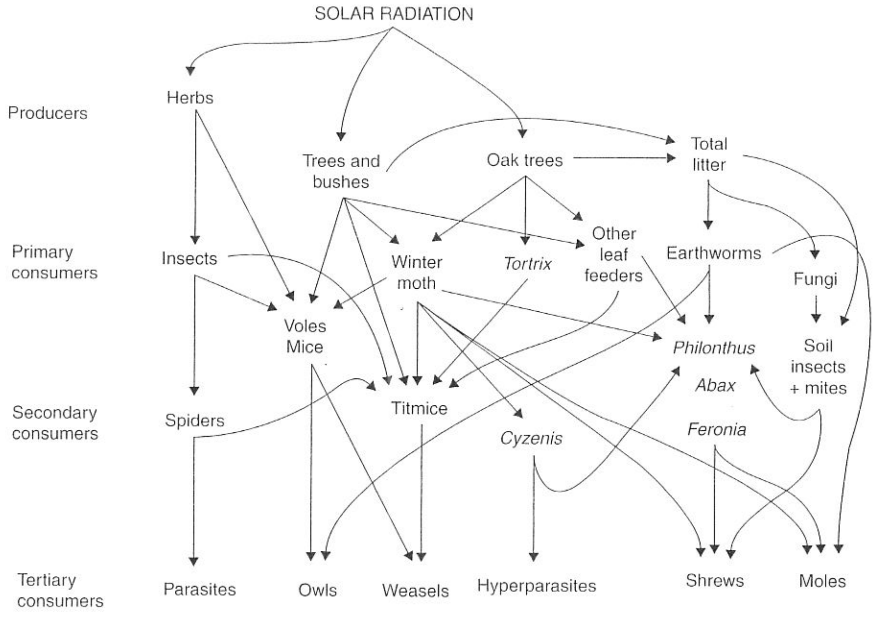
At whatever scale they are defined, ecosystems have generic structural features in common. It is these common features that we shall be looking at here. We look first at the way energy and matter move through ecosystems.

2.3.1 Energy and nutrient flows

We have seen that both plants and animals are thermodynamically open systems. Plants and animals are involved in a feeding chain. For a very simple example, in an aquatic context, we can think of plankton (plant) which gets eaten by a crustacean (herbivore) which gets eaten by a herring (carnivore) which gets eaten by a human (top carnivore, in this chain). In practice, feeding chains do not have such simple structures. Usually, for example, an organism at one level is an input to more than one organism at the next higher level. Figure 2.5 shows, still in simplified form, a more realistic set of relationships between producers and consumers in a foodweb. The particular foodweb here is a woodland ecosystem in the UK. However, the basic solar energy → producers → consumers structure applies to all ecosystems, at whatever level, from the local to the global, they are delineated.

In looking at plants and animals as open systems, we noted that the photosynthetic conversion of solar radiation to stored energy in plant tissue is (considerably) less than 100 per cent efficient, as is the conversion of plant tissue to animal tissue. There are, likewise, losses as we go from primary to secondary, and from secondary to tertiary, consumers. As a result, when we look at the structure of an ecosystem in terms of the chemical energy stored per unit time at the various levels of the foodweb, we get a **trophic pyramid** of the general form shown in Figure 2.6. 'Trophic' means 'of, having to do with, nutrition' – recall 'autotroph' for plant as producer, and 'heterotroph' for animal as consumer. In Figure 2.6 there are four 'trophic levels', corresponding to the four classes of organism whose nutrition is at issue.

The units of measurement at each level in Figure 2.6 are energy, stored in organic tissue, per unit area per unit time, say calories per square metre per year. Because, as discussed above, the consumption, assimilation and production efficiencies are all less than 100 per cent, for any ecosystem the amount of energy fixed in herbivore tissue over a given period of time must be less than the amount fixed in plant tissue. The same is true for the transition from any trophic level to the one above it, and hence the pyramid shape when we stack the energy stored by level as in



Source: based on Figure 9.1 in Jackson and Jackson (2000).

Figure 2.5 A foodweb for a woodland ecosystem.

Figure 2.6. Particular ecosystems vary from one another in the detail of their trophic pyramids, but all have the basic pyramid shape showing the necessary structure of the flow of energy through an ecosystem. The relative sizes of the rectangles in Figure 2.6 understate the narrowing of the pyramid that would be found in any actual ecosystem. The upper limit to the size of the rectangle at one level as a proportion of the size of the rectangle at the level below is estimated to be about 20 per cent. Assuming that upper limit holds at every stage, the size of the tertiary consumers' rectangle would be less than 1 per cent of the size of the rectangle for plants – the energy stored in the tissue of top carnivores would be less than a hundredth of that stored in the system's plants.

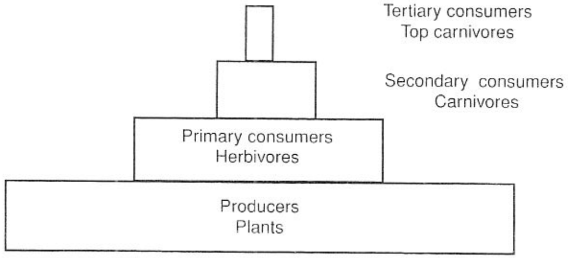


Figure 2.6 A trophic pyramid.

Table 2.1 Net primary productivities for selected biomes

Biome	*Area 10⁶ sq km*	*Net primary productivity per unit area tonnes per sq km per year*		*Net primary productivity World total 10⁹ tonnes per year*
		Range	***Mean***	
Tropical rainforest	17.0 (3.3)	1000–3500	2200	37.4 (22.0)
Temperate deciduous forest	7.0 (1.4)	600–2500	1200	8.4 (4.9)
Boreal forest	12.0 (2.4)	400–2000	800	9.6 (5.6)
Temperate grassland	9.0 (1.8)	200–1500	600	5.4 (3.2)
Tundra and alpine	8.0 (1.6)	10–400	140	1.1 (0.7)
Desert and semi-desert	18.0 (3.5)	10–250	90	1.6 (0.9)
Cultivated land	14.0 (2.7)	100–3500	650	9.1 (5.4)
Swamp and marsh	2.0 (0.4)	800–3500	2000	4.0 (2.4)
Total terrestrial	**149 (29.2)**		**773**	**115 (67.6)**
Open ocean	332.0 (65.1)	2–400	125	41.5 (24.4)
Continental shelf	26.6 (5.2)	200–600	360	9.6 (5.6)
Algal beds and reefs	0.6 (0.1)	500–4000	2500	1.6 (0.9)
Estuaries	1.4 (0.3)	2000–3500	1500	2.1 (1.2)
Total marine	**361 (70.8)**		**152**	**55.0 (32.4)**
Total	**510**		**333**	**170**

Source: based on Jackson and Jackson (2000), Table 9.4.

Abundance at different trophic levels can also be looked at in terms of numbers of individuals per unit area, or in terms of biomass per unit area. Biomass is simply the weight of living material. In terms of either numbers or biomass, the same pyramid shape is generally obtained as when looking at trophic levels in energy terms.

The base for the trophic pyramid of an ecosystem is net primary productivity, the amount of energy fixed as plant tissue per unit time. It follows that the relative abundance of life of all kinds in different ecosystems is mainly determined by relative performance in terms of net primary productivity. Table 2.1 gives information about this for some selected **biomes**. The data of Table 2.1 directly refer to plant life, but the relativities that they reveal will also apply, broadly, to the abundance of animal life given its dependence on plant life. The biomes listed in Table 2.1 do not account for the whole of the earth's surface: figures for the missing biomes can be obtained from the source for Table 2.1, which also gives data on biomass. The second column gives the area of the biome in millions of square kilometres, and, in parenthesis, as a percentage of the total surface area of the earth. The productivity data in Table 2.1 are in terms of equivalent mass rather than energy units. The third column gives productivity per unit area, i.e. per square kilometre, in terms of the range across the biome and the mean for the biome as a whole. The fourth column gives the biome's total productivity, in billions of tonnes, as the product of its area and its mean productivity, and, in parenthesis, its percentage contribution to global productivity.

The main points to be noted from Table 2.1 are as follows. The marine biomes account for 70 per cent of the earth's surface, but only 32 per cent of its total net primary productivity. This is mainly due to the fact that the open oceans account for over 90 per cent of the marine surface area, but have relatively low per unit productivity. The mean productivity per unit area for the open oceans is about the same as that for the tundra/alpine terrestrial biome, and not much greater than that for desert and semi-desert. Note, however, that algal beds and reefs and estuaries have per unit area productivities similar to those for the most productive terrestrial biomes. In fact, the upper end of the range for algal beds and reefs, which refers to tropical rain coral reefs, is higher than the upper end of the range for any terrestrial biome.

As regards the terrestrial biomes, per unit area productivity generally declines with increasing distance from the equator, reflecting declining receipts of solar radiation. It is estimated that more than 70 per cent of total terrestrial net productivity occurs between latitudes 30 °N and 30 °S. Tropical rainforest has the highest mean productivity. Note that the upper limit of the range for cultivated land is the same as the upper limit for tropical rainforest, but that, because cultivated land spans a wide range of latitude, its mean is well below the mean for tropical forest. Although tropical rain forest accounts for only 3.3 per cent of the surface, it accounts for 22 per cent of total global productivity. Cultivated land accounts for 2.7 per cent of the surface and 5.4 per cent of productivity.

Ecosystems can, then, be analysed in terms of the way that energy flows through them. As we saw in the discussion above of plants and animals as open systems, energy is not the only thing that crosses the boundaries of plants and animals. Minerals, or nutrients, are necessary for plants and animals to process energy conversions, and cross the boundaries of systems which are individual organisms. However, if we take the ecosystem as the system to be analysed, there is an important difference between energy flows and the flows of nutrients. Essentially, there is a one-way flow of useful energy through an ecosystem, whereas nutrients cycle around an ecosystem. As we shall see when we consider planetary processes below, this is strictly true only if we are looking at an ecosystem which is the whole biosphere – for other delineations of ecosystems, minerals do cross their boundaries. However, the statement is approximately true for most of the boundaries that would define interesting ecosystems, and it does make an important point in a simple way, so we shall proceed for now as if it were true without qualification.

Figure 2.7 shows energy and nutrient movement in an ecosystem. Energy as solar radiation enters the system, and is converted to organic matter by producers, and thus passed to herbivores and carnivores, as discussed above, and to decomposers, to be considered shortly. The heat flows shown in Figure 2.7 go from the system to its environment: they are the products of respiration processes, and are not useful energy. The flow of useful energy is unidirectional: energy is not recycled. Decomposers are the organisms that operate the decomposition processes in an ecosystem, which processes are the means by which nutrients are recycled round the system. **Decomposition** is the breakdown of dead organic matter, which releases the inorganic nutrients that it contains, making it available to be taken up from the soil by living plants. There are two classes of decomposers. Fungi and bacteria secrete digestive enzymes which break down the complex molecules of dead organic matter into simpler ones that they can utilise. These organisms are known as 'saprophytes'. The second class of decomposer organisms are animals, known as 'detritivores', that eat dead organic material. The material that these animals excrete is finer-textured than the material that they eat, which makes it easier for the fungi and bacteria to work on. Examples of detritivores are centipedes, earthworms, nematodes and woodlice.

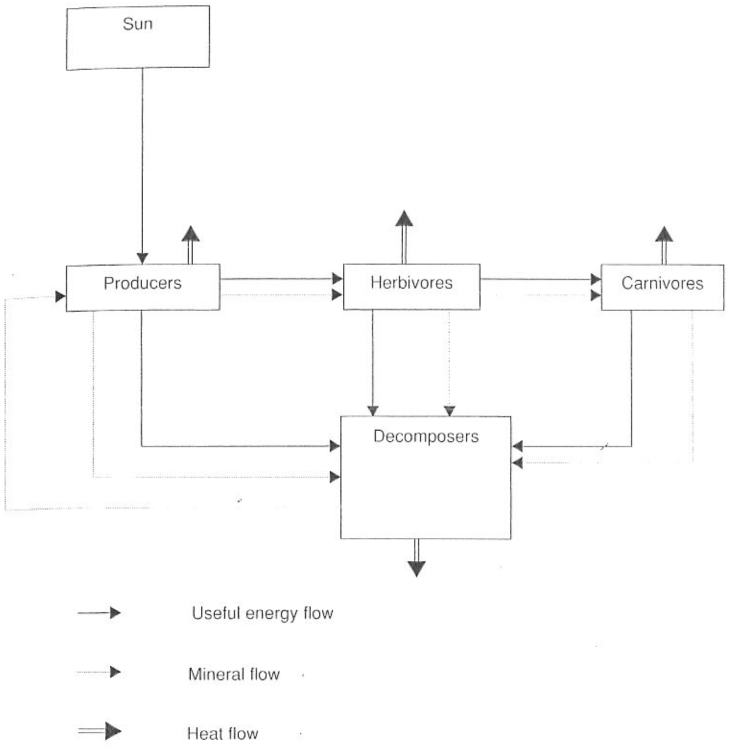


Figure 2.7 Energy and nutrient movement in ecosystems.

The complete decomposition of dead organic material is rare. Soils contain small quantities of organic material known as 'humus'. The processes involved in decomposition are inhibited by low oxygen availability, low temperature and high acidity. Where such conditions apply, the result is the accumulation of dead organic material. This is the starting point for processes, operating over millions of years, that led to the existence of the fossil fuels. Since modern industrial economies are characterised by extensive use of fossil fuels, we now discuss their origins.

2.3.1.1 *Origins of the fossil fuels*

The **fossil fuels** are coal, oil and natural gas. All are organic in origin, and all are the product of solar radiation that reached the surface of the earth over long periods of time a long time ago.

Coal was once vegetation, and particularly peat. Peat is an organic deposit which accumulates when the rate of production of plant tissue by photosynthesis exceeds the rate at which it is decomposed. Such a situation is usually associated with wetland areas, and peat is now found mainly in the higher latitudes of the northern hemisphere. Peat builds up over thousands of years. Coal was formed when ancient peat deposits were buried beneath sediment layers and thus compressed. There are several classes of coal according to the amount of compression, and hence the remaining moisture content. In order of increasing compression/decreasing moisture the classes are: lignite, brown coal, bituminous coal, anthracite. Bituminous and anthracite coal was laid down in the Carboniferous period (360 to 280 million years ago), and lignite and brown coal during the Cretaceous period (140 to 60 million years ago).

Oil was once animal tissue. It is thought to have originated with the accumulation on the sea bottom of the bodies of very small sea creatures. Under some conditions, decomposition was incomplete and the organic molecules were converted into hydrocarbon molecules, some of which accumulated as oil in porous rock formations. Oil being lighter than the water that saturates porous rock, it migrates towards the surface of the earth. Liquid oil deposits arise where this process leads to accumulation in reservoirs of porous and permeable rock capped by impermeable rock, so further movement towards the earth's surface is impossible. Oil shale is shale containing preserved organic matter that has undergone some conversion to hydrocarbons, but which has not migrated to a reservoir for liquid oil. Tar sand is sandstone in which some of the pore spaces are filled with heavy hydrocarbons.

Natural gas consists mainly of methane, which is released as a by-product during the formation of oil, and natural gas deposits are usually found in association with oil deposits. Methane is also produced during the process by which coal is formed from peat, and is sometimes found in association with coal deposits. Natural gas is so called because for many years the gas that was burned in homes and factories was produced from coal. In the UK natural gas, from fields under the North Sea, displaced 'towngas' produced in 'gasworks' in the 1960s.

The foregoing account of the origins of the fossil fuels is the standard account, accepted by the overwhelming majority of geologists. For oil and natural gas, an alternative has been proposed. Whereas in the standard account, oil and natural gas have biological origins, in the alternative their origins are inorganic. According to this alternative theory, the majority of the earth's oil and gas is the result of the entrapment on this planet of some of the primordial hydrocarbons dispersed through the debris that became the solar system. If true, this could imply that the amounts of oil and natural gas existing on planet earth are much larger than is currently estimated. However, it would also be true that most of the 'additional' oil and natural gas would be extremely difficult to exploit, given foreseeable technology. In this book we shall accept the standard story, and largely ignore the alternative account and its implications.

Given the standard account of the origins of the fossil fuels, we can think of solar energy as being like a flow of money coming in as income, and then the fossil fuels are like a savings account into which deposits were made (by means of photosynthesis), a long time ago, from that income. The fossil fuels are saved-up past receipts of solar energy, where the saving was made possible because some solar energy was converted to plant tissue by the process of photosynthesis. Now, we saw, when looking at plants as thermodynamically open systems, that of the solar radiation that reaches the earth's surface only a small proportion is converted to plant tissue by photosynthesis. And, we have just seen that in the case of coal, the saving of that tissue occurred only in some circumstances and only over some periods of geological time – all of the coal was laid down in two geological epochs with a joint duration of 160 million years, which is less than 5 per cent of the geological history of the earth. Similar considerations apply to oil, and to natural gas on the standard account of their origins.

Given those origins, the amount of energy that is stored in the savings account that is the fossil fuels must be finite, and is really quite small. Each year approximately 2,500 × 10^3 Ej of solar radiation arrives at the surface of the earth. Of this, photosynthesis is estimated as fixing 1.2 × 10^3 Ej as primary productivity. A central estimate of the size of the stock of fossil fuels prior to the start of their depletion by humans is 315 × 10^3 Ej. That is equivalent to just 260 years' worth of global primary productivity, and much less than one year's worth of the solar radiation arriving at the surface of the earth.

Given that the fossil fuels are incompletely decomposed organic matter, it follows from our discussion of plants and animals as open systems that the fossil fuels must contain carbon. Coal, oil and gas vary in their carbon content. And, for any one of the fossil fuels the carbon content varies across deposits, so that the following figures are averages. Natural gas has the lowest carbon content at 14.6 kg per Gj (kilograms per Gigajoule). Oil is next lowest at 18.6 kg per Gj. Coal is the most carbon-intensive of the fossil fuels at 24.1 kg of carbon per Gj of energy content. Taking natural gas as the base with value one, the relative carbon intensities are 1.27 for oil and 1.65 for coal. On average, and approximately, deriving a given amount of work or heat from coal results in the release into the atmosphere of 65 per cent more carbon than would be the case if gas were used.

2.3.2 Population dynamics

The biotic components of an ecosystem are populations of plants and animals. A **population** is a group of individuals belonging to the same species which live in a given area at a given time. A **species** is a set of individuals who are capable of interbreeding. Organisms which are physiologically incapable of interbreeding (or which produce sterile offspring when they do interbreed) are members of different species. A population is, then, a reproductively isolated subset of a species. Its reproductive isolation is due to location, as opposed to physiology. Different ecosystems may contain organisms from the same species, but different ecosystems contain different populations. One way of looking at ecosystem behaviour over time is in terms of the behaviours over time – the dynamics – of the populations that make up the ecosystem.

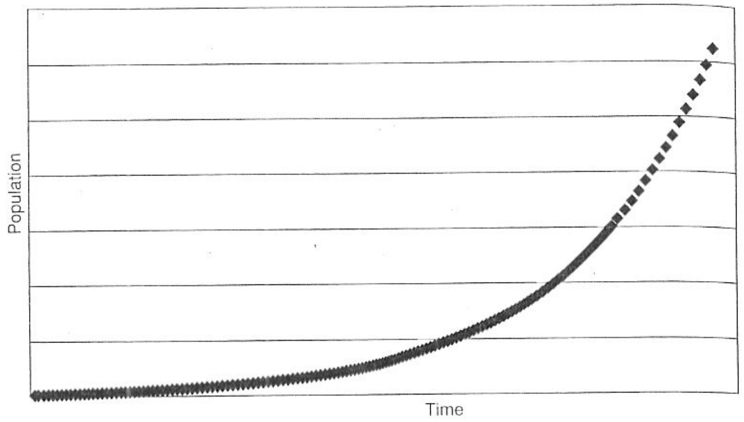


Figure 2.8 Exponential growth.

The actual dynamics of actual populations in actual ecosystems are determined by many factors, and disentangling the various processes at work can be very difficult. Ecologists try to understand the basic processes involved by constructing models, and by conducting controlled experiments in laboratories.

2.3.2.1 *Exponential growth*

A very simple model is **exponential growth**, where the proportional increase is the same in each time period, which means that the absolute increase keeps on getting bigger over time. Figure 2.8 shows a population growing exponentially. It is drawn using numbers from a simulation generated in an Excel^{TM} spreadsheet, as follows. The initial population size, 1, was entered in the cell A1. Then, the entry for cell A2 was generated using the formula palette as

$$A2 = A1^*1.05$$

which is the entry in A1 times 1.05, so that A2 is 5 per cent bigger than A1. Then, cell A2 was copied and pasted into cells A3 to A100. Because this is using relative rather than absolute cell references, the effect is that the entry in A3 is 1.05 times that in A2, the entry in A4 is 1.05 times that in A3, and so on and so on. Reading down the A cells gives exponential growth at the rate of 0.05 or 5 per cent.

Using some symbols, we can state the general exponential model in simple algebra. Let N_0 represent the population size at the beginning of the initial period, and the size at the beginning of the next period is

$$N_1 = (1 + r) \times N_0$$

and for the one after that it is

$$N_2 = (1 + r) \times N_1 = (1 + r) \times [(1 + r) \times N_0] = (1 + r)^2 \times N_0$$

Table 2.2 Annual growth rates and doubling times

Growth rate %	*Doubling time years*
1	69.7
2	35.0
3	23.5
4	17.7
5	14.2
6	11.9
8	9.0
10	7.3

and so on and so on, so that generally

$$N_t = (1 + r)^t \times N_0$$

gives the population size at the start of period t, N_t, for exponential growth at the rate r from a starting level of N_0. For two adjacent periods this is

$$N_t = N_{t-1} + (r \times N_{t-1})$$

where N_{t-1} is size at the start of one period and N_t is size at the start of the next period. The absolute amount of growth is

$$N_t - N_{t-1} = r \times N_{t-1}$$

and proportionate growth is

$$\frac{N_t - N_{t-1}}{N_{t-1}} = r$$

In this model, N, the population size, is the variable that we are interested in the behaviour of, and r, the growth rate, is a parameter. A parameter is something that is constant in one simulation. Different simulations of the same model arise with different values for the parameter – Exercise 2.1 at the end of the chapter asks you to plot graphs of exponential growth for different growth rates. A model with just one parameter is a very simple model. In the next sub-section we will look at a population growth model with two parameters.

A useful way of expressing the implications of exponential growth is in terms of the 'doubling time'. This is the number of periods that it takes for whatever it is that is growing at a constant proportional rate to double in size. Table 2.2 shows some annual percentage growth rates and the approximate corresponding doubling times in years. The same numbers would apply for different periods of time used consistently. Thus, for a daily growth rate of 3 per cent the doubling time would be 23.5 days. Money left to earn interest in a savings account grows exponentially – according to Table 2.2 if you could get 5 per cent per year which you never took out of the account, your money would double within 15 years. We will look at this kind of compounding in Chapter 8. The Appendix at the end of the chapter here explains how the entries in Table 2.2 are obtained.

2.3.2.2 *Density-dependent growth*

It is, of course, impossible for a population to experience exponential growth indefinitely. The population's environment sets an upper limit to the size that it can attain because there is an upper limit to available solar radiation, for a plant population, or to available food, for an animal population. The maximum population size that the environment can support is called its 'carrying capacity'. Figure 2.9 shows a simple model of population dynamics where there is an upper limit to population size, which limit is known as the environment's **carrying capacity**. Initially the population grows exponentially, but as it increases in size so the growth rate

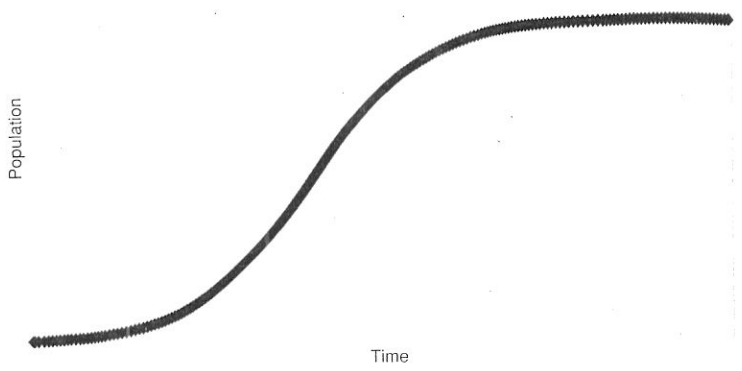


Figure 2.9 Density-dependent growth.

falls and goes towards zero as the population size approaches carrying capacity. This type of population dynamics is known as density-dependent growth because the growth rate depends on population size, which, given a particular environment, is equivalent to population density. Figure 2.9 actually shows an Excel^{TM}-generated plot of **logistic growth**, which is a particular kind of density-dependent growth. When we explained how Excel^{TM} was used to generate numbers for exponential growth at 5 per cent, instead of

$$A2 = A1^{*}1.05$$

we could have said

$$A2 = A1 + (0.05^{*}A1)$$

and got exactly the same results. In logistic growth, instead of a fixed number like 0.05, there is a number which varies with the difference between the size of the population and the carrying capacity of its environment. For the results graphed in Figure 2.9, the carrying capacity was 100, and the formula for A2 was

$$A2 = A1 + \left(0.1 \times \left(\frac{100 - A1}{100}\right)\right) \times A1$$

which in Excel^{TM} notation is

$$A1 = A1 + (0.1^{*}((100 - A1)/100)^{*}A1)$$

There is a growth rate 0.1 which is modified by a factor which is the proportion by which A1 falls short of the carrying capacity 100. For the simulation graphed in Figure 2.9, the entry in A1 is 1.

As for exponential growth, this formula is copied into cells A3 to A100, so that the entry in cell A3, for example, will be

$$A3 = A2 + \left(0.1 \times \left(\frac{100 - A2}{100}\right)\right) \times A2$$

or, in Excel^{TM} notation:

$$A3 = A2 + (0.1^{*}((100 - A2)/100)^{*}A2)$$

And so on, and so on. Going down the cells, 100 minus the entry for the cell above gets smaller, so the factor applied to the entry for the cell above to get this one gets smaller – the growth rate declines, and eventually it tends to zero.

Adding K for carrying capacity to the symbols introduced above for exponential growth, logistic growth can generally be represented as

$$N_t - N_{t-1} = r \times \left(\frac{K - N_{t-1}}{K}\right) \times N_{t-1}$$

where r and K are the parameters of the model. Comparing this with the general version of the exponential growth model

$$N_t - N_{t-1} = r \times N_{t-1}$$

you can see that a constant proportional growth rate r has been replaced by the proportional growth rate

$$r \times \left(\frac{K - N_{t-1}}{K}\right)$$

which varies with N_{t-1}, while r and K are fixed. In the logistic growth model r is referred to as the intrinsic growth rate, as it is the rate at which the population would grow (exponentially) if there were no environmental limits. When N_t is small, $(K - N_{t-1}) \div K$ is close to one, and the actual growth rate is close to the intrinsic growth rate r. As N_{t-1} increases towards the carrying capacity K, so $(K - N_{t-1}) \div K$ gets smaller and the actual growth rate decreases. For N_{t-1} equal to K, the numerator in $(K - N_{t-1}) \div K$ is zero so $(K - N_{t-1}) \div K$ is zero and the growth rate is zero.

2.3.2.3 Species types

Ecologists have found it useful to classify types of organism, species, in terms of the two parameters r, the intrinsic growth rate, and K, the carrying capacity, of the logistic growth model. They distinguish between r species, or strategists, and K species, or strategists. The idea is that species vary along a continuum with r species at one extreme and K species at the other. Most species exhibit some combination of the characteristics of a pure r and a pure K species.

The main characteristic of r strategists is a high value for r – given favourable conditions, they reproduce very rapidly. Another characteristic is that the population growth rate is not very sensitive to the population density, does not slow down very much as the carrying capacity is approached. As a result, there is a tendency for the population size to overshoot the carrying capacity, leading to a subsequent collapse. Individual members of r species tend to have relatively short lives, and to be small in size. For animals the length of time that offspring receive parental care is short. Examples of r species are annual plants and rabbits.

The main characteristic of K strategists is a low value for r, the intrinsic growth rate. Also, the actual population growth rate is more sensitive to the population density and K strategists tend to exist at population levels close to the carrying capacity of their environment. Individual members of K species tend to have relatively long lives, to be of large size, and, for animals, to provide extended parental care. Examples of K species are trees, elephants and humans.

2.3.2.4 Equilibrium and stability

The models presented above for exponential and logistic growth are examples of difference equations. A difference equation is an equation that gives the path taken by some variable over successive periods of time, as in Figures 2.8 and 2.9. There are many different types of difference equation.

One interesting question about the time path generated by a difference equation is whether there is an **equilibrium**, i.e. whether there is some level for the variable such that if it is attained the variable will, in the absence of shocks, remain at that level. With the exponential growth model, there is an equilibrium and it is zero. If you put 0 in the first cell when you simulate that model, all the subsequent cell values will be 0. Zero is also an equilibrium for the logistic growth model, as you can verify in the same way. However, this model has another equilibrium, which is K, the carrying capacity. You can verify this by putting the value for K in cell A1 when you do a simulation – see Exercise 2.2.

If an equilibrium exists, an interesting question about the time path generated by a difference equation is whether it has a tendency to return to an equilibrium if moved away from it by some external shock. This is the question of **stability**. Put another way, the question is: for an initial value which is not an equilibrium, will the variable move towards an equilibrium value? If there is more than one equilibrium, there is also the question of which the variable will move towards from a given initial value. These are questions about stability.

In the simulations for Figures 2.8 and 2.9 the initial values for the variable were not equilibrium values. In the exponential case, the equilibrium is 0 and the initial value is 1, and the variable grows away from 0. It is clear that this is what will happen for any starting level other than 0. Exponential growth is unstable – start it away from equilibrium and it will get further away from it.

In the logistic case of Figure 2.9, the population size moved over time towards carrying capacity, not towards 0. This is what will happen for any starting level other than zero. Logistic growth is stable with respect to the carrying-capacity equilibrium. Put another way, in the logistic growth model the carrying capacity is a stable equilibrium. Zero is an unstable equilibrium – start somewhere other than it, and the variable will move away from it.

Figure 2.10 shows simulations of some other types of population dynamics, where there are oscillations. In Figure 2.10(a) the size of the oscillations, known as the amplitude, is decreasing over time, and you can see that the variable is converging on an equilibrium, which is, then, a stable equilibrium. In Figure 2.10(c) the amplitude is increasing over time – there is an equilibrium, but it is unstable. Note that in this case the amplitude of the oscillations will eventually get so big that on the downswing the population size goes to zero and the population goes extinct. In Figure 2.10(b) the oscillations are of constant amplitude, and this pattern of behaviour is known as a 'limit cycle'.

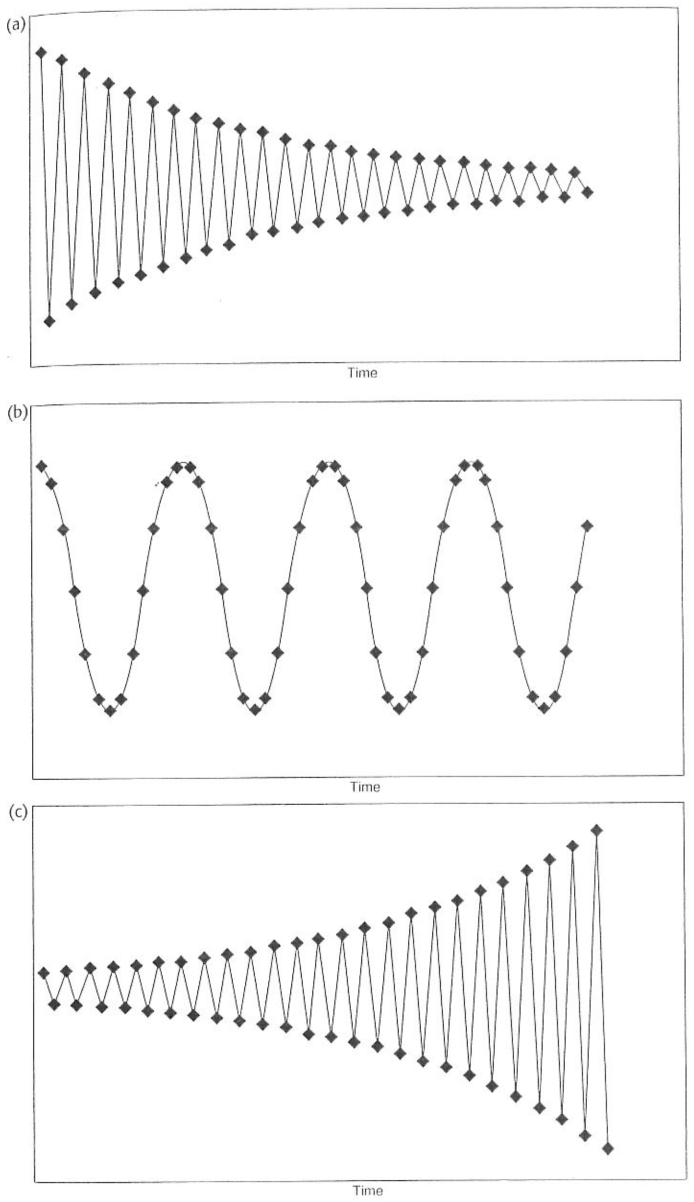


Figure 2.10 Some types of population dynamics.

2.3.2.5 Population interactions

The difference equation models for the simulations shown in Figures 2.8 to 2.10 refer to the behaviour of just one population. In ecosystems there are many populations, of different species, that interact with one another and with the abiotic environment. Ecologists have constructed more complicated models in which two, or more, populations interact in various ways, and have also run experiments where populations interact with one another.

One standard example is the model of inter-species competition for limited food. Analysis of this model shows four possible sorts of outcome, depending on the numerical values taken by the parameters which describe the intrinsic growth rates for each of two populations from different species, what the carrying capacity of the environment would be for each population if it alone existed in it, and how the size of one population affects the growth of its competitor. The possible outcomes are:

(1) There is no equilibrium in which both populations exist. Either population A completely out-competes population B, and B goes locally extinct, or vice versa.

(2) There is an equilibrium in which both populations exist, and it is a stable equilibrium. Starting from population sizes that are not the equilibrium levels, both population sizes will converge on their equilibrium levels.

(3) There is an equilibrium in which both populations exist, but it is an unstable equilibrium. Any disturbance to an equilibrium state will set in motion dynamic behaviour that involves the extinction of one of the populations.

Most, but not all, laboratory experiments with simple organisms in simple environments result in outcomes where only one of the competitive populations survives. The coexistence in field conditions of apparently competitive populations is generally taken to suggest either that the species are not fully competitive, that they are not both completely dependent on the same food supply, or that the supply of food is not actually a limiting factor.

Another standard example is the predator–prey model. In this model there is a prey population – rabbits say – that is the food for the predator population – foxes say. The solution in this model is that the sizes of both populations oscillate, with the turning points for the predator lagging behind those for the prey. The oscillations may be either damped or of constant amplitude. This type of behaviour can be produced in laboratory experiments, and is observed in the field.

2.3.3 System dynamics

An ecosystem is an assembly of many interacting populations, together with their abiotic environment. Even in a small localised ecosystem the population interactions will be many and complex, as illustrated by the foodweb shown in Figure 2.5. As well as studying the behaviour of the individual populations that comprise an ecosystem, ecologists study the dynamics of entire ecosystems. In doing that they focus on processes and functions of the system as a whole, as well as the characteristics of the component populations.

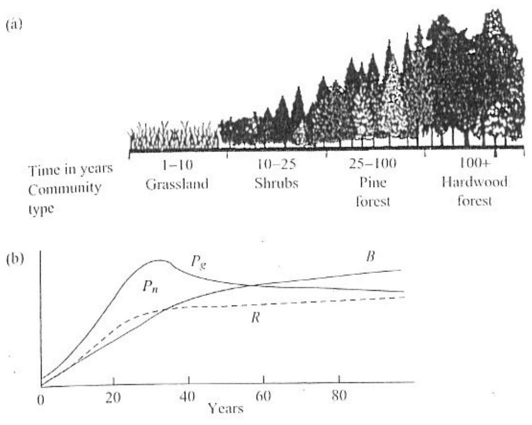


Figure 2.11 Forest succession.

Key: P_g - gross primary productivity
P_n - net primary productivity
R - respiration
B - biomass

Source: adapted from Folke (1999).

2.3.3.1 Succession

An important idea in much ecological thinking is **succession**. This refers to the way in which the species composition of an ecosystem occupying a particular area changes over time, converging on what is known as a 'climax state'. The process starts with an area with very little vegetation. This may be the result of natural events, such as fire or storm, or of human activity, such as clear-cut logging. Figure 2.11 illustrates the case where the process starts from the situation after clear-cut logging. The area is first colonised by annual plants, especially grasses, which are *r*-strategy species. This involves the expansion of remnant populations from the original state of the area, and/or invasion of the area by populations from outside. The species involved at this stage are also known as 'pioneer' or 'fugitive' species.

The pioneer species change the opportunities that the area offers to other species of plants and animals, and further colonisation takes place. The vegetation comes to be dominated by perennial plants, which in turn alter the opportunities available to various kinds of plants and animals, and eventually the area reaches a climax state in which it is dominated by *K*-strategy species such as, in this case, trees. Figure 2.11(a) shows an idealised forest succession in terms of the plants dominant at each stage: the animal species present, being dependent on plants for food, would also change as succession progressed. Figure 2.11(b) shows how primary productivity and biomass vary through the stages of this succession. The pattern shown there – biomass increasing to a stationary level at the climax state – is thought to be typical of successional processes in general. The climax state can be seen as the system equivalent of the equilibrium level for an individual population.

2.3.3.2 *Species functions*

From the perspective of the behaviour of the ecosystem as a whole, what is most interesting about the different species represented in it is the roles that they play in the functioning of the system as a whole. The functioning of the system has certain essential requirements – solar energy capture and decomposition are obvious examples. It appears that in any given ecosystem there is a subset of the total suite of species present that carries out the essential roles. Ecologists call such species 'keystone species'. A simple constructed example is as follows. Imagine an ecosystem in which solar energy capture is largely by a plant species which drops its seeds to the ground beneath. The seeds are dispersed by one particular species of bird which eats the seeds and then deposits them around the ecosystem in its faeces. In the absence of this species of bird, the plant species could not reproduce, and the ecosystem would suffer a major loss of its primary productivity, with major implications for the survival of other animal species. The bird species is, for this ecosystem, a **keystone species**.

It is tempting to infer from the existence of keystone species that the other, non-keystone, species in an ecosystem are functionally redundant. This inference would be incorrect. We will return to this question shortly, after introducing the concept of ecosystem resilience to which it relates. It can be noted here that while the identification of keystone species is in some cases fairly straightforward, in many circumstances it is very difficult and there is a great deal of ignorance about which are, and which are not, keystone species in most ecosystems. For example, given the necessity of nutrient recycling it is clear that the role played by the class of detritivores is essential in all ecosystems. However, there are in any ecosystem many detritivore species and it is not the case that those which are keystone in any particular ecosystem have been definitively identified.

Different species can perform the same role in different ecosystems. To continue with the seed dispersal by birds example: in ecosystem *A* plant species *X* is responsible for *z* per cent of solar capture and its seeds are dispersed by bird species *I*: in ecosystem *B* plant species *Y* is responsible for *z* percent of solar capture and its seeds are dispersed by bird species *II*. Different species playing the same role in different systems are known as 'ecological equivalents'.

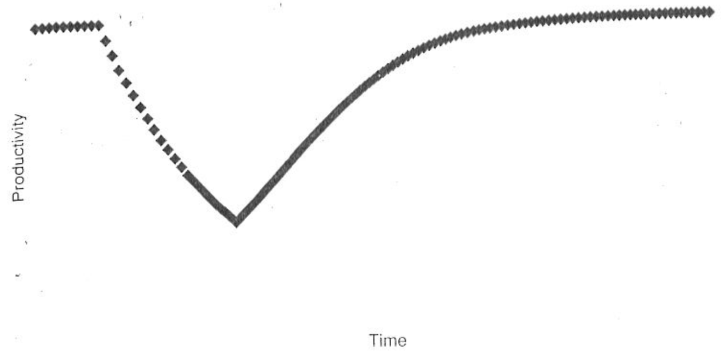
Australia, which became a separate land mass 60 million years ago, provides striking evidence on ecological equivalents. Australia has never had any placental mammals. In the rest of the world, the placental mammals (mostly) out-competed and displaced other types of mammals, notably the marsupials. The Australian marsupials did not face that competition. As a result, there evolved in Australia a whole range of marsupial species – herbivores, carnivores and top carnivores – which play ecological roles that are played elsewhere by placental species of mammals. A different way of putting this is to say that in Australia marsupial species fill ecological niches that are elsewhere filled by placental species.

2.3.3.3 *Resilience*

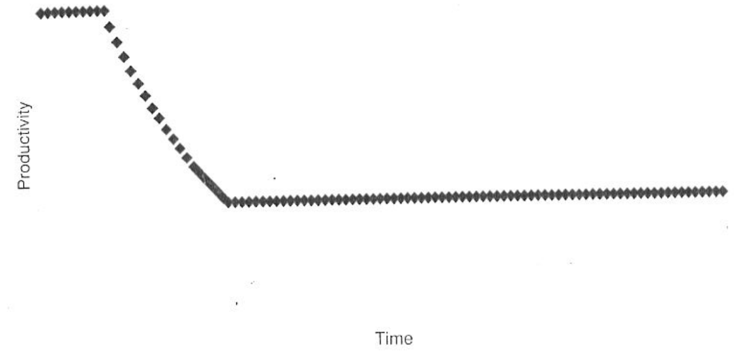
An ecosystem is said to be resilient if it tends to maintain its functional integrity when subjected to some disturbance. A resilient system is one that, when subjected to some shock, continues to exist and to function in the same essential ways. Note that it is not being said that resilience requires continued functioning in exactly the same way, nor is it being said that it requires that functions continue to be carried out by the same species. **Resilience** is, rather, consistent with some of the populations in the ecosystem going to zero.

Figure 2.12
Resilience.

(a) A resilient system



(b) A non-resilient system



How do we tell whether or not functional integrity is maintained? We need some kind of indicator. One candidate, which in fact ecologists do use quite widely, is primary productivity. Using this indicator, Figure 2.12 shows the difference between a resilient and a non-resilient system. In the former case, productivity recovers following some disturbance that reduces it. In the latter, it does not. To repeat the point made above, it is not necessarily the case that in the panel a situation

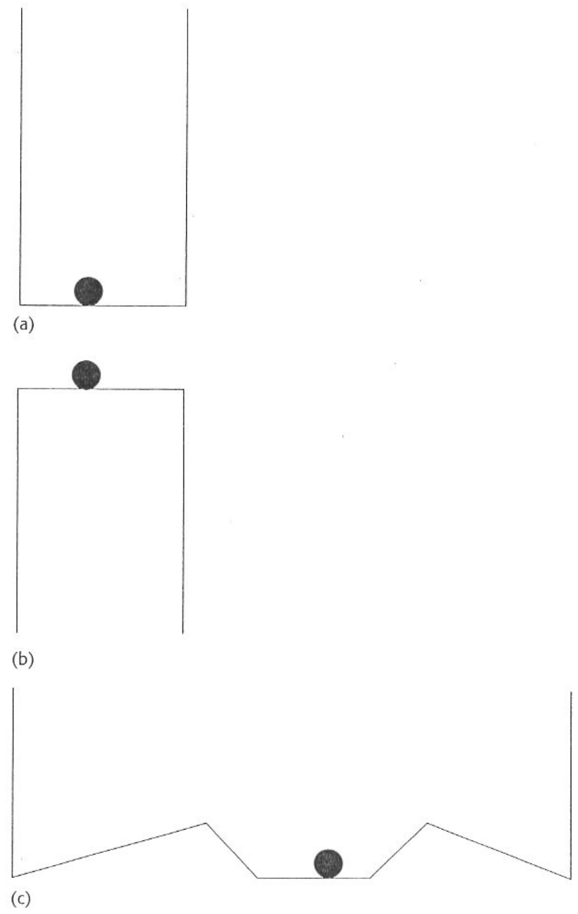


Figure 2.13
Another look
at resilience.

all of the populations in the system recover their former size. Some may 'go out of business'. The point is that the system as a whole 'stays in business' and that, given time, for the system as a whole 'business as normal', as reflected in primary productivity, is resumed.

Figure 2.13 provides another way of thinking about the idea of resilience, and introduces some further development of the concept. Again, Figure 2.13(a) refers to a resilient system, panel b to a system which is not resilient. Figure 2.13(a) shows a glass standing on a table in an upright position, with a ball in the bottom of it. Pick the glass up and shake it. The ball will roll around the bottom of the glass, within limits set by the sides of the glass. Put the glass back on the table and the ball will settle pretty much where it was before the disturbance. This corresponds to resilience – following a disturbance, the system returns to its original state. Figure 2.13(b) shows the glass upside down on the table, with the ball sitting on the outside of the bottom of the glass. Now the slightest disturbance will see the ball roll off the bottom of the glass, and down onto the table – there is no way back once this has happened. This corresponds to an extreme case of the form of non-resilience shown in Figure 2.12(b) – the system falls apart in the face of a disturbance.

Now look at Figure 2.13(c), where the bottom of the glass has been modified. Now, instead of being flat, it has a central circular depression around which the raised circumference has a slight downward slope out to the side of the glass. For a gentle shake, the ball will remain in the central depression, and the situation is as in Figure 2.13(a). A stronger shake will cause the ball to jump from the central depression onto the surrounding area. This behaviour is intermediate between that of Figure 2.13(a) and (b). It does not involve no change of state whatever the shock as in Figure 2.13(a), nor does it involve collapse for any disturbance as in Figure 2.13(b). Does this sort of situation get described as resilient or non-resilient? It gets described as non-resilient because the system does not revert to its original state, but remains in a different state after the disturbance. Going back to ecosystems, the idea that the only alternative to resilience is total collapse, as in Figure 2.13(b), is not correct. A non-resilient ecosystem will not necessarily collapse in the face of disturbance, but it will not regain its original state.

Resilience is a property of the system, rather than of its component parts. An obvious question is whether we can identify characteristics of ecosystems that promote resilience. If some systems are more resilient than others, why is that? This turns out to be, except in fairly general terms, a hard question to answer. One reason for this is that an ecosystem may be resilient with respect to one type of disturbance, but not to another – it may cope with fire but not with man-made pollution, for example. Similarly, a system may be resilient with respect to disturbance up to a certain level, but not beyond that – it may have a threshold level of disturbance beyond which it losses its resilience. In the present state of knowledge, the resilience or otherwise of an ecosystem is something that we can be sure about only after the occurrence of a disturbance, and then only that the system turned out to be resilient, or not, in the face of that particular disturbance.

Generally, many ecologists now take the view that species diversity promotes resilience. At one time it was generally agreed that more complex ecosystems were more stable in the sense that the sizes of the component populations fluctuated less, and that this promoted the ability of the system to persist in the face of disturbance. In that context, complexity was measured by the number of species in the system and the number of feeding links between them. The basic idea was that with many links, the removal of one link would do less damage. It is now understood that things are not that straightforward. It has been shown by mathematical modelling that increased complexity in this sense does not necessarily increase stability. It is now understood that resilience of the system is not directly and simply related to the stability of the component populations, and that ecosystems where there is low population stability can exhibit resilience.

A disturbance will threaten the functional integrity of an ecosystem to the extent that it threatens the existence of the keystone species. However, while the functions required for resilience are given and fixed, the identity of the species that carry out those functions need not be. The fact that the function of, say, seed dispersal is now carried out by species *x* does not, necessarily, mean that species *x* is the only species present in the system that could do the necessary seed dispersal. What are currently redundant species may be reservoirs of replacement keystone species, should disturbance severely reduce the ability of the current keystone species to carry out necessary functions. Also, currently redundant species may not be able themselves to exercise any necessary functions, but may be reservoirs of genetic material from which new species that can do that may evolve.

Ecologists admit to much ignorance regarding the nature and determination of resilience. But, among ecologists the majority view is that species diversity promotes resilience. That is one of the reasons why they, and other biological scientists, argue for the conservation of biological diversity. We return to this at various points in the rest of the book, and especially in Chapter 14.

2.4 NUTRIENT CYCLES

As already noted, as well as energy life requires the availability of certain chemical elements that get called **nutrients**. A nutrient is a chemical element taken up by an organism to maintain its functions. The 'macro nutrients', which collectively account for 99 per cent of human body mass, are: oxygen, hydrogen, carbon, nitrogen, calcium, phosphorous, sulphur, potassium, magnesium. Nutrients present in organisms in smaller quantities are called 'micro nutrients' and include sodium, iron, copper, zinc and iodine.

As seen when looking at ecosystems, nutrients cycle through the environment. Each nutrient has its own cycle which operates at the planetary level. The cycles involve both biotic and abiotic processes, and are for that reason sometimes referred to as 'biogeochemical'. Each cycle involves processes that connect it to other cycles. For reasons of space, we will look at just one cycle – the carbon cycle – here. Directions to descriptions of the other important cycles will be found in the Further Reading section at the end of the chapter. The carbon cycle illustrates the essentials of a nutrient cycle, and some familiarity with it is required for an understanding of the climate change problem.

2.4.1 The carbon cycle

There are basically two forms of carbon. The first is organic carbon, which is that found in living, and dead but not decomposed, organisms. Otherwise carbon is inorganic. Also, there are really two carbon cycles, a slow one and a fast one. Reference to 'the carbon cycle' is almost always a reference to the fast cycle, the slow one being so slow that for many purposes it can be ignored.

2.4.1.1 The slow cycle

The slow cycle is geological. More than 99 per cent of all terrestrial carbon is contained in the lithosphere. Most of this is inorganic carbon stored in sedimentary rock such as limestone: the organic carbon in the lithosphere is contained in fossil fuel deposits – recall that fossil fuels are incompletely decomposed organic matter. In geological time there is cycling of the inorganic carbon between the earth's crust, the oceans and the atmosphere. The crustal store gets added to by a process that first has precipitation taking carbon out of the atmosphere and (eventually) into the ocean, where it sinks to the ocean bottom, or is taken up by marine organisms the decomposed remains of which eventually also sink to the bottom. In both of these ways, sediment accumulates, which is converted to rocks such as limestone. Movement in the opposite direction occurs when, due to tectonic movements, such sedimentary rocks are subjected to heat and pressure releasing carbon dioxide into the atmosphere in volcanic eruptions.

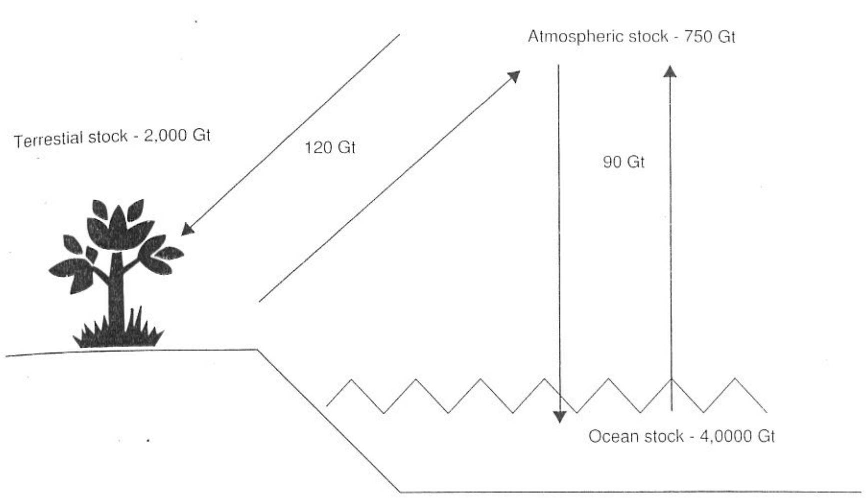


Figure 2.14 The carbon cycle.

As regards the organic carbon in the lithosphere, until very recently – 200 years ago – any significant exchanges between the fossil fuel deposits and other carbon stores also operated over geological time. This has changed since man began the large-scale extraction and combustion of fossil fuels, which releases the carbon that they contain into the atmosphere.

2.4.1.2 The fast cycle

The operation of the fast cycle in the absence of anthropogenic influences (i.e. caused by humans) is shown in summary in Figure 2.14. The flows and stocks of carbon are measured in Gigatonnes (Gt). There are three stocks, or reservoirs, of carbon – the ocean stock, the atmospheric stock and the terrestrial stock. The size of the annual exchanges between the stocks, often referred to as 'fluxes', are shown by the numbers between the relevant arrows. As shown in Figure 2.14, the flows in each direction are equal. Strictly this is not the case, but it is true to a close approximation, and in the absence of anthropogenic influences the relative sizes of the stocks would change very slowly over time. It is the atmospheric stock that determines the concentration of carbon dioxide in the atmosphere, which concentration influences the global climate. It is for this reason that, although it is the smallest of the three, it is the stock upon which interest is currently mainly focused. Note that total annual exchanges between this stock and the other two are approximately one quarter of the stock size, whereas for the other two stocks the flux/stock ratio is much lower. This indicates that the size of the atmospheric stock would be relatively sensitive to changes in the fluxes.

The terrestrial stock is the carbon contained in the tissue of land biota, in soil litter and in peat. Of the total of 2,000 Gt of carbon shown in Figure 2.14, about one quarter is accounted for by the biota. The exchanges between this stock and the atmospheric stock are effected by the processes of photosynthesis and respiration described above when looking at plants and animals as open systems.

The oceanic stock is by far the largest of the three – it is approximately fifty times the size of the atmospheric stock. Exchanges between this stock and the atmospheric stock are effected by chemical processes which establish an equilibrium between the concentration of carbon dioxide in the surface layers of the oceans and the concentration in the air above that surface. Some of the carbon thus absorbed into the oceans is taken up into the tissue of plankton. When these die some of the carbon they contain is carried down to the ocean bottom where it is effectively removed from the fast carbon cycle. This is one of the reasons why the fast cycle exchanges between the atmosphere and the oceans are not exactly equal.

2.4.1.3 Anthropogenic influences

Estimates of the quantity of carbon contained in fossil fuel deposits are in the range 5,000–10,000 Gt. As noted above, until recently this store did not figure in the fast carbon cycle. Now, due to human activity, it does, and its influence on the fast, global carbon cycle is significant.

Humanity's use of the fossil fuels essentially began with the start of the industrial revolution in the late eighteenth century. At that time the amount of carbon in the atmosphere meant that the concentration of CO_2 was 280 ppmv, where ppmv stands for 'parts per million by volume', so that this is saying that CO_2 comprised 0.028 per cent of the global atmosphere. Such has been the growth of fossil fuel use since the industrial revolution – to be discussed further in Chapter 3 – that the atmospheric CO_2 concentration is now (2004) approximately 370 ppmv. The concentration has increased by more than 30 per cent in a little over 200 years. According to the best estimates, the present concentration level definitely has not been exceeded during the last 420,000 years, and it is likely that it has not been exceeded in the last 20 million years. The rate of increase in atmospheric carbon over the last century is unprecedented in the past 20,000 years, and is likely to be high by the standards of a much longer period of time.

In the decade 1990–1999, anthropogenic releases of CO_2 into the atmosphere expressed in terms of carbon averaged (central estimates throughout this paragraph and the next) 6.3 Gt per year. As well as fossil fuel combustion, the production of cement from limestone contributed to these emissions, but fossil fuels account for over 96 per cent of total emissions. Note that the interaction of the atmospheric store with the fossil fuel store is unlike the interaction of the atmospheric store with the others shown in Figure 2.15 in that it is one way – there is a flow from the fossil fuel store to the atmosphere but not in the reverse direction.

In the decade 1990–1999, the amount of CO_2 in the atmosphere expressed in terms of carbon increased at an average of 3.2 Gt per year. The atmospheric stock increased by less than anthropogenic emissions. This is because some of the CO_2 released into the atmosphere by human activity was removed from it by the exchanges with the terrestrial and ocean stocks shown in Figure 2.15. Of 6.3 Gt per annum the oceans accounted for a net uptake of 1.7 Gt per annum, and the land for a net uptake of 1.4 Gt per annum. The net uptake by the land store consists of two elements. Land use changes by humans – mainly deforestation – reduced the rate at which the terrestrial store took CO_2 from the atmosphere. On the other hand, what is known as the 'residual terrestrial sink' increased the rate at which CO_2 was removed from the atmosphere. The term 'residual terrestrial sink' refers to a residual amount of CO_2 removal from the atmosphere which it can be established was not effected by exchanges with the ocean, but which cannot yet be definitively assigned to identified exchanges with the land store. One possibly important element in the operation of the 'residual terrestrial sink' is CO_2 fertilisation, whereby the rate at which plants take up CO_2 increases with the CO_2 concentration. Accounting for CO_2 is not a precise science. There is much, at the level of detail, that is not known.

While this is true, the big picture is clear. Human activity is affecting the global carbon cycle in a readily detectable way. Basically, carbon is being moved from the fossil fuel deposit store to the atmosphere. The increasing atmospheric concentration of CO_2 is affecting the global climate system. Discussion of this will come up at a number of places in the rest of the book. It is the subject of Chapter 14.

2.5 EVOLUTION

Evolution is the process of change over time. All kinds of systems undergo evolution. In this section we will be concerned mainly with evolution in the natural environment, and especially with biological evolution. We will look at evolution in human systems, especially economic systems, in subsequent chapters.

2.5.1 Biological evolution

An individual organism can be looked at in terms of its genotypes and its phenotypes. The **genotypes** are the organism's genetic inheritance, which at birth define the boundaries for potential development of the organism. The maximum height, for example, that a human organism can attain during his or her life is set by the genes that it is endowed with by the parents. The **phenotypes** are the organism's observable characteristics. To continue the example, the actual height of the human is one of his or her phenotypes. Phenotypes are determined by the genotypes and the organism's environment. An individual human may be phenotypically short notwithstanding genotypical tallness, due to inadequate nourishment.

The mechanism that drives biological evolution is **natural selection**, which works as follows. The individuals that comprise a population – members of the same species coexisting as a reproductive unit – differ genotypically and phenotypically. The reproductive capacity of a population generally exceeds the carrying capacity of the environment, and there is competition among individuals for the inputs needed for survival. Those individuals that are most fit will be the ones that survive. Fitness is directly a matter of phenotype, but generally has an underlying genetic basis. Individuals that survive to reproductive age can pass their genes to offspring, individuals that do not cannot. Hence, the struggle for survival, and for reproduction, over time shapes the genetic make-up of the population, as well as its phenotypical structure. By means of natural selection, a population becomes better adapted to its environment. A mutation is a random error in the process by which an organism inherits its genes from its parents. Mutations are occurring all the time in all populations. In most cases they result in an organism which is genotypically less fit than its parents, but sometimes they result in a better-fitted organism. If the latter outcome is an individual that can reproduce, then the process of natural selection towards a population better fitted to its environment is advanced by that mutation.

Natural selection is the generally accepted explanation for the proliferation of species, which process is known as 'speciation'. The basic idea is that a population splits spatially into reproductively isolated groups, i.e. becomes two populations. To the extent that the environments of the two populations differ, they will be subject to different adaptive selection. Also, the effects of any 'successful' mutation are confined to the population in which it occurs. The two populations diverge both in terms of genotypes and phenotypes. If the divergence goes to the extent that the two populations would be incapable of interbreeding then **speciation** has occurred.

As described here, the process of speciation, which is what is generally understood by 'biological evolution', works through adaptation driving natural selection operating on genotypes. An individual organism's fitness is determined by its phenotypes, but what it passes on are genes, and adaptation works because of the link between genotypes and phenotypes. The generally accepted position is that there is an effect from genotype to phenotype, but not from phenotype to genotype. The response of an organism to the environment that it is exposed to during its life, cannot, that is, affect the genes that it passes to its offspring. Phenotypical adaptation by an individual organism confers no benefits on its offspring. While this is true for biological evolution, it is not necessarily true of evolution in other contexts. In the evolution of human culture, for example, parents can pass to their offspring information that they have acquired. We shall come back to this in Chapter 3.

A final point needs to be made in this very brief account of biological evolution. Natural selection is very often referred to as involving the 'survival of the fittest'. It would be more accurate if the phrase that has become so widespread had been 'survival of the fitter'. The point is that what gets selected is that which, from among what is available, is relatively the best fitted to the relevant environment. It is not the case that what gets selected is the best possible

adaptation to the relevant environment, Natural selection can only operate on what is there.

2.5.2 Coevolution

Biological evolution is a complicated business. It is about organisms adapting to their environment. But the environment which is being adapted to is itself constantly changing. Indeed, adaptation itself drives environmental change, because for any population its 'environment' includes lots of other populations from other species, all of which are themselves subject to the pressures of natural selection. What is actually going on all of the time is 'coevolution'.

For a given population of a species, only a small part of the totality of the environment is directly relevant. The part of the environment with which interaction takes place is the population's niche. For example, a given species of bee has a niche which comprises a particular range of plants from among the many that occupy the space where they operate, while wolves have a niche which comprises many animal species, but still a lot less than the totality of the animal species in their territory. **Coevolution** refers to the fact that the niche for any one population is affected by evolutionary change involving other populations. For the bees new plants may appear, or existing ones vanish. A niche may be enlarged or reduced. Previously successful adaptation may be rendered obsolete, and a sufficient amount of niche reduction will lead to extinction. On the other hand, speciation will itself tend to create new niches, through, for example, new predation possibilities, thus promoting yet more speciation. Biological evolution has the potential to sustain itself through ongoing coevolution.

In fact, in the history of planet earth it appears that coevolution has involved non-living as well as living systems. The abiotic environment has affected the biotic, and the biotic has affected the abiotic. The nutrient cycles as they now exist are the result of coevolutionary processes involving non-living and living systems that took place many hundreds of millions of years ago. The atmosphere of the earth for the first few hundred million years after its formation contained no oxygen, but did contain a lot of carbon dioxide, as well as nitrogen, methane and ammonia. Given the presence in the atmosphere of a lot of carbon dioxide, it is supposed that the global temperature was then much higher than it is now.

The earth is thought to be about 4,500 million years old. For the first 1,000 million years there was no life on earth at all. It appeared, how is not really known, as a very primitive bacterial form, about 3,500 million years ago. The appearance of a form of life capable of utilising solar radiation came about 500 million years after that. These organisms took in carbon dioxide from the atmosphere and released oxygen into it. For the then existing organisms, oxygen was toxic. However, its rate of accumulation in the atmosphere was very slow. About 2,000 million years ago the first oxygen-tolerant photosynthesising organisms appeared. Had this not happened, the slow build-up of oxygen in the atmosphere would have extinguished life on earth. It did happen, and the composition of the atmosphere became more oxygen-rich, and eventually sufficiently so as to support animal life. With plants taking in carbon dioxide and releasing oxygen and animals taking in oxygen and releasing carbon dioxide, the carbon and oxygen cycles were linked. Simple forms of life, it appears, played a crucial role in creating the conditions for the existence of complex forms of life. The early simple forms are now almost completely extinct.

SUMMARY

This chapter has provided an introduction to some of the key ideas about how the natural environment works, focusing on those most relevant to ecological economics. The idea of a system is very important in the environmental sciences, and, as we shall see in subsequent chapters, in the social sciences as well. Thermodynamics is the study of energy conversions in systems. Living organisms are systems that perform energy conversions. An ecosystem is a collection of interacting populations of organisms, together with their abiotic environment. The biosphere is the global ecosystem considered in its entirety. The biosphere has evolved throughout 3,500 million years of the history of planet earth, and will continue to evolve. A major motivation for the study of ecological economics is the fact that the future evolution of the biosphere will be strongly influenced by human economic activity.

KEYWORDS

Biome (p. 39): a spatially large ecosystem defined by climatic and vegetative conditions.
Carrying capacity (p. 45): the maximum population size that a given environment can support.
Coevolution (p. 61): the process whereby the environment in which one population is evolving is itself changing due to the evolution of its constituent populations.
Decomposition (p. 40): the breakdown of dead organic matter into inorganic matter.
Ecosystem (p. 37): a system of living organisms and their non-living environment.
Energy (p. 26): the potential to supply heat or do work.
Entropy (p. 31): energy that is not available for conversion, a measure of disorder.
Equilibrium (p. 48): a population level that if attained will persist in the absence of disturbance.
Evolution (p. 59): the process of change over time.
Exponential growth (p. 44): growth at a constant proportional rate.
Fossil fuels (p. 42): energy sources of organic origin.
Genotypes (p. 59): an organism's genetic inheritance.
Keystone species (p. 52): species that carry out functions essential for ecosystem functioning.
Logistic growth (p. 46): a particular form of density-dependent growth with the growth rate declining as the population grows.
Materials balance principle (p. 28): matter can be neither created nor destroyed.
Natural selection (p. 60): genetic adaptation to the environment driven by relative reproductive success.

Nutrients (p. 56): chemical elements taken up by organisms to maintain their functioning.
Phenotypes (p. 59): an organism's observable characteristics.
Photosynthesis (p. 32): the process by which plants use solar radiation to convert inorganic to organic matter.
Population (p. 43): a group of individuals belonging to the same species living in a given area at a given time.
Primary productivity (p. 33): the rate at which plants create organic matter, usually measured as energy per unit area per unit time.
Resilience (p. 53): the maintenance by an ecosystem of its functional integrity when subjected to disturbance.
Speciation (p. 60): the emergence of new species. [illegible]
Species (p. 43): a set of individuals who are capable of interbreeding.
Stability (p. 48): the tendency of a population size to return to its equilibrium following a disturbance.
Succession (p. 51): the way in which the species composition of an ecosystem occupying a particular area changes over time, converging on a climax state.
System (p. 22): a set of interacting components.
Thermodynamics (p. 26): the study of energy transformations.
Trophic pyramid (p. 37): the decline in biomass moving from plants to herbivores to carnivores.

APPENDIX: DOUBLING TIMES WITH EXPONENTIAL GROWTH

From

$$N_t = (1 + r)^t \times N_0$$

the doubling time is the t value that is the solution to

$$2 = (1 + r)^t \times 1$$

Dividing both sides by 1 gives

$$2 = (1 + r)^t$$

and taking natural logarithms on both sides this is

$$\ln 2 = t \times \ln(1 + r)$$

The natural logarithm of 2 is 0.6931, and so the doubling time is

$$t = \frac{0.6931}{\ln(1 + r)}$$

If you solve this for $r = 0.01$, etc. you will get the results shown in Table 2.2. The reason for working with natural logarithms rather than logarithms to the base 10 is that it so happens that this answer lines up with an approximation that is easy to remember – divide 70 by the growth rate expressed as a percentage. So, for r as 5 per cent, for example, the approximation is, in whole numbers, 14.

FURTHER READING

Jackson and Jackson (2000) and Park* (2001) are two standard environmental science texts that deal with all of the topics dealt with in this chapter at greater length. Both are at an introductory level: Jackson and Jackson assumes some prior knowledge of chemistry. Rogers and Feiss* (1998) is an introductory environmental science text that approaches the material from the perspective of human interests. Bowler (1992) is a history of the development of the environmental sciences.

Thermodynamics is difficult for the non-specialist, for whom many accounts of the first and second laws have been written. Chapman (1975), Ramage* (1983) and Slesser (1978) are well-written books, intended for the non-specialist general reader, on energy matters, which contain reasonably straightforward expositions of thermodynamics and its implications. Although 'old' they are not 'dated' except in so far as they come at 'the energy problem' from the perspective of limited supplies of fossil fuels rather than that of the climatic implications of the use of fossil fuels. They all consider the technological limits that arise: see also Ayres (1978), Chapman and Roberts (1983), Hall *et al.* (1986), and Ruth (1999). Faber *et al.* (1996), chs. 6 and 7 especially, deals with the first and second laws in relation to ecological economics. Georgescu-Roegen (1971) introduced thermodynamics to economists, and is one of the seminal works in the development of ecological economics: it is not an easy read. The energetic data in the chapter is, where not otherwise cited, based on data from Ramage* (1983) and Georgescu-Roegen (1976).

The alternative theory of the origins of the fossil fuels is set out in Gold (1999). If true it has important implications for our understanding of the origins of life on earth, and for assessment of the prospects of life on other planets. Cole (1996) is a non-technical account of the controversy surrounding Gold's ideas, which is also very interesting for what it says about the actual practice of science. Despite its obvious practical, as well as scientific, importance, and the expenditure of lots of money on 'definitive' tests of Gold's hypothesis, the controversy remains unresolved.

Krebs* (2001) is a successful ecology text that is comprehensive but assumes no prior knowledge of the subject. Folke (1999) is a brief overview of ecological principles as they relate to ecological economics, and provides useful references to the literature. Krebs deals with the basics of the mathematical modelling of population dynamics. Gilbert and Troitzsch (1999) provides an overview of simulation modelling and available software. Hannon and Ruth (1994) is an introduction to the use of the Stella^{RM} software package for the simulation of dynamic models. The (Excel^{TM}) simulations for Figure 2.10 here can be found on the companion website. Our typology of the sorts of behaviour that difference equations can produce omits chaos. For some ranges of the parameter values, simple non-linear difference equations produce outcomes with oscillations where the amplitude is neither constant, constantly increasing or constantly decreasing, but varies over time. Also, the pattern of variation changes with very small changes in the initial conditions. There is now a large literature on chaos and its implications – Hannon and Ruth (1994) provide simple models that produce chaos. Closely related to the work on chaos is work on complex systems – roughly speaking a complex system is one whose behaviour is not predictable from the behaviour of its component parts. Kauffman (1995) covers much of the ground from the perspective of a biologist actively involved in the work: Gleick (1988) and Waldrop (1994) are journalistic, but informative, accounts.

As set out here, the idea of resilience as a property of an ecosystem was introduced in Holling (1973). It is further developed in Holling (1986). The paper by Ludwig *et al.* (1997) is a clear, but technical, exposition of the basic mathematics of Holling resilience and how it relates to another concept of resilience that appears in the ecology literature.

Nutrient cycles are covered in Park* (2001), Jackson and Jackson (2000) and Krebs* (2001); see also Ayres (1999), which, with Jackson and Jackson, gives more details on the chemistry. The information given on the carbon cycle in the chapter is taken from Houghton (1997) and Houghton *et al.* (2001). Further references relating to the carbon cycle will appear at Chapter 13 which deals with the problem of climate change.

Biological evolution and coevolution are dealt with Park* (2001) and Krebs* (2001). Faber *et al.* (1996) take a general formulation of evolution to be one of the distinguishing conceptual foundations for ecological economics. Norgaard (1994) looks at economic development as a process of coevolution involving economic and environmental systems. Kauffman (1995) looks at the way that the mathematical developments noted above can be used to understand the evolution of complex systems in nature and society. The historical coevolution of living and non-living systems in the history of planet earth is the source of the 'Gaia hypothesis' advanced in Lovelock (1979) and Lovelock (1988).

WEBSITES

The Encyclopedia at the ISEE website, http://www.ecoeco.org, includes two short entries relevant to this chapter – one on 'Entropy' by S. Baumgärtner and one on 'Resilience defined' by C. S. Holling and B. Walker. Holling is originator of the idea of resilience as set out in the chapter here, which is sometimes referred to as 'resilience in the sense of Holling'.

EXERCISES

1. Set up a spreadsheet simulation for exponential growth to confirm the doubling time results of Table 2.2, and that 0 is an equilibrium.
2. Set up a spreadsheet simulation for logistic growth with $K = 100$, and plot growth over time for $r = 0.1$, $r = 0.25$ and $r = 0.5$. Confirm that 0 and 100 are equilibria.
3. Set up a spreadsheet simulation for

$$y_t = (3.7 \times y_{t-1}) \times (1 - y_{t-1})$$

and do simulations for initial values for y of 0.5 and 0.501. This is an example of chaos – the small shift from 0.5 to 0.501 produces a completely different time path.