The Origin and the Evolution of Firms

The Origin and the Evolution of Firms

Information as a Driving Force

Joop A. Roels

IOS Press

ISBN 978-1-60750-956-1 (print)
ISBN 978-1-60750-957-8 (online)

Cover photograph: An impression of the headquarters of the former company Gist-brocades. Courtesy Mr. F. Zieck.

Published by IOS Press under the imprint Delft University Press

IOS Press BV
Nieuwe Hemweg 6b
1013 BG Amsterdam
The Netherlands
Tel: +31-20-688 3355
Fax: +31-20-687 0019
email: info@iospress.nl
www.iospress.nl

PRINTED IN THE NETHERLANDS

About the Author.

The author studied chemical engineering at the Eindhoven University of Technology where he graduated in 1968.

In Royal DSM, he is Vice President Life Sciences R&D and Strategic Planning until his retirement in 2008. In addition, he works in academic teaching and research at the Delft University of Technology from 1977 to 1984. For 10 years before his retirement, he participates in the management of a large R&D cooperation of a consortium of universities, research institutes and industrial corporations as a chairperson of the Program Council of the Top Institute Food and Nutrition.

He starts his industrial career in Royal Dutch Shell in 1968. In 1970, he joins the R&D department of Gist-brocades, a midsized biotechnology company that merges with DSM in 1998. In the early Gist-brocades career, he works on the application of mathematical models in the optimization of biological processes such as the penicillin fermentation. Modeling of biological processes is also the subject of his teaching and research activities at the Delft University. In his research career, he becomes conversant with biological systems, modern genetics, the nature of life in the perspective of the theories of Darwin, the nature of the genetic code and developments in non-equilibrium thermodynamics. The industrial career introduces the author to aspects of economic theory and the strategic management of industrial corporations including the strategic management of R&D. This importantly contributes to recognizing the role of information and uncertainty in the workings and the management of firms, science and technology.

In his early work on modeling biological systems, he becomes increasingly interested in thermodynamics and recognizes the power of this tool in the description of complex systems such as microorganisms. This results in his 1983 book (Roels (1983)) that importantly relies on the application of macroscopic models, including thermodynamics, to the modeling of biological systems.

The combination of his scientific interest and his exposure to decision making in an industrial environment contributes to the conviction that modern developments in thermodynamics allow an extension to apply to the processes that drive competition and evolution in industry. This provides the inspiration for this work.

Preface

The universe, the galaxies, the solar system, the earth, life, humankind, science, economies, markets and firms developed and develop in a spontaneous process we commonly term evolution. In this book, we attempt to develop a consistent theory that traces the evolutionary path that connects the birth of the universe in the Big-Bang to the complexity of today's socioeconomic system with, among others, its firms and markets. What are the regularities or, if you like the laws, and the forces that drive this evolution?

As far as living organisms are concerned, we owe a landmark body of theory to Darwin and Wallace. Darwin publishes his work "On the Origin of Species" in 1859. His work highlights the existence of a continuous "Tree of Life" that connects the present life forms to a common ancestor. In the evolution of life on earth, we clearly see a development in the direction of increasing complexity. Matter becomes increasingly organized. Evolution clearly involves increasing organization.

Later on, when we explore the molecular basis of life and its evolution, we realize that evolution derives from the processing and communication of information. Life largely rests on the evolution of DNA macromolecules that provide the information for the biological structure that engages in competition for scarce resources with other structures based on different DNA codes. In this book, we unveil that information processing and its perfection through competition forms the basis of the evolution of society based on the appearance of new sets of information beyond the DNA macromolecules that drive biological evolution. Information is the scarce commodity that drives economic progress. A continuous line of evolving information sets connects the Big-Bang to the present-day socioeconomic system with its firms and markets. We term the evolution beyond the replicating DNA exogenic evolution.

No doubt, the theory of evolution Darwin introduces in the 19th century marks a monument in the history of scientific thinking and for that matter the philosophy of science. Earlier in that same century, the French engineer Sadi Carnot analyzes the workings of the "heat engine". It is a contraption that allows the conversion of high temperature heat, obtained by the combustion of energy resources such as coal, into motive work. This greatly increases the availability of sources of motive power beyond those provided by human and animal muscles and e.g. waterpower. As we explain in this book, this increased availability of sources of useful work, triggers the industrial revolutions in the 18th and 19th century. These revolutions mark a strong increase in economic prosperity measured in terms of per capita Gross Domestic Product. This development becomes particularly manifest after 1850 when we witness the birth of many industrial initiatives that are the precursor of large industrial corporations that appear in today's economic landscape.

Sadi Carnot's analysis of the heat engine marks the birth of an important body of theory that, just as in the case of the theory of evolution, marks a revolution in the history and the philosophy of science. It spawns thermodynamics, the theory that governs the transformation of sources of energy, such as heat and the various flavors of work. The theory of thermodynamics introduces a number of laws that put restrictions to the transformation of said energy sources. The first law states that energy is conserved: It cannot be created or destroyed in any process. Hence, the total energy in the universe today is the same as the amount that exists when it emerges 13.5 billion years ago. In addition, thermodynamics introduces the intellectually evasive quantity entropy. The second law of thermodynamics states that entropy can only increase in the processes that take place in a closed system. Hence, the total entropy of the universe increases in the evolution of the universe after its birth and it continues to increase in the future. This leads to an apparent conflict with the process of evolution we introduce earlier. Evolution leads to a local decrease of entropy reflected in the increase of organization of matter. We see the evolution of organized systems, such as organisms.

Prigogine and his coworkers (see e.g. Nicolis and Prigogine (1977)) importantly contribute to removing the apparent paradox of the laws of evolution and the laws of thermodynamics. The theory of non-linear non-equilibrium thermodynamics reconciles the local evolution of "Order out of Chaos" with the second law direction of spontaneous processes. The theory shows that in non-equilibrium systems forces exist that locally drive processes in a direction against the natural direction the second law dictates. Such a development can take place as long as for the universe as a whole entropy increases. This introduces the concept of dissipative structures that organize themselves against the direction of the second law by using sources of energy in the environment. Organisms are examples of such dissipative structures. This book extends the concepts of thermodynamics beyond the conventional "physical" sources of work to include economic work. We thus identify organizations like firms as dissipative structures.

Another important result from thermodynamics introduces the notion that entropy intimately relates to the information we have about the details of complex systems. It identifies the information that is missing in our picture of reality as a scarce quantity that comes at a cost. In addition, this leads to the identification of the forces that drive evolution. This highlight the important role of information in thermodynamics and this closes the circle between evolution, based on the processing and communication of information, and our extended interpretation of thermodynamics.

We combine the information-based forces of thermodynamics with the information-based forces behind the theory of biological and exogenic evolution to arrive at a systems theory of evolution. This theory shows that evolution rests on the coupling of forces that drive processes in the natural direction defined by the second law, to processes that drive evolution against the natural direction dictated by the second law. This also holds for the coupling of economic work that drives our economy and the evolution of the socioeconomic system with its markets and firms, to the downhill processes that result from the non-equilibrium situation in the universe. Our theory allows tracing the path from the Big-Bang to the contemporary socioeconomic system. In this way, we trace the origin and the evolution of firms.

Table of Contents

CHAPTER 1. INTRODUCTION.

1.1. Introduction.

Daily observations show that reality is complex. At the intermediate level of sizes in the biosphere, we observe complex organized forms of matter. Plants and animals are examples of such organized matter. We hold the position, following Darwin (Darwin (1859)), that the biosphere on earth arose spontaneously in the process of evolution. Evolution introduces change, also in the direction of increasing complexity, through processes driven by random events, by chance. As we explain in this work, evolution is a necessary consequence of the conditions in the universe. Evolution towards increasing complexity is inevitable. The Nobel laureate Jacques Monod (Monod (1971)) uses the notion of "Chance and Necessity".

The second law of thermodynamics provides an arrow of time to the direction of spontaneous processes. Over the years, the literature shows a variety of formulations of this second law. One of the popular formulations is that systems, if left alone, develop in the direction of increasing disorder. In absence of maintenance buildings turn into ruins; we never observe the reverse process. The theory of evolution of Darwin, on the other hand, explains how complexity, such as in living systems, results under the conditions on earth and broader in the universe.

The apparent contradiction between the theory of evolution and the second law of thermodynamics puzzled many early investigators. Evolution generally proceeds in the direction of increasing complexity. Organized systems, such as humankind, evolve apparently spontaneously out of an initially unordered state. Fortunately, scientific developments in the 20th century lead to a reconciliation of thermodynamics and evolution. Prigogine and his coworkers formulate a thermodynamic theory of evolution (Glansdorff and Prigogine (1971), Nicolis and Prigogine (1977), Prigogine (1980), Prigogine and Stengers (1984)).

The discovery of the role of DNA and RNA in biology shows that biological evolution is of an informational nature. The processing and communication of information lies at the roots of the evolution of increasingly complex organisms. Evolution results for an extensive period of the history of life on earth exclusively from the further refinement of DNA (or RNA) macromolecules. Relatively recently, other ways of developing and communicating information emerge when the brain appears and evolves to sophistication when *Homo sapiens* appears. The brain triggers the so-called exogenic evolution; evolution based on transferring and developing information beyond the information carrier DNA. This also leads to the development of the socioeconomic system, with its institutions such as universities, economies, markets and firms. Communication of information through e.g. teaching and in written form complements the communication by DNA replication. It is the ambition of this book to investigate the relation between the theories mentioned above and the storage, processing and transfer of information. In this way, we grasp the dynamics of economies, markets and industries and trace the evolution of society as we know it today back to the birth of the universe.

Most of the systems that are of interest to physicists, chemists and biologist are far too complex to model in detail. In physics, this leads to the widespread use of macroscopic models that take only part of the microscopic details of complex systems into account. In using macroscopic models, science develops a reduced information picture of the systems of interest. As we explain later, this approach leads to limitations to the predictability of the future behavior of the systems we explore. In addition, this leads to limitations to the extent we can exploit the potential sources of energy and economic value (economic value is a concept we introduce and discuss in Chapter 5) in the system. The extent of this loss quantifies in the statistical entropy of the macroscopic description. It provides a quantification of the information that is lacking in the macroscopic description. We discuss the nature of the concept statistical entropy more extensively in Chapter 4.

In this book, we discuss some further limitations of modeling when applied to evolving systems. The first limitation relates to the observation that evolution develops in the direction of an increasing interaction between the evolving system and its environment. Here a traditional assumption of modeling breaks down. In most modeling exercises, relevant processes only take place in the system. The environment does not change due to the evolution of the system. We show that this is often not the case as the environment and the system engage in a process of co-evolution.

A second modeling problem is of a very different nature. One of the basic assumptions in modeling is that the system and its model are independent. This implies that the existence of the model does not influence the processes that take place in the system. This proves a dangerous assumption if the actors in the system know the predictions of the model and use these predictions to shape their behavior. This leads to a fundamental problem in the modeling of the socioeconomic system and the validation of the resulting models.

The ambition of this book is to develop a consistent theory of evolving systems with special reference to industries, markets and economies. We show that the basic driving forces behind the transactions that take place in our markets, industries and economies rest on the creation and maintenance of asymmetries in information. Furthermore, the value (and the cost) of the information is quantitatively defined using the concept of statistical entropy. This results in a general theory of evolution applicable to a wide range of systems. The theory allows tracing the origin of society.

We apply this basic formalism to systems in which asymmetries in information exist and develop. Furthermore, we analyze aspects of the theory in terms of accepted economic concepts, such as the perfect competition model, transaction costs economics, the concept of dynamic capabilities and the evolutionary approaches to organizations. This leads to the conclusion that the application of evolutionary approaches to markets, industries and economies does not rest on a mere analogy with biological evolution but is a reflection of a general evolution theory of complex systems. We argue that there are both similarities and differences between biological and socioeconomic evolution. The theories underlying the approach (thermodynamics of complexity, information theory, statistical thermodynamics and the theory of evolution) are not free from mathematical intricacies. We describe the formalisms in earlier work on the subject (Roels (2010)). This book differentiates itself by avoiding mathematical intricacy as much as possible without sacrificing rigor. It tries to address a multidisciplinary community of readers.

The main distinguishing feature of this book is that it develops a conceptually consistent framework for the existing concepts used in organizational economics in a way that should be accessible to readers not familiar with modeling approaches in physics, chemistry and biology. The author hopes that this book thus augments on and complements existing approaches in the literature on organizational economics and evolutionary approaches to organizations. In addition, we illustrate the application of the systems theory of evolution analyzing the evolution of the universe, the evolution of earth and life on earth, the scientific revolution and the industrial revolutions and derive the general features of these systems as non-equilibrium phenomena that lead to evolving information sets of an ever-increasing complexity. The book thus bridges the gap between the Big-Bang that creates the universe and our present-day socioeconomic system based on one universal theory of evolution.

Evolution is very pervasive, its influence stretches from the universe at a cosmic scale through our socioeconomic system with organizations like firms, to science and technology and the invisible viruses that caused a Mexican flu epidemic in 2009. Also our species, *Homo sapiens*, is a product of evolution.

We stress that evolution continues to progress as the forces that drive it still exist and continue to develop. Consider the development of language. Since the appearance of the ancestor language of Dutch, called Diets, it significantly changed and it is in a process of constant evolution today. Whether this leads to any improvement is a matter of taste and depends on the definition of improvement we choose to adopt. Certainly, evolution has an arrow of progression; it proceeds in a certain direction. Whether this direction leads to improvement depends, as said, on individual taste and the definition of improvement. To evolution itself, this question is not even marginally relevant.

In this book, we show that where conditions exist that necessarily lead to evolution, it exhibits certain regularities or laws no matter what the substance of the evolution may be, be it languages, industrial corporations, religions, human cultures, galaxies or minute viruses. Laws like the laws of physics govern the evolutionary path. Just as the law of gravitation invariably leads to planets orbiting stars, laws exist that govern all evolution processes. This book unravels some of the regularities in evolutionary processes. We discuss a systems theory of evolution (Chapter 6), show its foundation from the scientific perspective and illustrate its application to reality.

Before embarking on our journey, we provide the reader with a preview of the main conclusions we intend to reach. In addition, we highlight the flavor of the main elements of the organization of this work.

1.2. The main ambitions and features of this work.

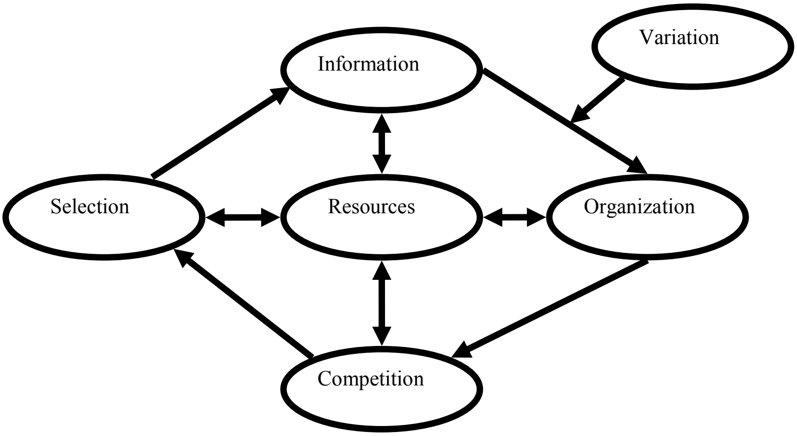

Fig. 1.1. Evolution through change by chance: Learning by doing.

In this book, we intend to analyze a wide range of evolutionary phenomena ranging from the very large, stellar objects, to the very small, e.g. the molecular machinery underlying life phenomena. We also argue that immaterial systems such as science and technology are subject to laws of evolution. In fact, also our economies, institutions and firms are subject to and result from processes following the general features of evolution. We intend to explore the sequence of complex events that connects the Big-Bang to our present socioeconomic system. The processes that created increasingly complex structures such as the objects in the universe, the life forms and economies and industries result from the operation of forces that drive these changes. We identified these forces in earlier work (Roels (2010)) and we further illustrate the nature of these forces as a direct consequence of the conditions that apply when the universe emerges. We thus arrive at the general features of evolution both in material systems, e.g. organisms, and in immaterial systems such as science and technology.

A concept we introduce and frequently use is the learning by doing cycle (Fig. 1.1) that is very basic to any type of evolution. In fact, it represents the archetype of an evolutionary cycle. We discuss this cycle in detail in Chapter 6 in the context of a systems theory of evolution. Fig. 1.1 highlights the pivotal role of information. All the systems we discuss in this work develop and communicate an information set, be it in a tangible form (e.g. an organism's DNA) or in an abstract or intangible form, such as in a scientific theory. The information translates in some kind of organization, e.g. the organism's structure, its phenotype, the understanding and the predictions that result from a scientific theory or the products and services and other aspects of the manifestation of an industrial enterprise. We will collectively indicate these manifestations of the information sets as "phenotype", also beyond the purely biological systems. In the process of communication of the information set changes appear, either by more or less directed change, e.g. by R&D or the gravitational force in the universe, or by error. This variation is vital, as it is one of the sources of progress. The phenotype competes for scarce resources that are the origin of the forces that drive evolution. Examples of such resources are energy and, more general, value, a concept we introduce in Chapter 5. We discuss these resources in a general way below and in far more detail in later chapters. This competition may also take place in the intangible world, e.g. where rivalling scientific theories compete for explanatory power in the light of experimental evidence. The competition for scarce resources leads to selection of the most competitive phenotype and hence its related, often changed, information set. This process leads to the selection of information sets that are coding for phenotypes that are more efficacious. It leads to closing a positive feedback cycle that is the basis of sustained evolution. We stress that closing the cycle leads to a blurring of cause and effect. Once the cycle closes, the familiar distinction between cause and effect disappears as these cyclically interact to drive evolution. The activities of the structure that the information sets code for both result from and are instrumental in the creation of the forces.

In addition, a cause and effect cycle closes between the structures and their environment. This in principle holds for all the elements of the cycle. The resources, the variation, the information set, the phenotype, the competitive environment and the selection process relate causally and are both cause and effect of the evolution taking place.

A second aspect that we introduce is the concept of sustained evolution (Fig.1.2) as introduced by Prigogine and his co-workers (e.g. Nicolis and Prigogine (1977)).

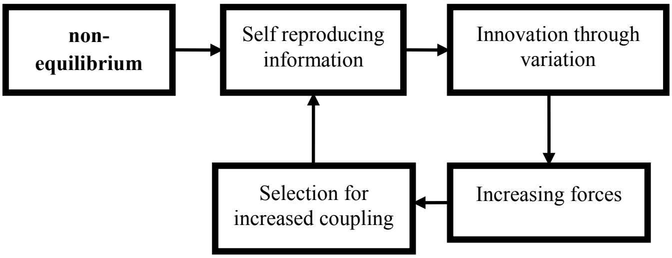

Fig. 1.2. Sustained evolution.

The drivers of the process of sustained evolution derive from the fact that we have a system that is not in equilibrium, i.e. in the system forces exist and evolve that drive processes. In the

Big-Bang theory of the evolution of the universe, this non-equilibrium situation emerges on the creation of the universe. It initiates a sequence of events in which information sets emerge that define the structures in the evolving universe. These information sets are self-reproducing and are subject to changes both in content and in nature. In the case of the evolution of the systems around us in the universe and on earth, the Bing-Bang creates a large amount of energy. This is the source of the forces that drive evolution and the origin of the non-equilibrium situation that arises on the creation of the universe. This amount of energy, the source of all the driving forces in the universe, remains constant in the history and in the future development of the universe. The famous first law of thermodynamics proclaims energy to be a conserved quantity, i.e. the amount of energy in the universe cannot change in any imaginable process. The amount of energy in the universe and the potential to perform work, including economic work to add economic value, thus remains constant in the estimated 13.5 billion years of evolution of the universe to date.

In the universe a background of electromagnetic radiation exists that is to a high degree of accuracy the same in any direction we choose to observe it. However, small variations in the energy of the background radiation do exist. These variations represent the initial information characterizing the early universe. The corresponding variation in energy density forms the initial self-organizing information. These differences in energy density amplify by gravitation and lead to increasing inhomogeneities in the local energy density in the evolving universe. This triggers the subsequent ignition of the stars by the energy that nuclear fusion processes release. Hydrogen was one of the early elements that form in the evolving universe and is the initial resource for fusion processes.

One of the systems that form in this way is our solar system that also contains all of the higher chemical elements beyond Hydrogen. In this way the constituents of the life forms, being mainly Carbon, Hydrogen, Oxygen, Nitrogen and Phosphorus appear. In the solar system, the planet earth develops. Life emerges on earth using the scarcely available sources of forces provided by mainly inorganic fuels and/or organic materials that form under the influence of solar radiation. DNA based self-replication appears and this leads to the first photosynthetic bacteria that harness the energy in solar radiation. This drastically increases, both in size and in number, the forces that drive evolution and fuels the development of organisms that feed on the organic material produced by the photosynthetic autotrophs. It leads to the food webs we know today. In the process of evolution on earth, the brain emerges and becomes perfected when the genus *Homo* starts to evolve and our species *Homo sapiens* appears. With the brain, new replicators appear beyond DNA and a so-called exogenic evolution emerges resulting in science and technology, the first and second industrial revolutions and our present-day socioeconomic system. In these processes the driver of evolution continues to derive from the forces that result from the initial non-equilibrium created in the early universe and the selection of sets of information that compete for the capacity to do useful work that derives from those forces. As is clear both the nature of the forces and the nature of the information sets, be it DNA, information stored in the brains of the life forms or information in written form or in computer files, change drastically in this extended process of evolution. We again stress that both the forces and the information sets are part of a cycle in which the information sets are both the consequence and the source of the forces and vice versa. In fact, the information contained in the information sets allows the transformation of the potential to do work, also in an economic sense, contained in the energy that emerges in the Big-Bang, into work that is useful from the perspective of the structures that feed on these energy resources. The structures for which the information sets encode are termed dissipative structures as these feed on the capacity to do work derived from the energy resources in the environment, i.e. these structures dissipate energy. These entities contain the information to create forces based on the available energy resources and the information to effectively couple to these forces to fuel their growth and development.

It is perhaps useful to illustrate these phenomena using an example. We imagine a landscape in which a river is flowing downhill. The energy contained in the difference in altitude characterizing the slope that drives the flow in the river, is the source of the potential to do work. In order to extract useful work from this potential we need additional information, e.g. a blueprint of a dam to build in the river and an electricity generator in the dam. Building these structures in the river allows the generation of electrical energy that is able to do work using a variety of electrical appliances. Without that additional information, the potential to de work is not available as a source of useful work.

The discussion above leads to the identification of a very important second concept that emerges in the theory of thermodynamics. It is the difficult and evasive concept of entropy. Entropy is a quantity that is of informational nature. It quantifies the amount of information that is lacking to transform the energy in a resource fully into useful work. In fact the essence of the second law of thermodynamics is that in any natural process in a so-called closed system, that is a system that cannot import energy resources from the environment, the information lacking to fully use the potential to do work contained in the energy in the system can only increase and the ability to extract useful work from that energy resource can only decrease. Entropy defines the direction in which natural processes spontaneously proceed. Any such spontaneous process results in the production of entropy.

The cycle of sustained evolution depicted in Fig. 1.2 clearly emerges in this broad-brush description of the processes that finally lead to our present-day socioeconomic system.

In Chapter 6 where we introduce the systems theory of evolution, we highlight the following conditions as characteristics of the evolution of dissipative structures such as organisms, firms and economies:

- The existence of sources of energy.
- The appearance of self-replicating structures that develop the information needed to transform available energy resources into useful work, also in an economic sense. This fuels the maintenance and further development of dissipative structures such as organisms and organizations.
- A mechanism allowing the development of new information sets by the introduction of variation in the information sets of the self-replicating structures. This introduces the elements of both change and chance in evolutionary processes.
- The developments of an increasing variety of information sets that both create and couple to an increasing variety of forces derived from the available energy resources. These information sets and the associated structures compete for ultimately scarce resources. This leads to the selection process that closes the cycle of evolution.
- The forces derived from the available resources start increasing in diversity and in fact, the structures that develop become resources for other structures. In biological evolution, this leads to primary producers such as plants that directly use solar radiation, herbivores that feed on plants, predators that feed on herbivores and omnivores such as humans that feed on both.
- In addition, the variety of information sets increases. Systems evolve based on a diversity of ways of storing, communicating and selecting information. Thus, in biological evolution structures outside the DNA, such as the brain and hardware and software detached from the biological systems such as books and computer files, succeed and complement the immortal coils of the DNA macromolecules. Exogenic software and hardware start complementing and dominating the function of DNA hardware.

Introduction.

In this book, we analyze the elements of the process of evolution and try to indicate regularities that appear in the variety of evolving systems on earth and in the universe.
It is important to recognize the ubiquity of the cyclic interactions when we analyze systems that are subject to evolution. In the remainder of this work, we identify a great variety of such complex and vital cycles. Some of the cycles we analyze are:

- The learning cycle that we introduce in the beginning of this chapter.
- Cyclic interactions in the evolution of the universe.
- The carbon cycle on earth.
- The co-evolution of life and the environment on earth.
- The cyclic co-evolution of humankind and society.
- The supply and demand cycle of industry.
- The evolution of science driven by the hypothesis-experiment cycle.
- The cycle of science and technology.

The forces that drive evolution derive from the capacity to do useful work including work to create economic value. The essence of these forces exists in information based differences in the ability to extract useful work from the available resources. In the systems we analyze we identify a wide variety of sources of forces. In the early stages of the evolution of the universe, the weak electromagnetic force is responsible for the formation of neutral atoms. This changes the nature of the interaction of matter and radiation to such an extent that radiation largely decouples from matter although, as we discuss later, the interaction between matter and radiation that remains is vital to the further evolution of life on earth as it allows photosynthesis. After the formation of the nucleons by the strong force, the electromagnetic force creates neutral atoms, primarily hydrogen and helium, the life supply of sources of matter for the universe. By the gradual decoupling of matter and radiation we progressively enter the matter dominated era and after the formation of neutral atoms gravitation becomes the force that lead to the formation of large concentrations of mass that ignite to form solar systems, initially by the transformation of hydrogen into helium and later on the synthesis of heavier elements unto iron. In a sequence of processes, this leads to the formation of our solar system and the earth some five billion years ago.
A vital element of the evolution of the socioeconomic system is the development of new forms of self-replicating information. What exactly fuels the emergence of life is still unresolved. Several competing theories exist for the explanation of this crucial step in evolution. The first life-like structures show some or all of the general features that drive evolution. The initial structures were autotrophic, i.e. could not depend on resources that were a product of already existing life forms. We do not delve further into these speculations and assume that RNA and ultimately DNA based replicators emerged on early earth quite early in evolution, at least some 3-4 billion years ago.
A very significant development needs to take place. Today, the overwhelmingly dominant source of non-equilibrium on earth derives from solar radiation. In order to access this intrinsic capacity to do useful work life has to develop an effective coupling mechanism to transform this potential to do work into a force that drives useful work. This depends on inventing ways of interaction between radiation and matter in a way not destructive to matter itself. Most probably photosynthetic bacteria of the cyanobacter species invented this coupling mechanism. The first photosynthetic bacteria do not evolve oxygen; they are not able to split oxygen from water. Later on probably some 2-2.5 billion years ago organisms appear that generate oxygen. This is a landmark event as it leads to massive reduction of carbon dioxide in the atmosphere and a concurrent increase of oxygen. The oxygen generating organisms arise in the oceans. When life moves from the seas to the land, some 400-500

million years ago, terrestrial photosynthesis starts to develop. This results in a quick increase of oxygen to its present level. Through this process, a massive new source of economic value, stored in plants, becomes available and evolution enters a new stage. Herbivores start feeding on the plant material and in their turn become a resource that invites the development of new ways of coupling when the carnivores appear. Also the human omnivore fits in this pattern as well as the development of the socioeconomic system.

Photosynthetic life on early earth also contributes in another way to the resources available to the socioeconomic system. Life of the past that decays fossilizes to resources such as peat, coal, crude oil and natural gas. Today this is the most important source of energy and resources for society. The majority of the resources we presently use, mainly as source of free energy, are of a fossil nature.

We show that the energy consumption of society is a good indicator for economic activity, as a definite positive correlation exists between the so-called Gross Domestic Product (GDP) and energy consumption. The correlation between per capita earnings expressed in the GDP and energy consumption is strong indeed, as we highlight later in this book. We conclude that energy is the ultimate driver of the generation of value in an economic sense. This conclusively ties the development and further evolution of the socioeconomic system to the energy created in the Big-Bang.

We will unveil some further features of the complexity that characterizes the reality of the universe and for that matter the socioeconomic system. A system is a part of reality that we study from a scientific, technological or economic perspective. The rest of the universe is the environment, it is important if it interacts with the system. It may be a source of energy resources, economic assets and information. Fortunately, we need to consider only that part of the universe that interacts with our system in a meaningful way. The very definition of the system involves assumptions that have consequences for the usefulness of the results of our analysis.

In virtually all cases of interest (certainly for complex systems such as microorganisms or the socioeconomic system), a full description of the system is not practical and even impossible. We need additional assumptions to reduce the complexity to a manageable level. We take resort to a representation of the system, a model, reflecting some but not all of the system's complexity. To introduce diligent approaches towards reducing complexity is at the heart of the art of modeling.

We usually study systems in which many interacting entities appear. A full microscopic model has to take the behavior of all these entities into account. This is in most cases an impossible task. We need a more clever approach to avoid facing a dilemma. Fortunately, there is a way out, although, as we indicated and further highlight later in this work, avoiding complexity and hard labor comes at a penalty. The penalty involves inability to harvest the full potential of the energy resources in the system, be it in terms of capacity to perform useful work in a physical system or in terms of harvesting of economic value in socioeconomic systems. In addition, we lose in almost all cases relevant in practice, part of the ability to predict the time evolution of the system if it is not in equilibrium in the thermodynamic or economic sense. We lose the ability to predict future behavior in detail. This becomes clear when we develop the systems theory of evolution in Chapter 6 and in discussing specific examples of evolving systems, e.g. the evolution of the cosmos and biological and socioeconomic evolution in Chapters 9-14.

There is at least on additional complication in the modeling approach to complex systems and this is particularly important when we study evolving systems such as the systems on earth. In the classical approach to modeling, we make a distinction between the system and the environment. In the system, the processes of interest take place. The environment is assumed given and is not subject to change due to the processes of interest to our study. We show that in complex systems in which the evolutionary cycle closes, the interaction between the system and the environment steadily increases and reaches a level in which changes in the environment start to take place due to the increased interaction with the system we have initially chosen. This leads to a situation in which the distinction between the system and the environment is no longer valid from the modeling perspective. The environment becomes part of the system and needs to be included in our modeling exercise. We show that this phenomenon, termed co-evolution, also importantly applies to the socioeconomic system.

1.3. Organization of this book.

Realizing the ambitions of this work requires covering of a broad area of experience and science. We plan to organize this as follows.

Chapter 2 discusses the nature of laws, models and theories and the intricacies of their development and analysis. We explore the origin of the need for such instruments. The approach and its methods find their roots in the limitations of our ability to understand and explore reality. We discuss the evolutionary relevance of this scientific methodology for sustaining our species. This intimately relates to the significance of the evolution of the brain that appears, albeit in a primitive form, some 500 million years ago and results in the intellectual capabilities, at least the roots of those capabilities, when our species appears some 150,000 to 200,000 years ago.

The subject of Chapter 3 is thermodynamics and its four basic laws. We introduce this notoriously difficult and mathematically intricate subject mostly in a verbal way. Understanding the philosophy and the results of thermodynamics greatly facilitates grasping the roots of evolution. Chapter 3 focuses on macroscopic thermodynamics. It provides a highly simplified but very useful picture of a reality far too complex to grasp in detail. There the macroscopic modeling method comes to the rescue. In this very fruitful scientific approach, we reduce the complexity of the molecular picture, the so-called microscopic picture, to arrive at a description in terms of a much smaller number of macroscopic averaged quantities. This combines manageable complexity with a picture that allows prediction of useful properties of the system. Classical thermodynamics is a prime example of a macroscopic theory and illustrates both the predictive power of such theories and the limitations thereof.

Chapter 4 discusses the informational foundations of the macroscopic description. We discuss statistics and probability theory, again avoiding mathematical intricacy as much as possible. Furthermore, we introduce information theory. This body of science allows us to introduce a quantitative measure of the information needed to bridge the gap between our macroscopic information about a system and the information needed to specify the system in its microscopic detail. This allows analysis of the relation between the macroscopic description and the microscopic reality of the system. The analysis leads to a deeper understanding of the almost mystical evasive macroscopic quantity called entropy. As we show, it is not so much a property of the system but a reflection of the way in which the observer interacts with the system. It reflects (limitations to) the information the observer can have about the system's state. In this chapter, we also conclude that the lacking information or the entropy that characterizes the macroscopic description, is a scarce quantity that comes at a cost.

Chapter 5 provides the full theory of value, identifies the forces that drive evolution, and shows the relation of these forces to the informational limitations of our knowledge about complex systems. The concepts of statistical entropy and the cost of information define such forces. We analyze the difference between intrinsic or potential value and economic value. Economic value is the value that is available to do economically useful work.

Chapter 6 provides a roundup of the theoretical consideration and introduces the thermodynamics and general dynamics of sustained evolution. It highlights the pivotal role of information processing and communication. We show that information asymmetries are the source of competition and selection. These information asymmetries are the driving force behind all evolutionary processes. Introducing a systems theory of evolution leads to an understanding of the key aspects of biological evolution and the evolution of the socioeconomic system. We profoundly discuss the theories and the reality of evolution. This leads to an analysis of the emergence of information storage and communication based on the immortal coils that comprise DNA molecules. We highlight the cyclic interaction between the genotype of an organism, i.e. its DNA (or in the case of some viruses RNA), as a carrier of the information code characterizing the organism, and its functional structure, the protein based structural elements and enzymes. An analysis of this cyclic interaction leads to unraveling its consequences. In addition, we introduce exogenic evolution, evolution of information sets beyond DNA.

Chapter 7 analyzes the evolution of markets and industrial corporations and broader the socioeconomic system. We identify the nature of the firm as an information processing structure. It proves to be an example of the dissipative structures that provide the mechanism of organization in all complex evolving systems including the universe, earth, the biosphere on earth and our present-day firms.

Chapter 8 provides an introduction into elements of economic theory and explores the limits of our understanding of the socioeconomic system. We live in an age that exposes the limitations of our understanding of the dynamics of our economic system. The world plummeted into a severe economic crisis in 2008. The governmental authorities resorted to unusual measures in an attempt to stabilize the economy and to reverse the downturn. To date these measures seem to have at least some degree of success. The specialists that apparently failed to spot the dangerous situation before the crisis hit us, do not agree about the question whether we are emerging from the economic slump. At the time of the conception of this book in 2010 and 2011, the jury is still out on this issue. In this chapter, we also discuss classical economic approaches such as the equilibrium model of perfect competition and unveil the inherent limitations of this approach. We introduce the neo-classical and Keynesian philosophies. We analyze theories of economic growth and discuss the evolutionary perspective that emerges in economic theory rather recently.

Chapter 9 highlights where it all started. The Big-Bang creates the universe from an initial situation in which time and space do not exist. We trace the evolution of the universe and witness the birth of stars and planets in stellar systems like the Milky Way. In the outskirts of the Milky Way, our sun resides. We see how the chemical elements emerge that form the substance of biological life and of the resources supporting life.

Chapter 10 discusses the earth and its biosphere that develops under the influence of the source of energy the solar radiation provides. We analyze how life and the environment on earth co-evolve and highlight the cyclic interaction between life and the conditions on our planet. Our discussion in this chapter ends with the emergence of the early ancestors of *Homo sapiens*.

Chapter 11 discusses the emergence of the Hominins, early ancestors of *Homo sapiens* that appear 5-7 million years ago and describes their evolution to modern man. We analyze the increasing importance of exogenic evolution, the evolution of human society beyond the information storage in DNA. This way of information processing and communication forms the basis for tool making, perfection of hunting, introduction of agriculture and animal husbandry, introduction and perfection of language and ultimately leads to our socioeconomic system, with economies, markets, industries and other institutions. We end this chapter when modern science and the industrial revolutions start to become visible.

Chapter 12 further analyzes the development, methodology and dynamics of scientific progress and highlight that it follows the evolutionary patterns of the systems theory of evolution.

In Chapter 13, we discuss the evolution of technology and further highlight its impact on the competitiveness of humankind as a biological species. It results in the precursors of the industrial revolutions.

Chapter 14 presents a discussion of the first and second industrial revolutions and shows how these shape the present-day socioeconomic system.

Chapter 15 completes the analysis of the functioning of markets and industries with a discussion of the historical development of a number of today's leading corporations from the perspective of the theories this book develops.

Chapter 16 presents a roundup and develops the full evolutionary perspective of the socioeconomic system. It discusses the dynamics of economic development. We reemphasize a number of regularities that result from the general evolutionary perspective.

In Chapter 17, we conclude on this work with a summary of the approach and the main findings. This chapter also discusses the prospects and limitations of the evolutionary perspective and highlights questions that are unresolved from the author's perspective.

1.4. Conclusion.

In this chapter, we present the background and the challenging ambitions of this work. This work presents a mental picture of human society as we see it today, explains how it evolved and discusses what we can and cannot say about its further development.

To achieve our ambition we need to cover a wide range of subjects from a broad variety of scientific disciplines. We now embark on this challenging journey and hope that you will enjoy the scenery.

CHAPTER 2. LAWS, THEORIES AND MODELS.

2.1. Introduction.

A system is a part of reality that we study from a scientific, technological or economic perspective. The rest of the universe is the environment, it is important if it interacts with the system in a meaningful way. The environment may be a source of energy resources, economic assets and information. Fortunately, we need to consider only that part of the universe that interacts with our system in a meaningful way. The very definition of the system we adopt involves assumptions that have consequences for the usefulness of the results of our analysis.

Systems can interact with the environment in a variety of ways. For isolated systems, we can ignore the environment, as there is no exchange with the environment. The other extreme is an open system. It in principle interacts with every item in the environment and understanding its relevant aspects becomes vital.

In virtually all cases of interest (certainly for complex systems such as microorganisms or the socioeconomic system), a full description of the system is not practical and often impossible. We need additional assumptions to reduce the complexity to a manageable level. We take resort to a representation of the system, a model, reflecting some but not all of the system's complexity. To introduce diligent approaches towards reducing complexity is at the heart of the art of modeling. It represents a vital creative act and is pivotal to arrive at useful models. Such models are the mainstay of the scientific and engineering approach to understanding and exploiting our world and its resources.

Models can be of a variety of natures. Tangible models, such as a downscaled version of the real system, are one example. We often use so-called pilot-plants in the chemical industry. These serve to mimic the behavior of a large-scale plant at a fraction of the cost of experimenting with the full sized plant. In the aircraft industry downscaled physical model of a plane serve to test aerodynamics in a wind tunnel.

Non-tangible representations, e.g. mental abstractions, are another example of modeling. We can use a verbal model based on the observations on the behavior of the system. Such verbal models express observations on the system's behavior and result from inductive or deductive reasoning in words. We stress the difference between the inductive and the deductive components of reasoning. Deductive reasoning relies on a system of logic and does not involve additional creativity. It follows the line of reasoning the logical framework prescribes. Inductive reasoning on the other hand relies on intuitive creative steps beyond strict logic. Such creative steps are again at the heart of the reduction of complexity inherent in modeling.

Mental models can also use mathematical equations to describe the behavior of the system. All empirical, verbal and mathematical models involve theories that rest on assumptions that make it possible to grasp aspects of the complexity of reality. Mathematical models are the hallmark of the scientific approach in physics, chemistry and engineering and to an increasing extent in economics and biology. These models often rest on widely accepted laws of nature or more encompassing scientific theories describing regularities in the behavior of the real world. Examples are Newton's laws of motion. These constitute the basis of the theory of classical mechanics. The theory allows modeling the solar system and explains the orbits of planets around the sun. However, the theory also describes a much wider range of phenomena. In note 2.1, we summarize the laws of Newton for the interested reader. Skipping the note will not jeopardize elements vital to the understanding of the remainder of this work.

Note 2.1. Newton's laws and classical mechanics.

Three laws of Newton are the roots of classical mechanics. These specify the dependence of the motion of a body on the forces acting on it. These laws of motion are:

1. *First Law: A body remains in a state of rest or continues its state of motion at constant velocity and in a straight line unless a force acts on that body.*
2. *Second Law: A body subject to a force (F) accelerates in the direction of that force. The magnitude of the acceleration (a) is proportional to the force and inversely proportional to the mass (m) of the body; or mathematically:*

   $$a = \frac{F}{m}$$

   A more common equivalent expression of this law is $F = ma$.
3. *Third Law: Whenever a body exerts a force on a second body, that second body exerts a force on the first body equal in magnitude and opposite in direction. This is the "action equals minus reaction" law.*

Newton first proposes the laws of motion in Philosophiae Naturalis Principia Mathematica, published in 1687. Newton applies these laws to explain and investigate the motion of physical objects and systems. In the third volume of the text, Newton shows that these laws of motion, if combined with his law of universal gravitation, explain Kepler's laws for the motion of planets in the solar system. The law of gravitation of Newton states that two bodies are subject to an attractive gravitational force proportional to the product of their masses and inversely proportional to the square of their distance.

We concentrate on mental models and mathematical models in particular. In addition, we always try to formulate the consequences of mathematical models in a verbal way to improve the reader's understanding of the concepts.

Models, theories and laws derive from assumptions beyond the mere observations. Both the quality of the empirical material and of the inductive and deductive reasoning determine the validity and usefulness of the model. Therefore testing models by discriminating new experiments is necessary. This is an established part of the methodology of modern science. It relies on abstraction and logical thinking and comparing the results of the deductions and inductions with discriminating experiments on the real system or a close enough image of it. This experimental verification is often not possible in socioeconomic systems or biological systems due to limitations of a financial, practical or ethical nature. This is an important hurdle in the development of predictive models of such systems. We return to this complication later.

Because of the fact that theories and laws rest on assumptions and a necessary reduction of the complexity of the real world, science does not represent an absolute truth. The assumptions and the reduction of complexity may be falsified when new conflicting experimental facts become available.

2.2. The microscopic and the macroscopic approach to modeling.

We usually study systems in which many interacting entities appear. Consider the number of water molecules in a glass containing 1 liter of pure water, the order of magnitude being 10^{25}, or the many actors involved in socioeconomic interactions. A full microscopic model has to take the behavior of all these entities into account. This requires specifying the state of the system in terms of the so-called state variables of all these entities. By definition, a full set of state variables, such as spatial coordinates, velocities and chemical or biological nature, fully specifies the properties of the entities deemed relevant in the modeling exercise. As three numbers specify the location of the molecules and three number their speed, we need of the order of 10^{26} numbers to specify the system's state if we consider the example of 1 liter of water. At least, if we consider these aspects to describe the state of the entities exhaustively (this is a significant simplifying assumption). This is clearly an impossible task. If we write down one number every second, it takes about 25 million times the estimated age of the universe to write down the state variables. We clearly need a more clever approach to avoid facing a dilemma. Fortunately, there is a way out, although, as we indicate earlier and substantiate further later in this book, avoiding complexity comes at a penalty. The penalty involves inability to harvest the full potential value in the system, be it in terms of capacity to perform useful work in a physical system or in terms of harvesting of economic value in socioeconomic systems. In addition, we lose in almost all cases relevant in practice, part of the ability to predict the time evolution of the system if it is not in equilibrium in the thermodynamic or economic sense. We lose the ability to predict future behavior in detail. This becomes clear when we develop the systems theory of evolution in Chapter 6 and in discussing specific examples of evolving systems, e.g. the evolution of the cosmos and biological and socioeconomic evolution in Chapters 9-14.

We return to our glass of water. If we want to ascertain whether it is safe to drink the water from the perspective of danger of burning our lips, we do not need to consider the vast number of state variables that specify the detailed state of the system. We only need one state variable, albeit a state variable of a very different nature as we see in a while. In addition, this state variable is readily accessible. We only need to measure the temperature, e.g. using a thermometer. Temperature is a macroscopic state variable. It results from averaging the microscopic state variables of the objects in the system, i.e. the state variables of the water molecules in the glass. Temperature describes the movement of the many molecules by averaging their kinetic energy, the energy contained in the movement of the molecules. The temperature provides an adequate answer to the question if it is safe to drink the water or if it is too hot. This reduction of the number of state variables is the basis of the macroscopic approach in physics and chemistry.

Temperature is as said an example of a macroscopic state variable. We further use it to present a preview of a few other concepts of the methodology of macroscopic modeling. As it happens, macroscopic variables appear in two flavors. There are extensive macroscopic quantities that depend on the size of the system and intensive macroscopic quantities that do not depend on that size. Temperature does not change if we consider two equal glasses of water of the same temperature. This makes temperature an intensive quantity. The total volume of water doubles when we consider two glasses. This makes volume an extensive macroscopic quantity.

There is a second feature of the macroscopic approach or for that matter macroscopic thermodynamics that we illustrate using our elementary example. How do we make certain that the second glass of water that we add to the first indeed has the same temperature? To do this we simply measure the temperatures of the two quantities of water, e.g. using a mercury-based thermometer. We measure the temperature of our original water by putting the thermometer in en we allow the thermometer to exchange heat with the water in the glass until the temperature reading on the thermometer stops changing. We consider the final reading as the temperature of the water in the glass. In doing this we introduce a concept and an assumption. The assumption is that the heat exchange between the thermometer and the water in the glass does not significantly alter the temperature of the water in the glass, i.e. the amount of mercury that heats or cools down must be very small compared to the amount of water in the glass. This is a rather obvious assumption and it generally is valid to a good approximation. The concept is that of thermodynamic equilibrium. If the thermometer exchanges heat with the system, it reaches thermal equilibrium with the water in the glass. If we make the thermometer part of the system and isolate the system from the rest of the universe, we see that finally everywhere in the extended system the temperature is equal. We can extend this concept. An isolated system is in thermodynamic equilibrium if the extensive and intensive state variables do not change anymore. We now proceed by measuring the temperature of the second quantity of water with the thermometer and we see that the reading on the thermometer is the same. We think we can safely assume that the temperature of the two glasses of water is the same. Probably without noticing, we introduced the zero-th law of thermodynamics. It allows the specification of an unambiguous concept of temperature by noting that if system A has the same temperature as system B, B being the thermometer, and system C has the same temperature as system B, then the temperatures of A and C are equal. This discussion formally introduced the concept of temperature. It turns out to be very important in thermodynamics. We use the symbol T for its value.

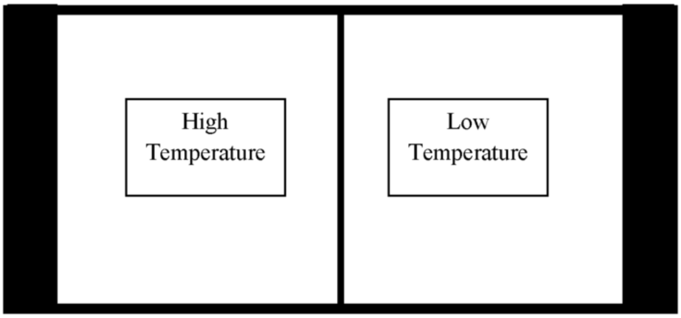

Fig.2.1. Contact between two pieces of metal of different temperature.

We proceed by using temperature to derive another important concept: The concept of a thermodynamic force. We consider the system depicted in Fig. 2.1. Two pieces of metal of different temperatures are in contact and isolated from the environment. The two pieces of metal exchange heat until their temperatures become equal. We know from experience that heat flows from the metal with the higher temperature to that with the lower temperature. This introduces the concept of a thermodynamic force. The difference in temperature, i.e. the difference in an intensive macroscopic quantity, provides a force that drives the flow of heat between the two pieces of metal. Later on, we see that to arrive at a consistent theoretical framework, we have to express the force in terms of the reciprocal of temperature but this does not matter at this stage. The discussion above leads to an important observation. Heat flows in a well-defined direction. The thermodynamic force sets an arrow of time. We always observe that heat flows in the direction of lower temperatures, never to higher temperatures. Coffee cools down if left alone and never spontaneously extracts energy from the universe to increase in temperature, irrespective the fact that energy is abundantly available in the universe. In the next chapter, we see that we stumbled on one of the alternative formulations of the second law of thermodynamics: In an isolated system, heat will flow from high to low temperature regions. We never observe the reverse, at least from the macroscopic perspective. This does, as we see later, not hold at the microscopic level as macroscopic laws only consider the average behavior of the microscopic entities. This becomes clearer when we study the statistical foundation of the macroscopic method (Chapter 4). In fact, to the thermodynamics purists the term heat flow that we used a few times in the last sentences is a little bit a loose and even dangerous term. Heat definitely is not a material substance (like e.g. water) that flows. When we say that heat flows, we mean to say that a process takes place by which heat exchanges between two systems. To avoid lengthy phrasing we often use the term heat flow and the reader should note that we therewith indicate the process mentioned in the previous sentences.

We summarize the main features of the macroscopic approach introduced so far. Heat and temperature turn out to be a macroscopic reflection of the speed of the molecules, more precisely their kinetic energy, the energy contained in their movement. (For the time being we will be sloppy and ignore that at this stage, we are not completely sure what energy really is). At the level of the individual molecules, the speed varies considerably and if we attribute the macroscopic concept of temperature at the level of a single molecule (something we cannot do) their "temperatures" vary largely. We resort to a concept from statistics: We invoke the so-called law of large numbers, i.e. if we consider a sufficiently large collection of molecules, the average speed of the molecules in two sufficiently large samples is almost equal for a system in thermodynamic equilibrium. In fact, if the samples are sufficiently large the difference becomes negligibly small. This is the essence of the macroscopic method: We simply average the behavior of individual entities to arrive at macroscopic properties that do not exhibit the large fluctuations taking place at the level of the individual entities. One macroscopic state variable takes the place of many microscopic state variables. The complexity of the microscopic level becomes manageable if we take a macroscopic perspective.

There is one further remark we need to make about the macroscopic method. The number of molecules we involve in the averaging has to be sufficiently large but the volume of the system involved in the averaging small compared to the total volume of the system. If that is not the case, problems arise if we consider systems that are not in equilibrium and spatial differences in the macroscopic variables are present. If we need to average over a significant part of the system's volume, to arrive at consistent average quantities, the variables are no longer continuous in the spatial dimensions of the system and we get into mathematical difficulty. Although this may be a problem in some systems, we ignore this complication.

2.3. Macroscopic balance equations and state equations.

An important modeling tool is the construction of macroscopic balance equations. These derive from the accounting principles of the macroscopic method. Consider the system depicted in Fig. 2.2. A number of macroscopic extensive state variables specify its state.

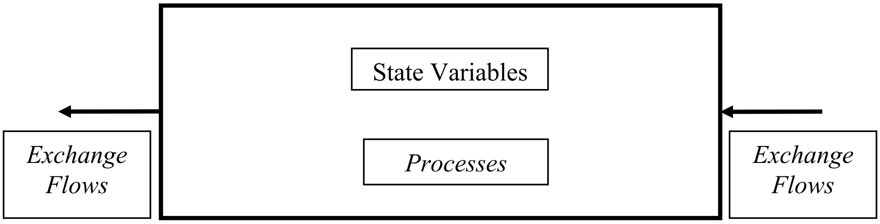

Fig.2.2. System for macroscopic accounting.

Macroscopic accounting considers all the effects that cause a state variable to change. This leads to equations for the rate of change of extensive state variables. For the general case, macroscopic extensive quantities change due to two types of causes. Firstly, there are flows to and from the environment that result in such change. Secondly, processes in the system may result in a change of the amount of the extensive quantity, e.g. chemical conversions may lead to such change. We can simply state:

Increase of state variable = Net rate of production in system + Net rate of exchange (2.1)

We promised to keep the mathematics to the bare minimum but here we introduce some mathematical shorthand notation. For each of the state variables under consideration eqn. 2.1 mathematically phrases as:

$$\frac{dY}{dt} = \Pi_Y + \Phi_Y \tag{2.2}$$

In eqn. 2.2, the differential at the left hand side stands for the rate of change of state variable Y. The two terms at the right hand side stand for the rate of change by transformation processes and the rate of change due to exchange respectively. Both terms at the right hand side of eqn. 2.2 can be positive as well as negative.

Eqn. 2.2 is the so-called state equation for the state variable under consideration. If we avail of expressions for the rates of exchange flows and processes, we can determine the time evolution of the state variable. This evolution defines the trajectory of the state variable. In fact, if we perform an exercise of the type discussed for every extensive state variable deemed relevant in the modeling exercise, we completely specify the time evolution of the system from the macroscopic perspective. This statement is only valid if we know the initial values of the state variables; we have to know the initial conditions. Later on, we see that this is only limitedly possible due to inherent limitations of the macroscopic method. Notwithstanding this complication, we achieve complete modeling of the system in a macroscopic sense if we perform the accounting exercise for all the relevant state variables, formulate rate equations for each of the exchange flows and the processes and have full knowledge of the initial conditions.

The rate equations derive from the theories and the underlying laws that govern the rates of these processes. We provide some examples of rate equations for exchange flows and transformation processes in the next note. Again, the readers only interested in the main arguments developed in this work, can skip this note. This applies to all augmenting and clarifying notes in this work.

Note 2.2. Examples of rate equations.

An example of a rate equation for a chemical transformation is the law of mass action kinetics. We consider a process in which a chemical compound A converts into a compound B and vice versa. Mass action law kinetics leads to the following equation for the rate of change of the concentration of A, i.e. the amount of A per unit volume, indicated C_A*:*

$$\frac{dC_A}{dt} = k_1 C_A - k_{-1} C_B$$

In this equation C_B *is the concentration of B and the k's are chemical rate constants.*

In fact, this equation involves a dangerous simplifying assumption as we assume the system's volume constant. For the general case, we cannot apply macroscopic accounting to concentrations, i.e. amount per unit system volume, but only to the amounts of chemical substances, being the product of concentrations and volume. The correct macroscopic balance equation reads:

$$\frac{d(VC_A)}{dt} = r_A V$$

In this equation r_A *is the net rate of production of A per unit time and volume. Mass action law kinetics leads to the following expression for the net rate of production of A per unit volume:*

$$r_A = k_1 C_A - k_{-1} C_B$$

Combination of the last two equation leads to the full formulation of the macroscopic balance equation. It shows that the first equation proposed in this note only follows if V does not change.
A rate equation for heat transport to a system derives from the Fourier law for heat conduction. It relates heat transport to the temperature difference between the system and its environment. If we indicate this difference by ΔT *the rate of exchange of heat follows as:*

$$\Phi_Q = S\lambda\Delta T$$

In this equation the term at the left hand side stands for the rate of transport of heat, λ *is the coefficient of heat conductivity, S is the surface area of the system.*

We conclude this section discussing a few special cases of the general balance equation introduced here. Firstly, we consider an isolated system, i.e. it exchanges nothing with the environment. In that case, eqn. 2.2 transforms to:

$$\frac{dY}{dt} = \Pi_Y \tag{2.3}$$

Eqn. 2.3 shows that only transformation processes contribute to the change of a state variable in an isolated system. Processes continue to proceed until the state variables become time independent and the system reaches equilibrium. This does not imply that the approach towards equilibrium in isolated systems is straightforward and smooth. Often the dynamics of the approach to equilibrium involves a very complex and lengthy evolution. We highlight this discussing the evolution of the universe (Chapter 9), a prime example of an isolated system, as there is nothing outside the universe to engage in exchange processes.
Another special example concerns the distinction between conserved and non-conserved macroscopic quantities. A conserved quantity is not subject to net production in the processes in the system. In that case, it changes by exchange flows only and the balance equation becomes:

$$\frac{dY}{dt} = \Phi_Y \tag{2.4}$$

The prime example of a conserved extensive quantity is total energy. The first law of thermodynamics, one of the cornerstones of thermodynamic theory, proclaims its conservation. We discuss this law in Chapter 3.
Finally, we introduce the concept of steady state, a state in which state variables have become

constant, just as in the case of equilibrium for isolated systems. In a steady state, the balance equation reduces to:

$$\Pi_Y + \Phi_Y = 0 \tag{2.5}$$

Steady states beyond equilibrium may occur in open systems if these exchange at least two different resources with the environment. There may be complex dynamics involved in the approach to steady states. In addition, steady states that conform to eqn. 2.5 may or may not be stable. We discuss these matters further in Chapter 6.

2.4. Linear and non-linear systems theory.

Application of the macroscopic balance equations to the extensive state variables results in the state equations describing the dynamics of the evolution at the macroscopic level. The construction of the state equations requires the formulation of equations for the rates of transport and the rates of the processes in the system. These derive, as an example, from established physical laws and theories and usually require further assumptions beyond those implicit in the macroscopic description. To predict the future evolution of the system we need, in addition to the state equations, the initial conditions, the state variables at the beginning of the evolution of the system. Only if we know the exact values of the initial conditions and the state equations describing the system, we arrive at a theoretical prediction of the evolution of the system. The state equations provide a reduced information picture in addition to an approximation to the real behavior of the system. We have a model mimicking the behavior of a real system. It generally does not behave completely the same as the very intricate reality. In the general case, the state equations may defy analytical solution resulting in a closed form mathematical expression for the evolution of the system. Often we rely on numerical computer calculations to solve the equations of the model.

In so-called linear systems, we encounter a class of state equations easily traceable by analytical methods (Hespanha (2009)). In these systems, linear first order differential equations result from the construction of balance equations. A simple case considers only one state variable, e.g. the number of organisms, indicated *N*, in a culture of a microorganism. A general equation describing such a system is, in its linear form:

$$\frac{dN}{dt} = \mu N \tag{2.6}$$

In eqn. 2.6, the term at the left hand side stands for the change of the number of organisms per unit time. The constant μ is the so-called specific growth rate. This equation states that the rate of the growth or decay of the number of organisms is proportional to, i.e. a linear function of, the number present. The linearity of the relationship makes it a linear description of the dynamics of growth. We can solve his equation straightforwardly and it results in exponential growth or decay (Fig. 2.3).

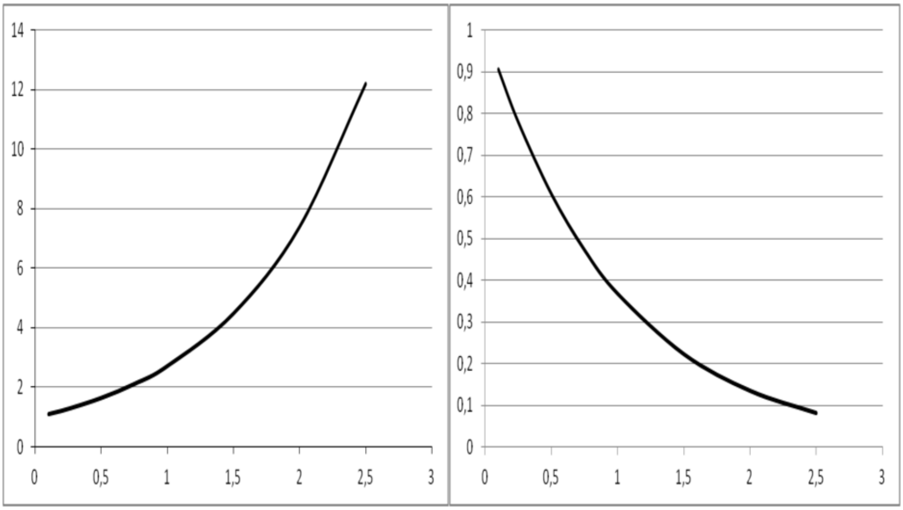

Fig. 2.3. Growth or decay of the number of organism (arbitrary units) according to a linear model.

We now consider a linear model involving two state variables, e.g. the numbers N_1 and N_2 of two organisms. We propose linear relations for the rate of change of the state variables of the following type:

$$\frac{dN_1}{dt} = C_{11}N_1 + C_{12}N_2 \tag{2.7}$$

$$\frac{dN_2}{dt} = C_{21}N_1 + C_{22}N_2 \tag{2.8}$$

In eqns. 2.7 and 2.8 the C's are constants.

For simplicities sake we study the time evolution of only one of the state variables. The calculated behavior depends on the values of the four constants. The theory shows that the general solution for the time evolution of the number of organisms is a sum of exponential functions of the form:

$$N_1 = a\exp(t/\tau_1) + b\exp(t/\tau_2) \tag{2.9}$$

In eqn. 2.9 *a* and *b* are constants. The τ 's are so-called relaxation times or time constants that define the dynamics of the system's evolution. In addition to real numbers that lead to exponential growth or decay, the time constants may also be complex, imaginary, numbers. In the latter case, the curves show sinusoidal fluctuations around an exponential growth or decay trend (Fig. 2.4). The graph at the right hand side of Fig. 2.4 indeed shows oscillations. We could consider this a model of fluctuations in the economic system, a line of reasoning we explore later in this book (Section 8.9). However, the fluctuations appearing in linear systems are regular and completely predictable unto infinity of time. The fluctuations may introduce an impression of irregularity and unpredictability. This is, however definitely not the case in linear system where complete predictability prevails.

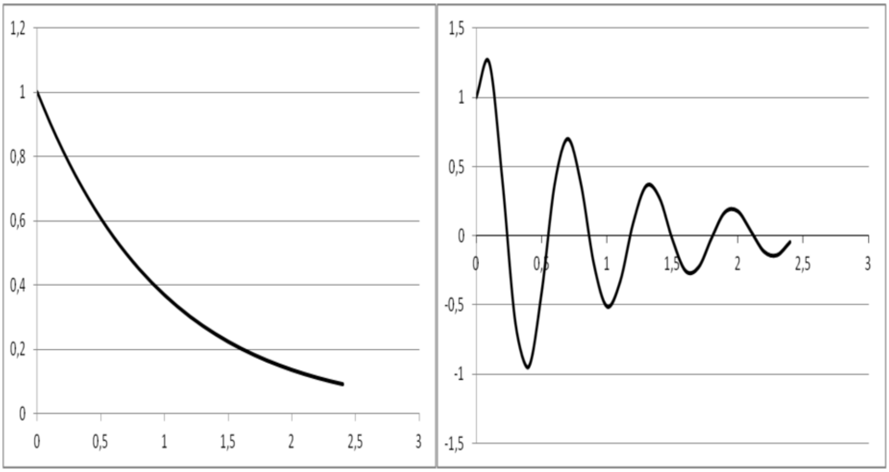

Fig. 2.4. Dynamics of a linear system involving two state variables.

For most systems of interest a linear systems approach does not adequately reflect the complexity of reality. For the general case, the state equations are non-linear in terms of the state variables. Already for systems of a rather limited complexity, this leads to much more complex behavior, e.g. oscillations that are far less regular than the sinusoidal ones appearing in the linear systems. However, for the non-linear region an additional complication kicks in. The systems may become infinitely sensible to the exact initial values of the state variables. Small differences in the initial conditions can lead to very large differences after a time that is characteristic for the system. These characteristic times may be short compared to the time horizon the observer has in mind for his predictions. The future of the system is no longer contained in the present value of the state variables that are a reflection of the past of the system. The future cannot be predicted from the present and the past, no matter how good a job of macroscopic modeling we do. This problem exists even if we have a perfect microscopic model involving the whole complexity of the system at the microscopic level. It becomes even more prominent if we avail of only a macroscopic reduced information picture of the system. We analyze these complications later in this work. This involves the chaos theories of complex systems, also termed the science of complexity (Gleick (1988)). These theories introduce the so-called “butterfly-effect” where the flap of the wing of butterfly in the US today may lead to a hurricane in China next week. The problem is also present in the famous three-body problem, a relatively simple system. It shows that we cannot accurately predict the evolution of a system in which three bodies move subject to the laws of Newton for an indefinite period of time. We also cannot prove its long-term stability. The complexity of reality escapes our modeling and theorizing possibilities.

2.5. Why do we need models?

The foregoing highlights some of the methodology and the problems of modeling. Here we revisit the question of the why of the use of models. The main reason is that we want to understand reality. However, this shifts the question to the query why we want to understand reality. The reasons are manifold but an important consideration is that relying on experimentation is often not economically feasible or to dangerous and unethical. We want to predict future behavior and to derive and implement measures that shape the future to our needs. A few examples are in order. Let us assume that we want to build an oil refinery at a cost of billions of dollars. Obviously, we cannot simply build one, see whether it works well and if it does not work properly, scrap it, build a second one, and repeat this process until we get something that shows reasonable performance. The amount of information we need is too large to obtain it by successive modification and selection of the proper one. Still this is the main approach on which progress by evolution relies before the era of the brain. It resulted in the extinction of most of the species that spawned from evolution. The same holds for experiments with e.g. the socioeconomic system. Random experiments may have unacceptable consequences. Hence, we need theories, laws and models to predict the future and to develop a future fitting our needs. This has a profound evolutionary significance. The "invention" of the brain in evolution makes building mental models of reality possible. This gradually results in an increasingly sophisticated ability to construct (mathematical) models that fit our need for prediction of the future. This aspect of so-called exogenous evolution is a new driver behind evolution. It complements the possibilities of molecular evolution at the level of DNA and RNA and is the main driver behind the further evolution of the competiveness of the species *Homo*. It enables us to exploit the economic potential the solar radiation provides, more extensively. The energy contained in solar radiation is the ultimate driver of evolution on earth. This brings us back to one of the main themes of this book.

It seems worthwhile to reemphasize a significant problem concerning the modeling of systems in which rational, or at least partly rational, actors exist. An important assumption underlying the philosophy of modeling is that the model and the system are independent. The existence of the model should not influence the outcome of the processes taking place in the system. This is a highly questionable assumption in systems containing actors aware of the predictions of the model. This influences their behavior, certainly if the outcome of the modeling exercise is an important factor guiding the actors. In that case, the assumption of independence of model and system is no longer valid. This results in an important philosophical problem that affects all modeling efforts in systems with actors that behave partially rational and are aware of the existence of the model. A potential solution is to take this part of human behavior into account in the modeling exercise. Apart from this solution being a difficult one, we get into the problem of infinite recursion as the model and reality start a process of co-evolution. The direction of this evolution is unpredictable from a modeling perspective. To date the author has no clues on approaches that avoid these problems and he probably never will arrive at such clues.

2.6. Conclusion.

In this chapter, we analyze the methodology of modeling of complex systems. We state that in almost all cases of practical interest a full description of real systems at the microscopic level, i.e. taking into account the properties of the many individual entities in the system, is unpractical and in most cases impossible. In contemporary approaches to the mathematical modeling of complex system, this leads to an averaging strategy resulting in reduced complexity macroscopic models involving macroscopic variables such as temperature. We show that two types of macroscopic variables exist: Intensive quantities not depending on the system's size and extensive quantities proportional to the system's size. We explore the macroscopic balancing or accounting approach for the amounts of extensive quantities that leads to state equations that in principle describe the evolution of a system in time.

We also show that differences in intensive macroscopic quantities define thermodynamic forces that determine the direction in which natural processes proceed.

The averaging process involved in macroscopic modeling leads to penalties of two natures. Firstly, part of the value intrinsically present in the system is no longer available to do useful work due to the limitations of our macroscopic information. Secondly, for systems not strictly linear, the evolution of the system may become infinitely sensible to the initial conditions. In those cases, a macroscopic model cannot predict the future evolution of the system.

CHAPTER 3. MACROSCOPIC THERMODYNAMICS.

3.1. Introduction.

This chapter summarizes the main features of macroscopic thermodynamics (Roels (1980, 1983, 2010)). Classical thermodynamics describes transformations of energy from a macroscopic perspective. An appropriate averaging process (Chapter 2, Chapter 4) reduces the microscopic complexity. We focus on three of the four basic laws of thermodynamics, numbered the zero-th to the third law of thermodynamics. We concentrate on three laws, as the fourth one is less relevant for our purpose. We introduce the zero-th law in Chapter 2 where we discuss temperature and thermal equilibrium. We need not discuss this in more detail at this stage so we focus on the first and second laws. The reader may wonder why numbering of the laws starts with zero. This derives from a historical mishap. The first and second laws already exist when the thermodynamics community realizes that formulation of the theory in a logically consistent way is not possible without the concepts of temperature and the associated thermal equilibrium.

The first law proclaims energy or internal energy (still undefined as we remember but we remedy that shortly) a conserved extensive macroscopic quantity. The second law defines processes and introduces an arrow of time in physics, i.e. it specifies the natural direction for processes in an isolated system. This involves the introduction of the state variable entropy. The third law, as said less important for our treatment, states that we cannot reach the absolute zero for temperature of about -273 degrees Celsius. The combined first and second laws of macroscopic thermodynamics allow the formulation of useful restrictions to the conversion of the different flavors of energy in nature.

In our discussion, we largely follow the line of reasoning of a very accessible treatment of thermodynamics titled “Four Laws that drive the Universe” (Atkins (2007)).

3.2. Energy and the first law of thermodynamics.

There are several ways to introduce the concepts of energy or internal energy and the first law. We use the symbol E to indicate energy. The preferred approach of this author introduces energy as an abstract notion defined by its conservation. In this way energy and the first law form a logical tautology. Every time science observes an apparent conflict between reality and the energy conservation law, a source of energy emerges restoring energy conservation.

A more convincing approach starts with relating energy changes to work exchanged between the system and its environment. The problem now becomes defining work. When we move a mass over a certain height in the gravitational field of the earth, we need to perform work. It equals the product of the gravitational constant, the mass of the object and the change in height. We then observe that in addition to mechanical work other types of work exist, e.g. electrical work and chemical work. Electrical work results if we move electrical charges over a gradient in an electric field. Chemical work results when we add or remove energy bearing chemical substances. To make a long story short it is possible to translate all other types of work into an equivalent amount of mechanical work.

We obtain the concept of energy as follows. We consider an adiabatic system, i.e. a system that does not exchange heat with the environment. We equate the change of the energy of the system to the work performed on it. This immediately implies that we cannot attribute energy an absolute value. We can define it by its change only. We arrive at a value for energy by relating it to an arbitrary datum level attributed zero energy. In contrast to work that is measurable in an absolute way, energy quantifies only with reference to a datum level. On definition of the datum level, energy becomes a function of state, i.e. it only depends on the, macroscopic, state of the system and not on the way it reached that state. The macroscopic state of a system defines its energy. Also in this respect, it fundamentally differs from work. We can define an amount of work performed on the system but we cannot define the amount of work that is present in the system, i.e. a function of state work does not exist.

Conversely, the change of a system's energy equals the maximum amount of work that we can obtain from the change of the state of a system. The change reflects the potential to do work.

The next step identifies another way by which the system's energy can change. We identify this mechanism of change by removing the insulation that made the system adiabatic and allowing it to exchange heat with the environment. The energy change in absence of work performed on the system, equals the amount of heat exchanged with the environment or, loosely phrased, the heat that flows from or to the system. We arrive at the conclusion that the change of the system's energy equals the sum of the work performed on the system and the heat it exchanges with the environment. This is the essence of the first law. If we relate to rates of change, i.e. changes per unit time, and rates of exchange of work and heat, we obtain the following expression as the mathematical equivalent of what we just stated in words:

$$\frac{dE}{dt} = \Phi_P + \Phi_Q \qquad (3.1)$$

The term at the left hand side of eqn. 3.1 is the rate of change of the state variable energy, the two terms at the right hand side are the exchange with the environment of the various forms of work and heat respectively,

With reference to the balance equation formalism introduced in Chapter 2, we see that by virtue of the first law, a production term does not appear in eqn. 3.1. This reflects the conservation of energy in transformations. The balance equation for energy involves no contribution due to processes taking place in the system, it only contains exchange terms. This leads to the alternative formulation of the first law: "The amount of energy does not change in any imaginable process". Energy can only change location: The total energy of the universe when it emerges equals to the total energy in the universe today and in the future.

3.3. The second law of thermodynamics: The concept of entropy.

The second law of thermodynamics is a monument of scientific thinking. It defines the direction of change and the forces that drive evolution everywhere in the universe, including biological and, as we argue, socioeconomic evolution. Before the second law, science had no definite view on time's arrow in the processes in nature. Our socioeconomic system is, as we argue, a direct consequence of the operation of thermodynamic forces. This argument is one of the main themes of this work. We state that the second law exerts its influence to a much broader range of phenomena than physics and biology alone. It also drives exogenic evolution, evolution beyond the evolution of material substances, such as the genetic code materialized in the DNA macromolecules. Exogenic evolution provides a decisive contribution to the creation and further development of the socioeconomic system.

The second law introduces the concept of the state variable entropy. Entropy, being a state variable, only depends on the state of the system not on the way the system reaches that state. We use the symbol S for entropy. It is notoriously difficult to grasp the physical and the philosophical meaning of entropy. Entropy enters thermodynamics as a consequence of using the macroscopic approach, i.e. it roots in the averaging process that translates the complex and numerous microscopic states into a vastly smaller number of macroscopic states. We suggest that we accept the concept of entropy as given for the time being. We postpone attempts to grasp the meaning of entropy to Chapter 4 when discussing the statistical foundations of the relation between the macroscopic and the microscopic approaches.

In developing the entropy concept, we follow history and put it in the perspective of the steam engine and the pioneering studies of the French engineer Sadi Carnot (Carnot (1824)). The invention and use of the steam engine was one of the key drivers of the industrial revolutions (Chapter 14). This example of the steam engine and the second law of thermodynamics illustrates important features of the interplay of science and technology. Firstly, the practical application in the steam engine precedes its scientific foundation in the concepts of entropy and the second law. Secondly, the theoretical development of those days rests on a misconception regarding the nature of heat. The then reigning theory considers heat a substance that flows and drives processes just as flowing water is a source for the generation of power in a hydropower plant. This proves a misconception. However, this had little influence on the usefulness of the technological concepts based on this misconception. Thirdly, it caused a revolution in the philosophy of physics with ramifications way beyond the economically important but scientifically rather mundane steam engine. So much for philosophy, we return to the storyline of this work.

The steam engine was the source of the need for understanding a part of reality that ultimately leads to the concept of entropy and the second law. A steam engine is a contraption in which heat flows from a high to a low temperature environment. The high temperature results from the combustion of a source of energy, such as coal. The high temperature heat fuels the conversion of water into steam. This results in a strong increase in volume and this expansion drives a piston to produce motive work. In essence, the random movement of the water molecules in steam changes into a useful directional movement of a macroscopic object, the piston. Heat transforms into mechanical work. Only part of the heat channels into motive power, waste heat needs to transfer from the system to a lower temperature sink. This sink may simply be the ambient temperature environment. Sadi Carnot's objective was the optimization of the transformation of the heat into useful work. The early nineteenth century engineers and scientists pursue various approaches to this objective. Some approaches involve a change in working fluid from steam to another substance. Others attempt operating at a higher pressure of the steam. Carnot argued that the efficiency of the perfect steam engine only depends on the difference in temperature between the source of heat and that of the sink. The appreciation for the importance of these findings of Carnot much later, make him recognized as the "father" of macroscopic thermodynamics.

The findings of Carnot trigger the formulation of a number of regularities contributing to the birth of the second law and the concept of entropy. One of contributors was Lord Kelvin. His name appears in the absolute temperature unit, the Kelvin. His observation is:

No cyclic process is possible in which heat is taken up from a hot source and converted completely into work.

This is an indication that, although heat contributes to energy, it has a lower quality than work, if we define that quality by the capacity of energy to do work.

The statement also implies that total energy is not a measure of the quantity of work we can derive from a source of energy. Energy reflects the potential to do work. It is not a direct reflection of the amount of work we can extract from the resource in a practical situation. The take-home message of Kelvin is that there exist different qualities of energy. We can convert high quality energy, e.g. mechanical work, completely into lower quality energy, heat. Heat, on the other hand, we can only partly transform into work.

We reach a further conclusion by combining the Kelvin and Carnot findings: High temperature heat has a higher quality in terms of its potential to do work than low temperature heat. This is the essence of the statement of another giant of thermodynamics, Clausius:

Heat does not pass from a body at low temperature to one at high temperature without an accompanying change elsewhere.

Both the statements of Kelvin and Clausius are consequences of the second law of thermodynamics. The formulations proposed above are logically equivalent. These are different ways of phrasing the same regularity, or if you like a law, regarding the way nature behaves concerning transformation of sources of energy.
We bring these statements together if we take refuge to the general balance equation for an extensive quantity as introduced in Chapter 2 (eqn. 2.2) and apply it to entropy *S*:

$$\frac{dS}{dt} = \Pi_S + \Phi_S \tag{3.2}$$

The term at the left hand side of eqn. 3.2 is the rate of change of the system's entropy. The terms at the right hand side are the contributions to that change of processes inside the system and exchange with the environment respectively. In discussing the first law, we note that energy is a conserved quantity. Its macroscopic balance equation shows no contribution due to processes in the system. In contrast, we introduce the conjecture that the statements of Kelvin and Clausius are a consequence of the second law of thermodynamics formulated as:

Any transformation inside a system results in production of entropy unless the system is in thermodynamic equilibrium where the entropy production and rates of macroscopic processes simultaneously vanish.

In shorthand, we express this as:

$$\Pi_S \geq 0 \tag{3.3}$$

We arrive at another frequently used formulation of the second law if we apply the combination of eqns. 3.2 and 3.3 (in this work, we use eqn. rather indiscriminatingly, i.e. also for inequalities like 3.3) to an isolated system for which the exchange of entropy with the environment vanishes: In that case, we reach the conclusion:

$$\frac{dS}{dt} \geq 0 \tag{3.4}$$

In words: For an isolated system, the entropy can only increase or stays constant if the system is at equilibrium. Alternatively, as the universe is according to the present view the only reality, hence has nothing it can exchange anything with, and furthermore is not at equilibrium, the entropy of the universe can only increase. This statement introduces a clear direction to the evolution of the universe. It was one of the ways in which Clausius expresses the second law.
To reconcile this statement of the second law with the observations we discussed earlier, we need to formulate an expression for the exchange term appearing in eqn. 3.2. First, we consider a system that only exchanges heat with the environment and identify the exchange of entropy with the environment. A fundamental thermodynamic result that we discuss further in Chapter 4, states that:

$$\Phi_S = \Phi_Q / T \tag{3.5}$$

Combining eqns. 3.2 and 3.5 for a system exchanging only heat with its environment, results in:

$$\frac{dS}{dt} = \Pi_S + \Phi_Q / T \tag{3.6}$$

We now consider a hypothetical so-called reversible process. Such process is so slow that the system always stays in thermodynamic equilibrium during the change. In such a mental construction that can exist only by approximation in the real world, the production term at the right hand side of eqn. (3.6) vanishes:

$$\frac{dS}{dt} = \Phi_Q / T \tag{3.7}$$

As we explained earlier the amount of heat is not a property, a state variable, of the system, it does not derive from the present state of the system. However, eqn. 3.7 shows that for a reversible change we can relate the amount of heat transferred to the change of the state variable entropy if we divide the heat exchange by the temperature. We term temperature an integrating factor for heat in the jargon of thermodynamics.

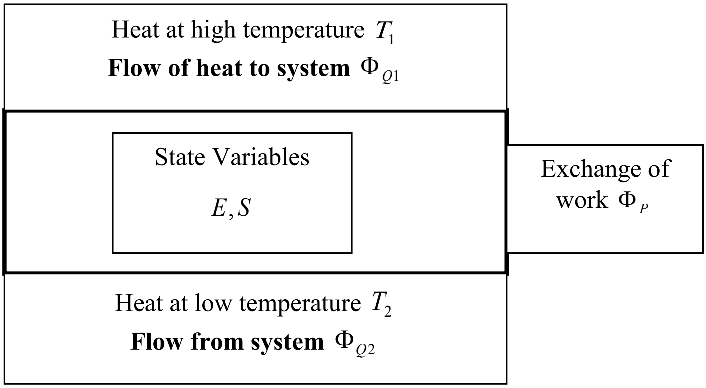

Fig. 3.1. System for thermodynamic analysis.

In Fig. 3.1, we introduce a more involved example in which a system exchanges heat with the environment at two temperature levels, i.e. we follow Carnot and study a model of the heat (or steam) engine.

The system in Fig. 3.1 interacts with a source of heat at a high temperature; the respective symbols indicate the amount of heat transferred and the temperature. The same applies to the transfer of heat to a sink at low temperature. The system exchanges work with the environment, its amount is Φ_P .The system's extensive state variables are energy and entropy. The system is subject to a cyclical operation. After one cycle the system returns to its initial state, i.e. the change of the state variables is zero. We now apply eqn. 3.6 to a cycle of the system:

$$\Pi_S + \Phi_{Q1} / T_1 - \Phi_{Q2} / T_2 = 0 \tag{3.8}$$

We also apply the first law, eqn. 3.1, to result in:

$$\Phi_{Q1} - \Phi_{Q2} - \Phi_P = 0 \tag{3.9}$$

If we eliminate the flow of heat to the sink using eqns. 3.8 and 3.9, it follows:

$$\Pi_S + \Phi_P / T_2 + \Phi_{Q1} / T_1 - \Phi_{Q1} / T_2 = 0 \tag{3.10}$$

If we multiply both sides of the last equation by the temperature of the sink, we obtain:

$$\Phi_P = -T_2 \Pi_S + \Phi_{Q1} (1 - T_2 / T_1) \tag{3.11}$$

We introduce the special case of a reversible cyclic operation. In that case, the first term at the right hand side of eqn. 3.11 vanishes:

$$\Phi_P = \Phi_{Q1} (1 - T_2 / T_1) \tag{3.12}$$

Several conclusions follow straightforwardly using eqn. 3.12:

- Due to the sign conventions, the system performs work on the environment if the term at the left hand side of eqn. 3.12 exceeds zero; in such case the system works spontaneously, i.e. no input of work is necessary. It is clear that this is only the case if:

  $$T_2 / T_1 < 1 \tag{3.13}$$

  This implies that heat flows spontaneously from a high to a low temperature. This is what we daily observe and it shows a consequence of the combined first and second laws of thermodynamics, the second law defines the direction of the flow of heat.

- We also conclude that a flow of heat from a low to a high temperature is possible if we are prepared to provide an input of work. E.g., a refrigerator works if we plug it into an electricity supply.

- We can agree with Carnot as the maximum efficiency of our engine, being the ratio of the work produced to the input of heat, indeed only depends on the temperature of the source and the sink. Normally, the efficiency is lower than the maximum as the entropy production in the system reduces the power output if the system does not follow a reversible cycle. This is immediately clear from eqn. 3.11 if we insert a positive entropy production. Unfortunately, a negative entropy production is not possible because of the second law.

- Returning to the statements of Clausius and Kelvin, we see that these indeed are a logical consequence of the second law. It requires that the net effect of all processes taking place in the system is production of entropy. We have to insist on the phrase “net effect of all processes”; if more than one process takes place, we have to remain silent about the effect of each process separately. This turns out to be an important observation as we show later on.

So far, we only considered systems that exchange mechanical work and heat with the environment. In the next section, we remedy that and introduce systems also exchanging chemical substances with the environment. Chemical substances carry both energy and entropy.

3.4. Constraints due to the combined first and second laws.

In this section, we analyze a system exchanging heat and chemical substances with its environment. We avoid a full mathematical treatment here and keep equations to the bare minimum necessary. The mathematically inclined can consult the literature (Roels (1983, 2010) and the references therein). In the system, transformation of material substances in chemical reactions takes place. In that case, introduction of a new concept that considers the combined effect of the energy and the entropy content facilitates identifying the restrictions posed by the first and second laws of thermodynamics, we refer to the concept of free energy that we discuss later. For simplicity's sake, we consider a system where two transformations involving two substances appear. This simple system reflects all the features of the general case involving a multitude of compounds and chemical reactions. Fig. 3.2 reflects the system.

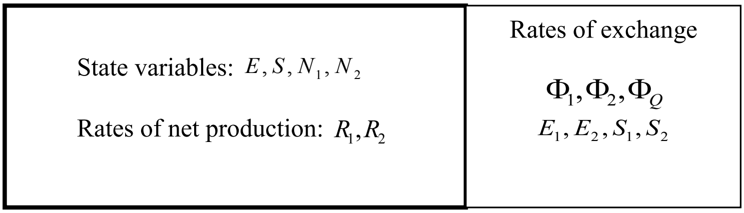

Fig. 3.2. System involving chemical transformation of material substances.

In this figure Φ_1 and Φ_2 are the flows of substances 1 and 2 to the system, Φ_Q is the heat flow to the environment. The N's are the amount of the two substances in the system. The R's are the net rates of production of the two substances in the transformations in the system. E_1 and E_2 are the energy contents per unit substance 1 and 2 respectively. S_1 and S_2 are the entropy contents per unit of those substances.

We write the first and second law based balance equations in the usual way:

First law:

$$\frac{dE}{dt} = \Phi_1 E_1 + \Phi_2 E_2 - \Phi_Q \tag{3.14}$$

Second law:

$$\frac{dS}{dt} = \Pi_S + \Phi_1 S_1 + \Phi_2 S_2 - \frac{\Phi_Q}{T} \tag{3.15}$$

For the amounts of compounds 1 and 2, balance equations of the same structure apply, for compound 1:

$$\frac{dN_1}{dt} = R_1 + \Phi_1 \tag{3.16}$$

For compound 2:

$$\frac{dN_2}{dt} = R_2 + \Phi_2 \tag{3.17}$$

We now invoke a fundamental result of thermodynamics, the so-called Gibbs equation of state:

$$\frac{dE}{dt} = T\frac{dS}{dt} + G_1\frac{dN_1}{dt} + G_2\frac{dN_2}{dt} \tag{3.18}$$

In eqn. 3.18 new state variables appear. G_1 and G_2 are the free energies of compounds 1 and 2 respectively. We return to the nature and definition of free energy below.
Combining eqns. 3.14-3.18 results in (see Roels (1983, 2010) for the mathematical detail):

$$T\Pi_S = -(R_1 G_1 + R_2 G_2) \tag{3.19}$$

Combing this result with the second law that requires entropy production positive if processes take place in the system, i.e. if it is not at equilibrium, the following restriction results:

$$R_1 G_1 + R_2 G_2 < 0 \tag{3.20}$$

Eqn. 3.20 states that free energy decreases because of the transformations in the system. We discuss the significance of this inequality shortly.
A similar restriction applies to a system in a steady state, i.e. a state in which the changes in the state variables become zero. That implies by virtue of eqns. 3.16 and 3.17 that the flow of each substance from the system equals its net rate of production, i.e. $R_1 = -\Phi_1$ and $R_2 = -\Phi_2$. Substitution of these equalities in eqn. 3.20 results in:

$$\Phi_1 G_1 + \Phi_2 G_2 > 0 \tag{3.21}$$

For a system in a steady state, the net effect of flows of energy bearing substances to the system must result in a positive net flow of free energy to the system if the system is not at equilibrium. In the words of the Nobel laureate Schrödinger (Schrödinger (1945)) when referring to living systems: The system "feeds on negentropy". With this statement, he indicates that transport of free energy bearing substances to the system, casually called negentropy ("negative" entropy), serves to keep the system in a non-equilibrium steady state by compensating the entropy production in the processes that take place in the system.

These restrictions are the counterpart of the ones we derived for system exchanging heat and work with the environment and prove to be important in analyzing complex patterns of transformations such as in an organism. Later on in Chapter 5, we argue that such restrictions have counterparts in systems where complex patterns of economic transactions take place. So much for mathematics that I hope remained sufficiently elementary.

The system we analyze here involves only two transformations and two compounds. What we find if we do the math for more complex systems (e.g. Roels (1983, 2010)), are the following statements that apply to any transformation pattern involving an arbitrary large number of compounds and reactions:

In any system, the net effect of a pattern of transformations can only be the destruction of free energy.

In addition, again with perfect generality:

In such systems, we can put no restriction to the direction of individual transformations. Transformations in which entropy decreases, i.e. reactions against the natural direction defined by the second law, are perfectly possible if supported by transformations in the natural direction. What we can say is that at least one transformation must proceed in the natural direction and that the effect of the total pattern of transformations must result in the production of entropy.

The phenomenon we phrase above, henceforth called coupling, i.e. combining a process in the natural direction with one against this natural direction, shows of crucial importance to the understanding of evolutionary phenomena in chemistry, biology and economics.

The restriction to the reaction pattern applies to any system be it isolated or open. Hence, it also applies to the universe. The direction of its evolution involves an irreversible destruction of free energy that sets the time arrow of evolution.

For a system in steady state, we formulate a restriction that does not depend on the complexity of the transformations in the system. We analyze the constraint based on a black box approach. We police the boundaries of the system and simply apply free energy accounting to the flows going in and out. This usually leads to a vast reduction of the complexity of the analysis. The present author (Roels (1980, 1983)) applies this principle to functioning microorganism that engage in very complex patterns of chemical reactions to build up their structure and to produce metabolic energy to support their maintenance and growth. By the very nature of their successful functioning, the significant exchange flows with the environment are vastly smaller in number than the myriad of compounds supporting their metabolism. This observation leads to powerful tools for the analysis and optimization of processes in which microorganisms are used.

For systems in a steady state the restriction due to the combined first and second laws reads:

For a system in steady state, the perfectly general conclusion is that the net effect of the exchange flows with the environment must be transport of free energy to the system. Again, we can say nothing about individual flows.

For a complete understanding of the nature of the restrictions introduced above the concept of free energy that we left unexplained, needs further analysis. This is the subject of the next section.

3.5. Free energy.

Free energy is a powerful concept. It arises from combining of the first law, the second law and the concept of temperature. The definition of free energy is relatively straightforward. Firstly, it is a macroscopic function of state and it thus solely depends on the present state of the system. The trajectory the system follows in arriving in that state, its history, is immaterial. Let us assume that the system is originally in the reference state for energy, i.e. the arbitrary datum level for energy, and let us additionally assume that the entropy is initially also zero. We further assume that the system travels from the initial to the final state by reversibly exchanging heat with the environment. In that case, following the reasoning in the Section 3.3 that leads to the identification of the relation between entropy increase and heat transferred to the system, we write:

$$\frac{dS}{dt}=\frac{\Phi_Q}{T} \tag{3.22}$$

When we phrase this a little bit casually, we can identify the product of temperature and the change of entropy, and hence the product of temperature and entropy as we started from a reference state in which the entropy was zero, as the low quality, heat related, portion of the system's total energy. Thus, it seems reasonable to diminish the system energy with the low quality portion defined by the product of temperature and entropy. In this way, we define free energy G as the difference between internal energy and the product of temperature and entropy. The author realizes that the reasoning above is far from being rigorous but it conforms to the result of a full mathematical analysis and may be of use to the less mathematically inclined reader. Anyhow, we arrive at free energy using the following expression:

$$G=E-TS \tag{3.23}$$

Formally, eqn. 3.23 is the definition of the Helmholtz and not the Gibbs flavor of free energy but that is immaterial to the treatment here. We now suspect that G is an expression of the portion of the available energy that we can transform into useful work, i.e. is interesting from the perspective of "relevant" energy. Total internal energy thus not directly relates to the real available capacity to obtain work. The restrictions we derive in the previous section put a limitation to transformation and transfer of relevant energy or free energy. In Chapter 4 when discussing the relation between the macroscopic and microscopic approaches, we provide a clearer explanation of the difference between energy and free energy.

We summarize the root cause for the difference between potentially available internal energy and the fact that only a portion is accessible to do useful work as follows. We resort to a macroscopic description, as the microscopic picture of reality is far too complex to analyze in detail. Therefore, we seek refuge to an approach that involves far less state variables by lumping together many microstates in one macroscopic state. In the associated averaging process, information gets lost. Entropy is the state variable that quantifies that loss of information. Entropy is proportional to the amount of information lost in the averaging process (an amount of information can be quantified using information theory as we show in Chapter 4). The proportionality constant is the famous Boltzmann constant indicated with the symbol k. It is a mishap of the history of thermodynamics to use the Celsius degree as the unit for temperature. On choosing the product of the Boltzmann constant and our present-day temperature as the temperature scale, equality results between entropy and lack of information on using the macroscopic approach. This reveals that the product of temperature and lacking amount of information represents the energy penalty resulting from the macroscopic approach. That penalty results in a portion of the internal energy becoming “hidden” by the lack of information in the macroscopic approach to reality and hence that portion is unavailable to do useful work.

Entropy is less a property of the system than a property of the model of reality we can realistically develop. Our picture is always of a reduced information nature and we cannot avoid paying the penalty that this involves.

3.6. Conclusion.

This chapter summarizes the basic framework of macroscopic thermodynamics. It introduces the concepts of internal energy, entropy and free energy, the first and second laws of thermodynamics and the nature of temperature as a so-called integrating factor for the heat flow. We discuss the relation of temperature to the cost of the information that is lacking in the macroscopic approach. These principles result in restrictions to transformations and the associated flows of work, heat and energy bearing substances to and from the systems’ environment.

CHAPTER 4. MACROSCOPIC AND MICROSCOPIC MODELS: STATISTICAL THERMODYNAMICS.

4.1. Introduction.

We discussed the vast complexity of almost all physically and socioeconomically relevant systems. A detailed modeling of the microscopic picture of reality is not practical. That is why physicists resort to macroscopic models containing vastly less information than needed to specify the microscopic richness of the details of a system. These methods are one of the hallmarks of modern science and an important driver of scientific and technological progress. This chapter unveils the relation between the macroscopic and the microscopic approaches by the presentation of a statistical analysis. This allows the reader to obtain a better understanding of the power and the limitations of the macroscopic approach. To this effect, we need a background in statistics and information theory. This chapter treats these formalisms as well as the foundations of a statistical theory of energy transformations.

4.2. Statistics and probability theory.

Statistics is a largely empirical branch of science. We observe a property, or a number of properties, of a collection of objects called the ensemble. We develop a picture of the characteristics of the population of objects in the ensemble by calculation of the probability that the properties have a defined value. As an example, we collect data on the age in years of all people on earth. The earth's population is the ensemble for statistical analysis. We count the number of people with a given age and divide by the total number of people on earth to arrive at the probability of that age. As you see this does not involve rocket science, it just is very tedious if you singlehandedly have to do this for all the people on earth. You go on and use the probabilities to calculate the average age of the population on earth. This comes closer to a theoretical consideration, but it is a rather elementary exercise. Calculating the average involves multiplying the probability of a certain age with the age and adding the resulting numbers. Of course, this average of the age of the people on earth, a macroscopic quantity, contains far less information than the probability distribution or probability density function of age. Average age is an example of a reduced information macroscopic quantity describing the age of the world's population at the macroscopic level. It is useful for some purposes but not for others. Identifying the number of people that are above or below a prescribed age needs the probability distribution and the average age is of little use.

We discuss the macroscopic approach and its reduced information characteristics in Chapters 2 and 3. As our example regarding selection of age categories shows, the macroscopic variable average age is of little use when we want to select an age category according to a range of values of that property. This proves to be an important observation concerning the limitations of the macroscopic method. As we show later, selection mechanisms that introduce a preference for objects with given characteristics, are one of the main drivers behind Darwinian evolution. This leads us to suspect that he macroscopic method has some limitations in systems where selection favors objects with non-average properties. We conclusively show later on that this is indeed the case. It is not the only problem in attempts to predict behavior in complex systems but it certainly is important and in its own right already prohibitive.

We already noted that the entropy of the macroscopic description reflects the uncertainty that remains about the exact properties of the objects, if we only know the average value of those properties. In our example, the statistical entropy defines the amount of information that is lacking to pinpoint the exact age of a specific member of the population. In our further

discussion about measuring amounts of information, we introduce the concept of the statistical entropy of a probability distribution as a reflection of the information that is lacking in a macroscopic approach and show its relation to the thermodynamic concept entropy.
In statistics, the average value of a property of an object is only one of the relevant numbers used to characterize a probability distribution function. Another frequently used concept is the so-called standard deviation of the probability distribution function. The standard deviation is the square root of the average of the squared differences between the properties of objects and the ensemble average of that property. We illustrate the concept of the standard deviation using the example of a ubiquitous probability distribution function, the normal or Gaussian distribution function. The term normal reflects that the statistics of properties frequently obey such a distribution function. Later on, we analyze the root causes for this proliferation of the normal distribution function. Fig. 4.1 depicts the normal probability distribution; it plots the probability of the property on the y-axis against the difference of the property of the object and the average value, expressed in units of the standard deviation of the probability distribution, on the x-axis.

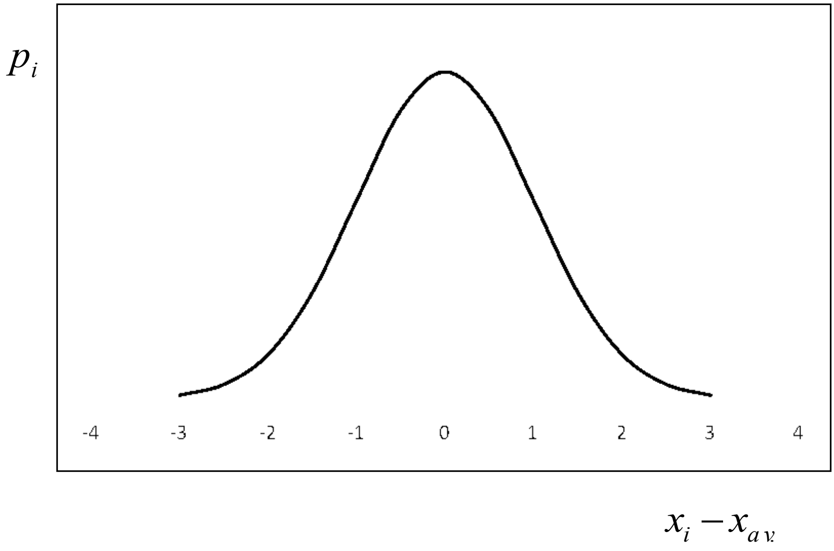

Fig. 4.1. The normal distribution function.

The normal distribution is bell shaped. The standard deviation defines the width of the bell. As it happens, there is a 95% probability of the objects' property having a value in the range between the average plus or minus two times the standard deviation. As standard texts (Roels (2010)) show, the statistical entropy of the normal distribution is equal to the logarithm of its standard deviation. In fact, the energies of the movement of molecules in an ideal gas follow, according to the Maxwell-Boltzmann distribution for the velocities of molecules in an ideal gas, a normal distribution with the standard deviation increasing with increasing temperature. The bell shaped curve puts a perspective on the macroscopic method that only considers the average, if selection favors properties at the tails of the distribution function that have a low probability of occurrence but become dominant by the magnifying action of the selection process. This becomes particularly important if the number of objects is large. In that case, a highly improbable object property may increase in dominance due to the selection process.

Before embarking on a more detailed analysis of the characteristics of probability distributions, averages and standard deviations, we discuss the difference between probability theory and statistics. These branches of science are similar but different. Firstly, we look at the historical dimension. What causes statistics and probability theory to emerge? The mathematical theory of probability traces back to an old problem: Is it possible to predict the outcome of games to be able to realize economic gain in such games? The development of these approaches traces back to the sixteenth century and further develops in the 17th century. Notably Pascal, Fermat and Huygens contribute. The objective of probability theory is to predict the likely outcome of games or events subject to uncertainty.

The origin of statistics in the 17th century goes back to the prediction of mortality rates to support insurance practices and the need of governments to predict the development of populations, i.e. demographic applications. Statistics focuses on the question what is going to happen, not why it is happening. Probability theory tries to predict what is going to happen by analyzing why things are happening. We touch on the essence of the difference between the drivers of industrial research and academic science.

We illustrate the difference by a simple example. Consider a die with six squares numbered from one to six, i.e. an apparently proper die. The objective is predicting the average of the numbers that appear on throwing the die a large, in fact an infinite, number of times. The statistician starts throwing the die noting the numbers that consecutively appear. He faces a problem. He must find the average after an infinite number of throws. Of course, he cannot throw the die an infinite number of times, because that takes forever. The statistician being a clever chap starts calculating the average after every throw. Soon he sees that the average closes in on 3.5 and after a time fitting his liking of accuracy, he states that the average is 3.5. In itself, it is remarkable that a completely random process results, from the macroscopic perspective, i.e. in terms of the average outcome, in a completely predictable result.

The adept of probability theory takes a different approach. After some pondering, he concludes that, for a perfect die, the probability of one of the faces coming up is equal and hence it is one sixth for each of the faces. He quickly calculates the average to be 3.5. There is another difference between the two approaches that needs discussion. The statistician may come up with a number of 3.6 and he confronts the probability theory gay with his finding. This chap, as he is a scientist, tells the statistician that he is wrong and that he probably did not throw the die a sufficiently large number of times. The statistician goes on throwing the die but continues to come up with 3.6. Who is finally right? The answer lies in the fact that the truth of the matter always lies in verifying a theory experimentally. Also in science, the proof of the pudding is in the eating. The problem rests on an assumption that the theoretician sneaked in, probably without even realizing that he made a crucial assumption. The theoretician assumed that the die was perfect and that indeed each of the squares had equal probability of coming up. The empirical approach shows that this is not the case; at least one of the higher numbered squares must have a higher probability to appear if we want to explain the empirical result. Empirical evidence leads to the conclusion that the die is not perfect.

We can derive more from the empirical data of our statistician. He discovered the Central Limit Theorem, which states that the average of the outcome of random events, such as throwing a die, converges on an average that obeys the normal probability distribution. In addition, the standard deviation of the average, if we repeat the process of consecutive throws many times, decreases with the number of throws in the sample. In fact, it decreases with a factor equal to the reciprocal of the square root of the number of throws. This means that an outcome equal to the average becomes extremely probable if the number of throws is large. This is comforting if we look at the problems we face in physics, where we use the average kinetic energy of an extremely large number of molecules to calculate the macroscopic quantity temperature. It is extremely likely that the average is the measured value we observe using a thermometer. The statistician also discovered the law of large numbers. If he analyzes his list of data carefully and calculates the probability of each face coming up he will find a value of one sixth for each of the faces if he takes a large enough sample.

There are further facts we deduce from the experiment with the die. The first one is at face value a rather peculiar one. If we throw the die say ten times, the chance of 10 sixes coming up is very small, but in fact it is equally probable than any other specified sequence. At first, this may seem counterintuitive but it is a fact. The problem rest in the illusion of order in the sequence of 10 sixes. It is difficult to accept that this beautifully "ordered" set comes out by pure chance. In fact, every other specific sequence is equally ordered. The second point we make is that even if a sequence of ten sixes appears the likelihood that the next throw will be a six stays one sixth. This apparently conflicts with the fact that the averaged throws in the end result in 3.5. Therefore, a significantly lower number of sixes will have to come up in future throws. That is certainly a truism, but still the outcomes of the past have no influence on the next throw's outcome.

Another important feature of the die example provides clues about the nature of time and the nature of the second law that provides an arrow to time, i.e. time always moves to the future it never reverses. This is apparent conflict with the fact that the laws of motion show time reversibility, i.e. if we substitute minus time in the equations of motion a particle exactly traces back the trajectory to its initial position. We know that this never happens in practice, it is at least very improbable. Are the laws of mechanics wrong or are we missing a point? The answer rests in the nature of our probabilistic information about the system, i.e. it lies at the very heart of the fact that we have limited information. This limited information leads to the observation that prior to an event we have only a probability that an event will happen. We only know on the average what will happen. This notion breaks down when the event happens and we are completely sure about the outcome. Time reversal includes erasing that information and that is impossible once we have it. This is an admittedly sloppy, but most probably correct, way to explain the nature of the arrow of time; it reflects the probabilistic nature of our knowledge about future behavior and our certainty about past behavior.

We close this section by returning to the bell shaped normal probability distribution function. From the perspective of probability theory this bell shape derives from the fact that the outcome of a complex process often depends on a large number of random events that all have an incremental positive or negative effect on the outcome. The probability that we reach a certain age depends on staying alive a large number of seconds and surviving all the events that happen in this multitude of seconds. If we do the math this indeed shows to lead to the expression that defines the normal distribution function. This result also comes up in a number of other situations. If we consider the equations governing diffusion of molecules and heat transport, it shows that these are mathematical equivalents of the normal distribution. This is not a mere analogy; the processes of diffusion and heat transport arise from the same principle: A sequence of in principle random consecutive steps. This is the famous random walk principle. If we observe a person standing at a lamppost who is so drunk that the consecutive steps he takes are in a randomized direction, the probability of finding him at a given distance of the lamppost follows the normal distribution. Random behavior at the microscopic level often results in a very predictable outcome at the macroscopic level. This provides the ultimate justification of the application of macroscopic approaches.

4.3. Information theory: Quantifying information.

We start out analyzing a classical problem. Consider a pile of eight coins. The coins are identical, except for the fact that one weighs more than the others that are of equal weight. A simple weighing instrument is available. It has two scales and allows us to judge, which of the scales contains the heavier object. In fact, the scales allow us the make one binary decision, i.e. it allows choosing between to equally likely alternatives. We define the amount of information needed to make such a binary decision as the unit of information. Information theory defines this amount of information as one bit, one binary information unit. We divide the coins in two groups of four and conclude that the heavier coin is equally likely in the first or the second group of four. The scales allow us to identify which of the two groups of four contains the heavier coin, i.e. one weighing allows us to gather one bit of information. We repeat the procedure on the group of coins we identified as containing the heavier coin, i.e. we divide the four coins into two groups of two. We obtain another bit of information and ascertain in which group of two the heavier coin is located. We now have two coins of which one is the heavier coin. We gather one final bit by weighing the two remaining coins and we have identified the heavier one. We used three bits of information. This makes sense because two to the third power happens to be eight. This leads to the conclusion that the amount of information in bits equals the base 2 logarithm of eight, or, if we use the natural logarithm, it equals the natural logarithm of eight divided by the natural logarithm of 2. Standard texts on information theory (Shannon and Weaver (1949), Brillouin (1962)) generalize the discussion as follows. Each binary information unit allows the selection of one outcome out of two equally probable ones. Thus, the following formula defines the amount of information, I, needed to select one specific object out of an ensemble of W equally likely candidates:

$$I = \frac{\ln W}{\ln 2} \tag{4.1.}$$

To be in contact with thermodynamics we suppose the ensemble contains W microstates that are consistent with the macroscopic state of the system and for the time being we assume that these microstates are equally probable. In that case, eqn. 4.1 represents the statistical entropy of the macroscopic picture of the system. The next step is to formulate the statistical entropy in terms of the probability distribution of the microstates. If the probabilities of the microstates are indeed equal, the probability $p(i)$ of each of the W microstates is the reciprocal of the number of microstates. In that case, it follows that $\ln W$ equals $-\ln p(i)$ .We generalize this to the case in which the probabilities of the microstates are different. In that case, we have to substitute the average value of the logarithm of the probability of the microstates:

$$I = -\frac{< \ln p(i) >}{\ln 2} \tag{4.2.}$$

In eqn. 4.2, we use the shorthand notation <quantity> for the average value of a quantity. As we already mentioned a calculation of the statistical entropy of the normal distribution leads to the conclusion that its statistical entropy is the natural logarithm of the standard deviation of the probability distribution.

The material presented above summarizes everything that we need to know about information theory for the purpose of this book.

4.4. Statistical thermodynamics.

Consider a macroscopic system. It can exist in a large number of microstates that comply with the macroscopic information we have available. We assume the probability density function known as well as the energy E_i of each microstate. We straightforwardly obtain the average energy by summing up all the energies of the microstates weighted according to their probability. This average is the value of the energy we most likely observe macroscopically. To be in contact with thermodynamics we simply denote that average as the macroscopic energy *E*.

The problem of statistical thermodynamics is now to find the distribution of E_i over the states that are consistent with the macroscopic information we have about the system (Hill (1960), Andrews (1975)). In statistical thermodynamics, we assume that the probability of a microstate uniquely depends on the energy of that microstate. This is a basic postulate and rests on the circumstance that the energy of the microstate is the only distinctive feature of that state from the perspective of the macroscopic description. To assume anything else than sole dependency on the energy of the microstate assumes availability of information we do not have and this introduces undue bias. We do not present the mathematical detail but only cite the major results. The literature cited provides ample access to the mathematics, Roels (2010) provides a summary.

The statistical analysis reveals that the change of the macroscopic energy of a system has two contributions:

- The first one represents the change of the macroscopic energy due performing work on the system. We discuss this as the starting point for defining energy in Chapter 3.
- The second one represents a change in the information the observer has about the probabilities of the microstates; it stands for the effect of the changes in the probability density function. We identify the second term as the contribution due to "information work" performed on the system. It represents the change in the information we have about the system's microstate.

In Chapter 3, eqn. 3.1, we formulated an expression for the change of the macroscopic energy. It contains a contribution due to work performed on the system and a term due to heat exchanged with the environment. For convenience sake, we reproduce eqn. 3.1 here.

$$\frac{dE}{dt} = \Phi_P + \Phi_Q \tag{3.1}$$

The first term at the right hand side of this equation corresponds to the first bullet point above, it represent the exchange of energy with the environment due to work performed on the collective microstates of the system.

The second bullet point above thus has to represent the exchange of heat. It represents changes in the information we have about the microstate of the system as reflected in changes of the probability density function. Heat is a statistical rather than a mechanical concept. Heat emerges as a direct consequence of the information that is lacking in a macroscopic approach to reality.

The analysis leads, using eqn. 4.2., straightforwardly to the following expression for the exchange of heat between the system and the environment:

$$\Phi_Q = -\frac{\dfrac{d < \ln p_i >}{dt}}{\beta \ln 2} \tag{4.3}$$

We discuss the nature of the constant β later.

Combination of eqns. 4.3 and 4.2, leads to the following equality:

$$\Phi_Q = (\frac{1}{\beta})\frac{dI}{dt} \tag{4.4}$$

Eqn. 4.4 relates the heat flow to the rate of change of the lacking information in our reduced information macroscopic picture of reality. It again highlights the statistical nature of heat. As the next step, we return to eqn. 3.7, the expression for the change of the system's entropy:

$$\frac{dS}{dt} = \frac{\Phi_Q}{T} \tag*{(3.7.)}$$

Combining eqns 3.7 and 4.4 results in:

$$\frac{dS}{dt} = (\frac{1}{\beta T})\frac{dI}{dt} \tag*{(4.5.)}$$

If we integrate both sides of eqn. 4.5 starting from a reference state that we attribute zero entropy and zero amount of lacking information, we obtain the following relation:

$$S = \frac{I}{\beta T} \tag{4.6}$$

This expression confirms the proportionality between entropy and lack of information.
A fundamental result of statistical thermodynamics shows that the constant β in eqn. 4.6 follows as:

$$\beta = \frac{1}{kT} \tag{4.7}$$

Where k is the Boltzmann constant.
Combining eqns. 4.7 and 4.6 leads to the famous equation that originates from the work of Boltzmann, it relates entropy and lacking information by the relation $S=kI$.
The literature (see e.g. Brillouin (1962)) debates the nature of the factor β. kT is equal to the energy cost of one bit of information; at least it puts a minimum to the energy cost to obtain information beyond the macroscopic information. Careful analysis of e.g. the concept of the "Maxwell demon" , a genie assumed able to avoid paying the cost of information and hence escaping the constraints due to the second law, leads to the conclusion that there exists no way to avoid paying this minimum cost. This exorcises the demon effectively (Brillouin (1962)).
To close this section we remind the reader again of the nature of the concept of statistical entropy that underlies the entropy state function as used in thermodynamics. The statistical entropy, a term we frequently use in this work, measures the information that our model of a system lacks if we want to pinpoint the microstate, i.e. the description representing the microscopic details of the system. The statistical entropy provides a quantification of the lack of information.

4.5. Conclusion.

In summary, we reach the following conclusions in this chapter. Eqn. 4.6 presents the statistical interpretation of entropy as introduced in macroscopic thermodynamics. It ties entropy to the limitations of the information about the exact microstate the system is in due to having only a reduced information picture in its macroscopic description. This limited information results in the distinction between macroscopic intrinsic energy and free energy as introduced in Chapter 3:

$$G = E - TS \tag{4.8}$$

Combining eqns. 4.8, 4.7 and 4.6 results in the statistical interpretation of free energy:

$$G = E - kTI \tag{4.9}$$

The rationale behind eqn. 4.9 develops as follows. As we have a reduced information picture of reality, a restriction exists to the useful work that can result from a given quantity of potentially available energy. We pay a penalty for that lack of information; this penalty equals the product of the amount of information that is lacking and the value, in energy units, of that information. An in depth study of the nature of that value (Brillouin (1962)) shows that we cannot avoid paying that penalty. Even if we use the most clever way of gathering additional information about the exact microstate of the system, in order to free up part of the additional “true energy“ of the system, we pay an energy cost which exceeds its value in energy units or is at best equal to its value in energy units.

By combining this interpretation with the second law, we see that this reduced information also sets the direction that natural processes take. Natural processes proceed in the direction of decreasing information or increasing entropy, at least in isolated systems. Alternatively phrased, if we combine the system and its environment (the environment being the rest of the universe outside the system) the information naturally decreases and entropy increases in any possible process. This means that the statistical meaning of the second law is very logical indeed. In any natural process, i.e. if there is no further interaction between the system and the environment, the system will evolve in the direction of the state that is the most likely one in view of the limited information we have; assuming anything else rests on unfounded bias.

CHAPTER 5. A GENERAL THEORY OF TRANSFORMATIONS AND ECONOMIC TRANSACTIONS.

5.1. Introduction.

In Chapter 3, we introduce macroscopic thermodynamics. Here we investigate the extension of this theory to include economic transaction. In this way, we arrive at a theory to describe transactions and other processes that appear in our economic system. We henceforth use the name economic value theory, shorthanded EVT, to indicate this theory. In addition, we attempt to merge thermodynamics and EVT into one body of theory that encompasses both economic processes and processes in the purely physical sense. Is it possible to combine thermodynamics and Economic Value Theory (EVT) in one theory? We introduced this theory in earlier work (Roels (2010)), where we termed it Value Transaction Theory but we argue that the term Economic Value Theory reflects the nature of the theory in a better way. In fact, classical thermodynamics hardly deserves a name in which dynamics appears. Thermodynamics is, in its classical sense, an equilibrium theory; hence, there is nothing dynamic about it. It introduces infinitely slow reversible changes and that is not very dynamic at all. We consider dynamic processes in irreversible thermodynamics, a subject that we introduce in this chapter. Classical thermodynamics has an unfortunate name, as it is a static theory and is not limited to processes involving only thermal effects as the adjective “thermo” leads to believe.

To investigate the formulation of an extended theory we return to the basic concepts of thermodynamics, i.e. energy, entropy and temperature and the laws of thermodynamics. In addition, we return to the heat engine that spawned thermodynamics following Carnot’s analysis of that contraption 185 years ago. In the heat engine, we find the inspiration to generalize our analysis.

5.2. A general Economic Value Theory.

This section introduces a rather bold line of reasoning. We extend the theories of macroscopic thermodynamics to functions of state termed potential value, or in shorthand value, and economic value respectively. These are the counterparts of energy, potential value, and free energy, economic value, as these feature in classical thermodynamics (Chapter 3). In this book, we use the term economic value instead of free value, a term we use in earlier work (Roels (2010)). The definitions are straightforward if we assume equivalency between economic work and thermodynamic work. In that case, potential value becomes the potential capacity to generate economic value. It is the analogue of internal energy in thermodynamics. Economic value is the analogue of free energy. Economic value is potential value corrected for the statistical entropy of the macroscopic picture of reality that we adopt in view of the overwhelming complexity of the system at the microscopic level. Economic value is the actually available capacity to do economically relevant work based on our reduced information picture. Thus, we obtain complete symmetry between thermodynamics and our general theory of economic value, EVT. We propose a macroscopic theory of economic value; we do not consider the microscopic complexity. The symbol W indicates potential value. It is the counterpart of the concept of energy in thermodynamics. For economic value, we use the symbol G to remind us of the thermodynamic counterpart free energy. We assume that we have an objective unit of value, just as in thermodynamics where we use gravitational work, i.e. the movement of a mass in the gravitational field, to define the units and the equivalence of different forms of work.

We straightforwardly proceed with the formulation of the first law of EVT. It defines the conservation of potential value. Its production or destruction in processes in a system in an economy is zero. Henceforth, following terms used in economic reality, we use the term transactions for all economically relevant processes. We introduce the following expression for the first law of EVT for a system part of the economic reality:

$$\frac{dW}{dt} = \Phi_W \tag{5.1}$$

Although the notation is perhaps already familiar, we reiterate that the term at the left hand side stands for the rate of change of the amount of value, the term at the right hand side is the increase of value due to exchange with the environment. At the right hand side, no contribution due to transactions in the system appears by virtue of the conservation of value.

Considering potential value conserved may be confusing as it implies that its creation or destruction does not take place economic transactions. This may conflict with the illusion that we create value in our economy. Explaining this feature involves returning to energy and thermodynamics. In addition, we move to some 13.5 billion years ago when, according to the Big-Bang theory of its creation, the universe instantaneously appears with all the energy that it has today. By virtue of the first law the universe's internal energy today cannot be anything else than equal to that when the universe emerges. The first law does not allow a change as the universe is an isolated system. The energy in the universe fueled the emergence of the socioeconomic system, as we argue later, and the potential value that exists in the universe is constant and equal to the amount of energy present in the universe when it emerges. Earth with is socioeconomic system is an example of an open system and the solar radiation is to a good approximation the only source of resources on earth. The solar radiation is the product of the minute part of the energy present in the universe that constitutes the sun. If we think that we observe the creation of value, we see transformation of potential value into economic value that is available to perform economic work. We see the creation of economic value out of the potential value that emerges with the creation of the universe.

We can complete the conservation equation according to the first law of EVT if we formulate an expression for the exchange of value with the environment. It turns out that we have three types of contributions to the exchange of value:

$$\Phi_W = -\Phi_I - \Phi_P + \Phi_1 W_1 + \Phi_2 W_2 \tag{5.2}$$

The first term at the right hand side is the equivalent of heat in thermodynamics; in EVT, it represents a flow of information (in fact information that is lacking in the macroscopic description). In doing this we follow the formalism of statistical thermodynamics introduced in Chapter 4 where we revealed the informational nature of heat. The second flow represents the equivalent of work in thermodynamics. In EVT, we can think of liquid capital and other assets of which we know the economic value with certainty. The third part represents the sum of the values of the assets of which we only know the expected value subject to uncertainty that is reflected in the EVT equivalent of entropy, i.e. the information that is lacking to complete specify the value microstate the asset is in. These assets can be stocks and bonds or industrial and consumer products that we sell or buy. We again limit the number of such assets to two as this serves to illustrate the properties of more complex situations in which many assets appear.

A new feature in EVT is economic work to complement the purely physical sources of work such as mechanical, electrical and chemical work. It is worthwhile to explore the nature of economic work somewhat further. My thoughts on this subject are still incomplete so part of the remarks made here are speculative in nature. The present view of this author is that at least an important contribution to economic work is information work. It reflects the ability to gather information and hence increase the economic value of a given amount of potential value. We return to this concept in Chapter 7 when discussing the nature of the firm.

Taking the next step in the development of EVT involves introducing the second law of EVT. It will not come as a surprise that this law states that every economic transaction results in the destruction of information reflected by production of the EVT equivalent of entropy. The information that is lacking to specify the exact value microstate of everything in our system can only increase. This leads to the following macroscopic balance equation for that lacking information:

$$\frac{dI}{dt} = \Pi_I - \frac{\Phi_I}{C_I} + \Phi_1 I_1 + \Phi_2 I_2 \qquad (5.3)$$

The term at the left hand side stands for the rate of change of the amount of lacking information. The first term at the right hand side of eqn. 5.3 stands for the creation of uncertainty or the destruction if information that must be positive due to the second law. The second term at the right hand side is the equivalent of the contribution of heat in thermodynamics. We stressed that temperature in thermodynamics is of an informational nature, the product of Boltzmann's constant and temperature stands for the minimum of the energy cost for obtaining one unit of information about the microstate of the system beyond the information contained in the macroscopic description (Chapter 4). The symbol C_I is the EVT equivalent of the product of the Boltzmann constant and temperature; it represents the cost of additional information about the value microstate of the system.

The economic value follows, in analogy with free energy in thermodynamics, as:

$$G = W - C_I I \qquad (5.4)$$

The interpretation of eqn. 5.4 runs as follows. Any economic asset has a macroscopic potential value or in shorthand a macroscopic value. That value is conserved in every possible transaction. We only know the expected macroscopic value and the asset can exist in many value microstates. To specify the exact value microstate in view of the uncertainty contained in our macroscopic picture, we need information beyond that contained in that macroscopic picture, its amount is *I.* That information is not a free commodity it comes at a cost of C_I units of value per unit of information. Hence, the economic value is obtained if the product of that cost and the amount of lacking information is subtracted from the expected macroscopic value. In fact, that cost is a minimum. In general, the value expenditure to obtain the information is higher.

By reasoning analogous to the thermodynamic discussion in Section 3.4, we arrive at the general restriction due to the EVT formalism:

In any economic system, the net effect of a pattern of transactions can only be the destruction of economic value.

In addition with perfect generality:

In an economic system, there are no restrictions to the direction of individual transactions. Activities in which information increases, i.e. against the natural direction set by the second law, are perfectly possible if transactions in the natural direction of sufficient magnitude support these. All we can say is that at least one transaction must proceed in the natural *direction and that the overall effect of all the transactions in the system must be a decrease in information.*

We termed this coupling in the thermodynamic treatment. Coupling shows of crucial importance to the understanding of evolutionary phenomena in any system, including the socioeconomic system.
The restrictions apply to any system be it isolated or open. Hence, it also applies to the universe; the direction of its evolution involves an irreversible destruction of economic value that also sets the time arrow of evolution of the universe and the socioeconomic system.
For a system in steady state, we can formulate a restriction that does not depend on the complexity of the transactions in the system. Just as in thermodynamics, we can analyze the constraint based on a black box approach. We can police the boundaries of the system and simply apply economic value accounting to the flows going in and out. This leads to a vast reduction of the complexity of the analysis.

For a system in a steady state outside equilibrium, the perfectly general conclusion is that the net effect of the exchange flows with the environment must be transport of economic value to the system.

This completes the development of the EVT formalism. We apply it to the analysis of economic transactions later in this work. This formalism also allows us the derive the force that drives evolution in socioeconomic systems just as the thermodynamic forces that drive evolution in physical system, such as in the example of heat transfer that we discussed earlier.

5.3. The Capital Asset Pricing Model and Economic Value Theory.

The theory of finance presents the Capital Asset Pricing Model (CAPM) as a decision-making tool in investments subject to uncertainty. It is an example of a macroscopic approach that analyzes a complex decision-making problem in terms of a reduced information description. Here we analyze the differences and correspondences between EVT and CAPM. We start our analysis with a brief outline of the CAPM.
A number of excellent texts on the subject (e.g. Sharpe (1970), Elton and Gruber (1984)) summarize the Capital Asset Pricing Model. This section explains the main features of this model and compares it to EVT. CAPM presents an equilibrium model of markets for capital assets, for example stocks and bonds and liquid capital. It is a simple model for these complex markets. Here we discuss the standard capital asset pricing model; it is the original and most transparent version of the model.
The development of CAPM rests on a number of simplifying assumptions. We outline some of these assumptions here, for a more complete account we refer to Elton and Gruber (1984). Firstly, we assume that there are no complications due to income tax and transaction costs, these are negligible or ignored in the treatment. We further assume that buying and selling actions of individual buyers and sellers do not affect the price of an asset. The total community of buyers and sellers determines the pricing of assets. This is the perfect competition assumption that we discuss in Section 8.8. The decision making of buyers and sellers solely derives from the asset's average value and the uncertainty about this average value. The actors in the market see the probability density function of the past returns on an asset as an estimate of the probability distribution of the returns they expect in the future. A further assumption is that these returns follow the normal probability distribution. The theory assumes the expected value of the returns equal to the average of the returns in the past and the standard deviation of the past returns

representative for the standard deviation of the returns in the future. The standard deviation of the past returns defines the uncertainty whether the future returns indeed materialize as the average value of the past returns. These seem prudent assumptions in absence of information to assume something else. The reader will start to recognize this approach as a characteristic of macroscopic modeling. You always have to assume that the outcome that materializes is the one most probable in view of the information available. In fact, the statement reflects the essence of the second law. If you isolate a system from the environment, the system always evolves, at least from the macroscopic perspective, to the microscopic configuration most likely based on the limited information available. Evolution to a maximum of the statistical entropy leads to a situation in which all allowable microstates, i.e. the microstates that comply with the initial information, become equally probable. The ability to predict the microstate it actually finds itself in has decreased to an absolute minimum. Casual formulations of the second law often state that a system evolves to a situation of maximum disorder. This is an ambiguous statement: It is more a definition of order than an interpretation of the meaning of the second law. It derives its attractiveness from it compliance with aspects of daily experience. Buildings, if left alone, turn into ruins; the reverse process never takes place spontaneously. So far this brief excursion, we return to CAPM.

In addition to the foregoing, we assume that investors are greedy; they prefer a higher return to a lower one. Everything left the same the investors prefer more value to less value. To me this seems of fair summary of at least some aspects of human nature. A further assumption is that in investment deals that lead to the same expected return, investors tend to prefer investments with the lowest uncertainty of the returns. This is the risk aversion hypothesis. In absence of anything better, the CAPM approach suggests this uncertainty to relate to the standard deviation of past returns. This rings a bell! In the EVT formalism, we quantified the lack of information, a scientific phrasing of the word uncertainty, as the statistical entropy of the distribution of value. In addition, we stated that, for the normal distribution, the statistical entropy equals the natural logarithm of the standard deviation. Here CAPM and EVT seem to agree about the fact that the standard deviation of the distribution function is a measure of uncertainty. The approaches disagree about the exact mathematical shape of that relationship. In earlier work (Roels (2010)) we show that for investments with an risk that is close to the average risk of all investment opportunities in the market, the two approaches lead to uncertainty estimates that are, also quantitatively, almost the same. Furthermore, CAPM assumes that the investors can unlimitedly lend or borrow liquid investment capital at the so-called risk free interest rate. CAPM attributes this risk free rate a non zero value but assuming a value of zero, which comes very close to the 2011 economic reality, does not affect the line of reasoning. Hence, for simplicity's sake, we again choose the easy way out and assume it zero.

Under these assumptions, CAPM shows (Elton and Gruber (1984), Chapter 11) that all investors end up with portfolios of assets that lay along the so-called capital market line. At least when they have the same information about the probability distribution of past returns and accept this distribution as representative for the distribution of future returns. This line characterizes efficient portfolios with an in the eyes of the collective investors optimal tradeoff between expected return and risk or uncertainty. Indeed, the returns of the past reflect present-day expected value in the mind of the investors. However, they do not pay the expected value of the future returns, as they want a compensation for the uncertainty about the future returns as measured by their standard deviation. The collective investors do not want to pay average value, but rather average value corrected for uncertainty. This conforms to the distinction between potential value and economic value as discussed here. It seems that two different approaches, lead, for economic equilibrium, to the same conclusion about the behavior of actors in the market.

Elton and Gruber (1984) show that the capital market line, if the risk free return is zero, defines as follows:

Expected return = (compensation for uncertainty) x (standard deviation of distribution of returns)

The literature cited shows the full mathematical expression of the capital market line, it is a straight line with a slope equal to the compensation for uncertainty. We will not introduce the mathematics here as these contribute very little to the line of reasoning in this book. The reader can consult my earlier work to get access to the full story (Roels (2010)), Chapter 4).
If we perform a full analysis of the investment problem described above from the perspective of EVT we obtain the following result with perfect generality, i.e. we do not have to introduce assumptions about the probability distribution of past returns:

Expected return = (cost of information) x (statistical entropy of the distribution of returns)

The expressions in the two formalisms are very much alike indeed. We already highlighted the difference in the appreciation of uncertainty. A second point is that for CAPM the quantification of uncertainty as standard deviation comes in as an assumption without any justification. In EVT, it appears as a consequence of the theoretical development. Therefore, we conclude that EVT is at least as valid to describe behaviour in financial markets as the well-accepted CAPM approach.
We feel the need to make an additional remark on the application of models of financial markets. It involves an important methodological problem that we highlighted earlier in Section 2.5. A model is a tool to describe a situation of more or less involved complexity. It predicts features of reality beyond the data from which the model derives. A very basic assumption or requirement is independency of model and the system, i.e. the existence of the model should not influence the processes that take place in the system. This is a questionable assumption for models of a system in which human actors participate that know the model. Certainly, in a situation where the financial community accepts the predictions of the model, it influences the outcome of the transactions between the players in the market. This introduces a significant methodological problem also in independently verifying the validity of the model. It may take the disguise of a self-fulfilling prophecy. This holds for both the CAPM and the EVT approaches.

5.4. The nature of energy and value revisited.

We first need to perform a more in depth analysis of the concept of energy. We stated in Chapter 3, Section 3.2, that energy is concept allowing accounting principles that keep track of the exchange of various forms of work (including exchange of chemical substances and other energy bearing matter or radiation) and of the informational quantity heat. We cannot directly measure the energy of a system, we can only calculate it if we do careful bookkeeping of all the work performed on the system and all the heat exchanged with its environment and equate the result to the change of energy. We can do that by the convenient introduction of a first law that defines energy as only being able to change by exchange of heat and work. We argue that energy conservation is not so much a natural law but conservation is the very property that defines energy. Every time when something seems amiss with its conservation, we start looking for sources of work or heat that lead to restoration of conservation. Until today, we remain successful in insisting on conservation. In many instances, where we encountered those problems in analyzing the evolution of the universe, the scientific world got through a painstaking exercise to prevent energy conservation to fail. In fact, the jury is still out on a large portion of the energy that is apparently present in the universe but that we presently cannot identify. We refer to the concepts of dark matter and dark energy.

We need to analyze the accounting principles regarding the exchange processes that lead to changes of energy. In order for work to be possible, we need a process at a non-zero rate at the macroscopic level. A process with a macroscopic rate is only possible if non-zero macroscopic forces exist. A force is only possible if we have gradients in an intensive macroscopic state variable, e.g. pressure or temperature. Heat exchange requires a temperature difference; moving the piston in a heat engine requires a pressure difference caused by the expansion of a medium. If such gradients fail to exist, only infinitesimally slow reversible changes occur and accounting becomes a very lengthy procedure indeed.

What we propose is to include value in the economic sense in addition to physical energy in a concept of total value. To avoid confusion we propose a different symbol whenever we mean that merged sum: The Greek capital omega, Ω.

The macroscopic amount of this extended concept of value follows a conservation principle due to the first law of what we term economic value theory (EVT). This theory encompasses both energy and value in the economic sense. This procedure directly results in the need to include economic work as a flow of value in the accounting for total value. For the various flavors of work in thermodynamics, there exists the principle that these are equivalent and we can convert these in each other. In fact this holds more generally if we consider processes at a non-zero rate (and in case of a zero rate we reach a conflict with the very definition of process) where dissipation always takes place. Hence, one flavor of work never transforms completely in another one in a process with a non-zero rate.

To be complete we stress another important issue. We cannot always distinguish different flavors of work and heat in an unambiguous way. For example, magnetic work and electrical work result from the same basic force. This may also hold for economic work and the other types of work. Also for heat and work, we need an additional convention to prevent ambiguity (Roels (1983), Chapter 2). We need to be careful to use a consistent set of definitions to avoid flaws by double counting in our accounting. In fact, the forces driving these varieties of work al derive from the Big-Bang that creates the universe and drives its evolution to date.

We now write the full balance equation for our extended value concept as:

$$\frac{d\Omega}{dt} = -\Phi_I - \Phi_P + \Phi_1\Omega_1 + \Phi_2\Omega_2 \tag{5.5}$$

Most of these terms should not come as a surprise. Exceptions are the third and the fourth term at the right hand side that introduce the total value flow associated with exchange of material substances and economic assets and their associated total value rather than energy and economic value alone. For simplicity's sake, we consider only two such assets. Again, as in Chapter 3, this simplification does not lead to loss of the features of the general case of many assets. The omega's stands for the sum of value in economic sense and energy in the thermodynamic sense associated with the exchange of one unit of these substances. Furthermore, we show one term for heat and economically valuable information indicated by one flow of information in the first terms at the right hand side.

5.5. The second law of Economic Value Theory.

In Section 5.4, we present a generalization of the first law to a situation where we consider both value, in the economic sense, and energy, its thermodynamic counterpart. We can do the same for entropy and the information concept introduced in EVT. In discussing information both in the context of thermodynamics and EVT, we mean to say lacking information, i.e. the amount of information we need, in addition to the information contained in the macroscopic picture of reality, to be in a position to specify the microstate the system is in. Furthermore, we argue that the concepts of entropy and temperature are the product of a mishap in classical thermodynamics. Entropy is the product of the amount of information and the Boltzmann constant. This results from the introduction of temperature expressed in the Celsius degree based Kelvin scale. When we use energy units, i.e. the product of Boltzmann's constant and temperature, entropy would be equal to the statistical entropy of the probability distribution of microstates. We now introduce the perfectly generalized second law of natural processes considering both energy and value, as follows:

$$\frac{d\Sigma}{dt}=\Pi_{\Sigma}-\frac{\Phi_I}{C_I}+\Phi_1\Sigma_1+\Phi_2\Sigma_2 \tag{5.6}$$

Here Σ is the statistical entropy of the distribution of energy and value combined. The first term at the right hand side reflects the production of statistical entropy as the sum of the thermodynamic and the value part. The second law of EVT stipulates that it can never be negative. Any possible process or transaction or a combination thereof results in production of statistical entropy. The second term applies to the generalized concept of "heat". It contains both a thermodynamic and an EVT contribution. The last two terms account for the exchange of information due to exchange of material substances and economic assets. The sigma's represent the statistical entropies of the total value of those substances and assets.

Eqn. 5.6 accounts for information regarding the microscopic structure of physical entities as well as "non-physically bound" information that we need to specify the value microstate of an asset. This resembles the difference in concepts of DNA bound information and exogenic non-DNA bound information that defines the evolution of the socioeconomic system. We return to these concepts in Section 6.10.

We now introduce an element of speculation but the author feels safe in introducing it. The production term for statistical entropy contains two contributions. One represents physically bound information, i.e. the purely thermodynamic part, the other is socioeconomic part reflecting information not necessarily bound to physical objects. The same applies to the second term; it includes both physically bound and "socioeconomic" information. Does the extended second law necessarily imply that in the total production of statistical entropy both the energy and the value part in isolation must be positive? Defining this restriction for the two separate parts in isolation is a more stringent requirement than putting a restriction on their sum only. As we show later on, this leads to the question whether coupling exists between the forces driving the processes underlying economic evolution and those driving purely physical evolution. I am of the opinion that economic value forces can drive energy transformation processes and free energy forces can drive economic transactions. We therefore conclude that only the sum of the separate contributions to the productions of statistical entropy must be positive, one of the contributions can be lower than zero. The argument derives from the following logic. In the creation of the universe, it emerges as a purely "physical" system. Hence, purely physical forces resulted in the very emergence and evolution of the socioeconomic system in all its complexity at some moment in the evolution of the universe. Most probably even attempting a distinction between these two types of forces is futile. If such a possibility for coupling between physical and economic processes necessarily existed in the past, there are no reasons to assume its impossibility today. We conclude that only the sum of the contributions to total production of statistical entropy must be positive and we consider the very attempt to make a distinction unjustified. We find further arguments for this thesis in e.g. Chapter 6 where we discuss the emergence of exogenic evolution to complement physically bound DNA based evolution. In addition, Chapter 16 provides clues when we analyze the dependence of Gross Domestic Product on energy consumption.

5.6. The heat engine revisited and more.

Fig 5.1 shows the classical heat engine introduced in Section 3.3. High temperature heat results from e.g. a process of combustion of a fossil resource, coal in the days of the old heat engine. We recognize coal as a product of solar radiation in the record of evolution in fossil deposits. We can consider the ambient temperature of the environment as the low temperature sink. What drives the heat engine? Is it the heat derived from the combustion of fossil resources, or is there more to the story? The answer results if we introduce a constraint. We consider the heat engine and its environment as part of a larger isolated system, i.e. the combined heat engine and the environment cannot exchange energy with anything. In that case, we identify at least two problems. These both relate to the problems in the press at the moment of the writing of this book. I refer to the Climate Summit in Copenhagen of December 2009. This conference regarded both the finality of fossil resources and the heating of the earth by the greenhouse effect. Of course, a prerequisite for the heat engine to remain in operation is that the source of high temperature heat remains in place, i.e. fossils exhaustion has a "game over" effect on the heat engine if we do not avail of another source of energy to generate high temperature heat.

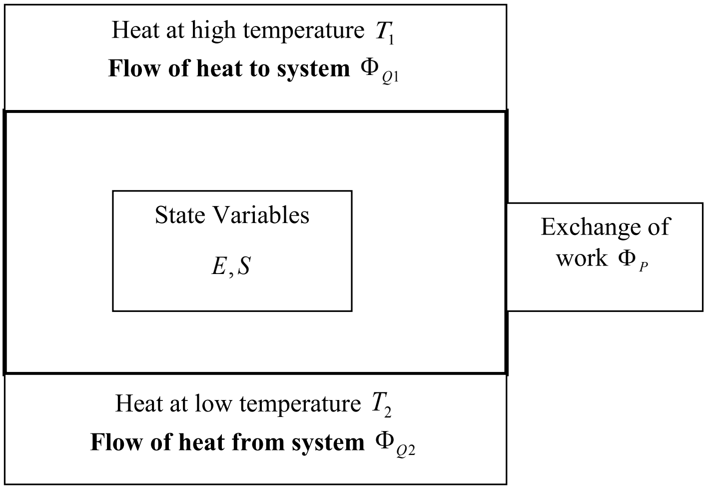

Fig. 5.1. The heat engine.

The second point is of a different nature. The heat engine only works if the heat not transformed into work transfers to the environment. Inevitably, this heats the environment, but as the earth is rather large, this involves a problem only if it is a very large heat engine. Even then, without too much interference of greenhouse gasses like carbon dioxide, the earth radiates heat into the universe; a big sink indeed. This is not the case if we confine the heat engine to a much smaller, isolated, environment, then the environment heats up and a gradual decrease of the production of power by the engine results. Eventually, power production stops when the temperature of the environment reaches the temperature of the hot source. The combination of the heat engine and the environment, the isolated combined system, reaches thermodynamic equilibrium and all processes, also the production of power, stop. We see that both the hot side and the cold side are crucial. We see, as we explained earlier (Section 2.2) that it is a gradient in a state variable, in case of the heat engine temperature, that creates a thermodynamic force that drives the system out of equilibrium and allows the production of power to do work. In the case of the heat engine, the difference of the reciprocal of absolute temperature drives an irreversible process that generates power.

Can we expand this example to get some first feeling about the performance of economic work to drive the creation of economic value? In introducing EVT, we noted that the cost of information is the EVT equivalent of temperature. We thus suspect that EVT predicts the possibility of performing economic work if we identity a source of value with a high cost of information. What we also need, however, is a sink with a low cost of information. In that case, information related gradients exist and as long as these remain in place, the creation of economic value by doing economic work is possible. We discuss this more exhaustively in the remaining sections of this chapter. We also understand that equilibrium results in absence of a sink with a cost of information different from that of the source. In case of an economic transaction, this results in an equilibrium transaction in which no economic gain is possible for any of the parties. Only where gradients in information exist, also called information asymmetries in economic jargon, meaningful transactions are possible. The world of the equilibrium economic system does not allow creation of value. Above we only discussed asymmetries in cost of information. Later on, we introduce asymmetries due to differences in statistical entropy and identify differences in the ratio of economic value to the cost of information as the driving force for doing economic work.

5.7. Forces in thermodynamics and Economic Value Theory.

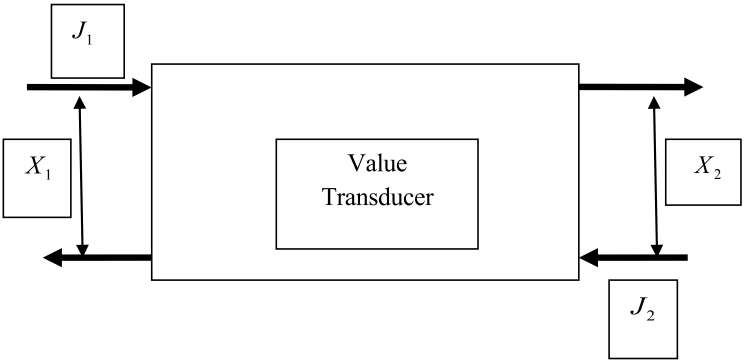

Fig. 5.2. A value transducer.

In thermodynamics a branch of non-equilibrium thermodynamics, irreversible thermodynamics, emerges in the beginning of the 20th century, based on the pioneering work of Onsager (1931a, 1931b). Before that, thermodynamics restricts itself to equilibrium situations or idealizations such as reversible processes. Irreversible thermodynamics analyzes situations beyond thermodynamic equilibrium where processes take place driven by thermodynamic forces. It turns out that the ratio of free energy, G, and temperature, T, defines thermodynamic forces, indicated by the symbol X:

$$X = \Delta(G/T) \tag{5.7}$$

In this equation, the symbol *Δ* stands for a difference. For the general case, there are many forces and corresponding processes. For simplicity's sake we stick to the case of two forces driving two processes. This case stays manageable from the (mathematical) complexity point of view and shows all the fundamental features necessary for and relevant to the subject matter of this book. As said, the forces drive processes and we illustrate this using the example of a value transducer (Fig. 5.2). We adopt the EVT analogue of the thermodynamic forces to define forces in terms of differences in the ratio of economic value and cost of information.

The input side of the value transducer is an exchange flow due to a force X_1 defined in terms of economic value, its rate is J_1. The input force is positive and the process proceeds spontaneously from the perspective of the second law as the process produces statistical entropy. We term such process downhill from the perspective of the second law. The term downhill reflects the analogy of a river flowing downhill from a mountain, i.e. in the direction we naturally expect it to flow. The output side of the value transducer involves a process that proceeds against the direction of the second law, as the force is negative. However, the flow is positive because the mechanisms inside the value transducer couple the downhill process at the input side to the uphill process at the output side. In the heat engine example, this output is the transformation of heat into work. This process of coupling of uphill to downhill processes is fundamental to the understanding of evolution. To illustrate this we further explore the relation between the second law and structures like the value transducer introduced here.

Our analysis concerns the simplest case of value transduction. It takes place in analogy to the so-called linear region of irreversible thermodynamics. The linear value transducer operates beyond equilibrium but we stay relatively close to equilibrium where linear relations between the flows and the forces prevail. In that case, the following equations describe the behavior of the transducer in Fig. 5.2:

$$J_1 = L_{11}X_1 + L_{12}X_2 \tag{5.8}$$

$$J_2 = L_{12}X_1 + L_{22}X_2 \tag{5.9}$$

In eqns. 5.8 and 5.9 the L's are constants not dependent on the forces, i.e. the flows are true linear combinations of the forces.

Several interesting features appear in these equations. Firstly, we need to introduce some jargon. X_1 and X_2 are the so-called conjugate forces of the input and the output flows respectively. These forces directly drive their conjugate flows. Additionally, these forces are non-conjugate to the flows at the output and the input respectively. The forces also drive their non-conjugate flows. This so-called coupling is, as said, crucial for understanding complex real systems. As is apparent from eqns. 5.8 and 5.9, the coupling coefficient of flow 1 to force 2 equals that of the reverse coupled pair. We can visualize this so-called reciprocity if we assume coupling based on a friction mechanism. When to surfaces move relatively to each other they both experience the same friction force. Among others the present author (Roels 2010, 1983), analyzed the linear value transducer extensively. Here we avoid the mathematics of the analysis and only cite the conclusions.

As said, the heat engine is an example of coupling of a heat flow and generation of power. Power is the work per unit time and that is the quantity of interest in practice, reversible infinitely slow exchange of work has no practical meaning. Power output or input is proportional to the product of a flow and its conjugate force. An example is the hydrodynamic generation of electrical power by the coupling of uphill and downhill processes. In this case, a dam in a river flowing downhill contains a power generation plant. The water flowing through the dam drives an electrical generator and electrons move up a gradient in voltage. This process results in useful energy by the coupling of the generation of electrical power to the natural process of the water flowing downhill. This positive force drives electrons uphill against the gradient in voltage.

Of course, as all systems, the transducer is subject to the second law and as irreversible processes take place in the transducer, the entropy production in the combined coupled processes must exceed zero; the equality sign cannot apply, as we are not in equilibrium. This leads to conclusions about the magnitude of the coefficients in eqns. 5.8 and 5.9. The coefficients of coupling of the flows to their conjugate forces must be positive. Furthermore, there is an upper limit to the coefficient of cross coupling, the coupling of the flows to their non-conjugate force. This upper limit defines a so-called degree of coupling that must always be smaller than one.

In the value transducer a dissipation of power takes place equal the product of the cost of information and the production of statistical entropy. This dissipation is necessarily positive because the value transducer being beyond equilibrium must show a positive entropy production due to the second law. Hence, Prigogine (1980) coined the term dissipative structures for systems like the value transducer. They escape the forces of the second law by dissipation of part of the power derived from their coupling to a positive force in their environment. The structures have lower statistical entropy than the one characteristic for thermodynamic equilibrium. A careful analysis of the linear value transducer reveals that these structures exhibit the phenomenon of so-called maintenance dissipation, i.e. they need power dissipation even if they do not produce power. If they do not have access to sufficient power to dissipate, their structures degrade and the system's entropy increases until it reaches equilibrium.

This author (Roels (1983)) used the model of the value transducer to study so-called oxidative phosphorylation in microorganisms. Microorganisms oxidize an energy source such as sugar to carbon dioxide and water; this is equivalent to a process of low temperature burning just as the heat engine uses coal. However, it is a very intricate and clever process of burning in which energy only partly emerges as heat; a part serves to transform ADP in ATP, complex chemical substances whose naming does not have to bother us. ATP is the free energy currency of the cell. It supplies it with free energy to grow and maintain its non-equilibrium structure. The efficiency of the generation of power in oxidative phosphorylation is of the order of 60-65%. This compares with the efficiency of the generation of electrical power in a power plant that is in the order of 50%. Roels followed the Nobel laureate Schrödinger (Schrödinger (1945)) in the view that microorganisms are dissipative structures that feed on "negentropy", the term Schrödinger coined to indicate free energy. Furthermore, Roels identified the maintenance dissipation in the energy transducer as the known feature of maintenance energy in microorganisms. Recent work (Roels (2010)) identifies industrial corporations as examples of dissipative structures that feed on sources of economic value, such as the need for a product in the market that allows an economic transaction resulting in the capacity to do economic work. We analyze this analogy further in subsequent chapters (particularly Chapter 7).

We can use the example of microorganisms to further illustrate the feature of coupling of a thermodynamically downhill process to an uphill process against the natural direction dictated by the second law. Consider a microorganism growing on the sugar glucose (of course, it also needs sources of the other chemical elements in biomass but we can conveniently ignore this for the purpose of the discussion here). As said, part of the glucose is oxidized to result in carbon dioxide and water. Part of the energy that this combustion process releases conserves in a source of free energy, ATP. This part of the glucose drives a thermodynamically downhill process. It is the input of the value transducer of our organism. Another part of the glucose is used for building new biomass. This represents an uphill process as the free energy of the product, biomass, exceeds that of the glucose used. However, the free energy in the ATP the combustion of glucose generates drives this process. As ATP is part of a closed cycle in the organism, we can ignore it in our thermodynamic analysis and the net effect is that an uphill process, the production of biomass from glucose, is driven by a downhill process, the combustion of glucose. In fact, ATP provides the mechanism by which the downhill and uphill process couple.

We conclude this section by pointing out some differences between dissipative structure and equilibrium structures. Also at equilibrium, ordered structures appear. The classical example of the difference between dissipative and equilibrium structures is the distinction between a snowflake and a microorganism. Snowflakes, structures composed of ice crystals, have a beautifully ordered appearance. Snowflakes are equilibrium structures. Microorganisms, on the other hand, are marvels of order at the molecular level and need support by a constant flow of a source of energy, such as sugar. Microorganisms are dissipative structures. The same applies to a city. A city only maintains its structure and order if a continuous flow of food and energy from the environment exists. Another difference rests in the ability to generate work. Snowflakes being equilibrium structure do not dissipate and produce no work. Dissipative structures can do work. Prigogine and his co-workers proposed dissipative structures as a model of complex organized structures such as organisms. In earlier work of this author (Roels (2010), we propose these structures as a model for the evolution of the socioeconomic system, including organizations like industrial corporations. We return to these matters in e.g. Chapter 6 discussing a systems theory of evolution and in Chapter 7 where we discuss the nature of the firm.

5.8. Transformation of information in economic value: The concept of price.

This section treats creation of economic value using the EVT formalism. Fig. 5.3 presents the system under consideration.
An investor invests in a risk-bearing asset characterized by a cost of information, an intrinsic value and a statistical entropy. The investor pays a price *P* using liquid capital with zero statistical entropy. We further assume that buyer and seller a characterized by the same cost of information and the same statistical entropy concerning the traded asset. The system exchanges a flow of information with the environment.

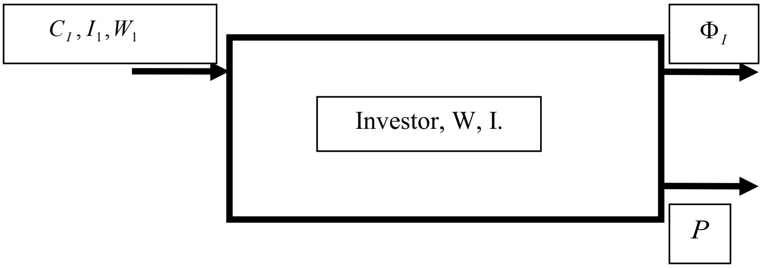

Fig. 5.3. A transaction involving a risk bearing asset.

Application of the first law of EVT results in:

$$\frac{dW}{dt} = W_1 - \Phi_I - P \tag{5.10}$$

The second law results in:

$$\frac{dI}{dt} = \Pi_I + I_1 - \frac{\Phi_I}{C_I} \tag{5.11}$$

We assume that the statistical entropy that characterizes the investor does not change, i.e. the left hand side in eqn. 5.11 vanishes. Furthermore, we ask ourselves what the price is the seller accepts for the asset. The conclusion must be that he at least requires a price reflecting his perception of the economic value of the asset. Hence, his price is $W_1 - C_I I_1$. If we do the rather elementary math we conclude that the gain in economic value of the investor is equal to $- C_I \Pi_I$. This result is rather troubling for the investor. As the production of statistical entropy must be positive, the investor loses economic value if he pays the price the seller expects. Hence, a transaction perceived beneficial by both buyer and seller never takes place.

This discussion leads in summary to the following observation. If buyer and seller have the same perception of statistical entropy and costs of information, a transaction can at best lead to a zero change of the amount of value to the investor. This situation is a transaction where the production of statistical entropy reaches the minimum of zero allowed by the second law. Generally, the investor looses economic value if the production of statistical entropy in the transaction is positive. The result corresponds to the conclusion we reached earlier: We need a gradient in cost of information to achieve a gain of economic value. The discussion in this chapter shows that both a difference in cost of information and a difference in the appreciation of the risk associated with the asset can lead to a gain or loss of economic value. The problem is that the information set available to all players in the market does not allow a prediction whether the product of cost of information and the amount of risk will go up or go down. In a situation of an efficient market, i.e. a market in which buyer and seller avail of the same information and use it in an optimal way, the average gain or loss of economic value for the collective buyers and sellers is zero in the limiting case of no production of statistical entropy.

5.9. A more generalized market transaction.

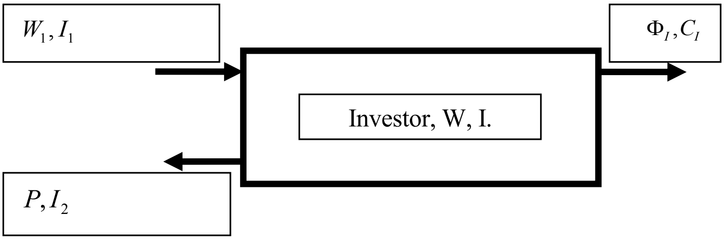

Fig 5.4. A generalized market transaction.

Fig. 5.4 represents a more involved market transaction. An investor, characterized by a value and statistical entropy and a cost of information acquires one unit of a risk bearing asset. The value per unit is W_1 and the statistical entropy per unit I_1 . The investor pays a price P per unit in risk free value. The cost of information of the seller and the buyer is equal and we use its familiar notation. However, the probability distributions of value of the buyer and the seller are different as their states of information differ. The seller's statistical entropy is different from that of the buyer. Information exchanges with the environment. Application of the first and second laws of EVT and realizing that the economic value, and hence, the expected price, to the seller is $W - C_I I_2$ leads, for a situation in which the statistical entropy of the buyer does not change, to the conclusion that the gain of value by the investor is:

$$\frac{dW}{dt} = C_I(I_2 - I_1) \tag{5.12}$$

We see that contrary to the situation in the previous section a positive increase of the value for the investor is possible if the right hand side of eqn. 5.12 is positive. This applies if the statistical entropy of the picture of reality of the seller is higher than that of the buyer. This applies if the buyer avails of superior information. In economic theory, this is termed a case in which asymmetries in information exist. The derivation of eqn. 5.12 follows in note 5.1; the reader can skip it if (s)he accepts that the equation is right.

Note 5.1.

With reference to Fig. 5.4, we write the first law of EVT as:

$$\frac{dW}{dt} = W_1 - P - \Phi_I$$

As we indicated, the seller accepts a price according to his perception of the economic value of the asset:

$$P = W_1 - C_I I_2$$

If we substitute this in the first law equation, we get:

$$\frac{dW}{dt} = C_I I_2 - \Phi_I$$

The second law of EVT reads:

$$\frac{dI}{dt} = \Pi_I + I_1 - \frac{\Phi_I}{C_I}$$

If the entropy of the investor remains unchanged, the term at the left hand side of the equation above vanishes and it follows:

$$\Phi_I = C_I(I_1 + \Pi_I)$$

Substituting this in the last first law equation shows:

$$\frac{dW}{dt} = (-C_I \Pi_S) + C_I(I_2 - I_1)$$

Eqn. 5.12 is the limiting case that we obtain if the production of statistical entropy is zero.

These asymmetries in information are the forces that drive transactions in a non-equilibrium situation where gradients in statistical entropy occur. In a practical situation of non-equilibrium in an economy, many forces as defined in eqn. 5.7 exist and these are the driver behind socioeconomic evolution. We discuss this more exhaustively in Chapters 6 and 7.

5.10. Conclusion.

It is instructive to compare the result obtained in the preceding section with the classical microeconomic model of perfect competition. We include this model in our analysis of economic theories in Section 8.8. One assumption underlying the theory of perfect competition is that the market shows a large number of buyers and sellers. The entry and exit barriers facing suppliers are negligible. All producers and consumers avail of the same information and use that information in the most effective way. No inhomogeneities in product and production methods exist; the suppliers produce one product in a standardized way. Considering the situation in real life markets, these assumptions have a questionable degree of realism.

These assumptions lead to a market in which a balance between supply and demand exists. The sellers in the market earn only a "normal" return; probably this means that the players only earn the risk free return and hence there is no overall creation of economic value. Furthermore, the situation is at equilibrium socially optimal, it leads to an optimal allocation of resources in the economy. We show later that these conclusions generally do not apply to the harshness of economic reality.

The result concerning the absence of creation of economic value agrees with our discussion of an EVT equilibrium transaction as discussed in Section 5.8. In addition, EVT shows that this critically depends on the information of the various players. In case of asymmetries in information, over average returns are possible. Asymmetries in information do not exist in the perfect competition world.

Clearly, the perfect competition model as well as the EVT formalism applied to an equilibrium situation, lead to severe discrepancies between the theory and some observations regarding real life situations. It is difficult to find the rationale for the existence of firms as these offer no clear advantage over transactions completely based on market exchange. In addition, the fact that some firms earn over-average returns, prosper and grow, whilst others disappear, is problematic from the equilibrium perspective.

Perhaps the best reflection of the problems with efficient markets and equilibrium theory is the way in which Jensen (1972) defines an efficient market. A market is efficient with respect to a given set of information if none of the players can earn a profit if they optimally use the available information and have access to the same information. This agrees with the information based EVT approach. If the information set of all players is equal, no creation of economic value by transactions is possible. EVT, however, also shows that over-average returns are possible if asymmetries in information exist. In EVT, asymmetries in information follow from the nature of the probability distribution of value. The probability distribution relevant to an actor depends on his information about the state of the system. As this level of information may be different for suppliers and sellers, such differences are very likely to exist. In fact, in later chapters (e.g. Chapter 6) we show that these are the very driver of evolution and constitute the substance that makes economies tick.

The EVT formalism has, in addition to the ability to tackle information asymmetries, a lot more to offer in removing the limitations of the perfect competition and the equilibrium approaches to markets. Clearly, an equilibrium approach is not adequate to describe situations in markets in which firms and other institutions appear. Such markets are definitively not in equilibrium and firms and other organizations are the product and the source of this disequilibrium. In a situation away from equilibrium, competition between firms creates, maintains and destroys gradients in information and costs of information. The underlying asymmetries in information lead to a situation in which over-average returns appear. In the general non-equilibrium case firms, markets and industries constantly evolve under the pressure of competition between the actors that contest the sources of economic value.

Based on our EVT approach we also identify the forces that drive market transactions and the appearance of industries and markets. In analogy with thermodynamics, the force that drives transactions is a gradient in the ratio of the economic value to cost of information (eqn. 5.7). It leads to the interpretation of firms and other organization as dissipative structures that inevitably appear if forces exist that result in evolution beyond equilibrium.

CHAPTER 6. SYSTEMS THEORY OF EVOLUTION.

6.1. Introduction.

In Chapter 5, we introduce the linear value transducer. We show that it exhibits features that resemble those of forms of organized matter such as microorganisms and broader some of the organizations existing in our society. The treatment further substantiates the concept that organizations and organized forms of matter, such as firms and organisms, are examples of dissipative structures that feed on the opportunity to create a source of free energy or economic value to maintain their structure and to grow and develop. Several authors, e.g. Schrödinger (1945) and Prigogine and co-workers (Glansdorff and Prigogine (1971), Prigogine (1980), Prigogine and Stengers (1984) and Nicolis and Prigogine (1977)), reach similar conclusions. Prigogine coined the term dissipative structures. However, we need to overcome a significant hurdle. In this chapter, we follow the approach of Prigogine to illustrate the problem. Firstly, we reemphasize that the formation of complex ordered structures that evolve, is incompatible with equilibrium (at equilibrium the statistical entropy reaches a maximum). More significantly, these are also incompatible with non-equilibrium states that are too close to equilibrium. In the range of conditions where the linear phenomenological rate equations hold (i.e. the equations governing the linear value transducer), formation of dissipative structures takes place but complex forms of organized matter or organizations and markets never evolve. This sheds some doubt on the usefulness of the linear model of the economic value transducer in the description of phenomena of complex organization. In this chapter and the next chapter, we analyze and resolve these problems. We largely follow the line of thinking in the above-mentioned works of Prigogine and his coworkers. In the next section, we investigate the stability of non-equilibrium steady states (states that, on the macroscopic level, no longer change) in the strictly linear region.

6.2. The linear value transducer revisited.

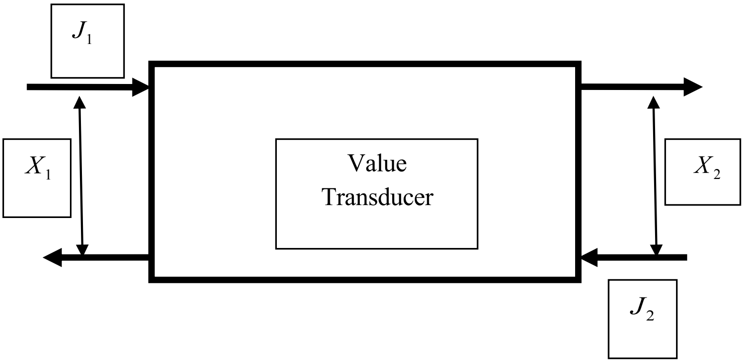

Fig. 6.1. A linear value transducer.

The subject of our analysis is a system in which two forces operate to drive two transactions or transformations according to linear phenomenological laws. Chapter 5 introduces this system (Fig. 6.1) and analyzes some of its properties. It illustrates the coupling of a downhill force, e.g. derived from the need for a given product functionality by consumers, to an uphill process, e.g. the realization of a profit by a company that supplies a product satisfying that market need.

Initially the system is in equilibrium. At the beginning of the evolution of the system, the input force becomes operative that drives the system away from equilibrium. There is production of statistical entropy equal to the sum of the products of the flows and their conjugate forces. We assume the force at the downhill side of the transducer constant whilst allowing the output force to vary. If we wait for a sufficiently long time a situation evolves, in which the macroscopic state variables characterizing the system no longer change. As a force is operative this so-called steady state cannot be equilibrium, it is a non-equilibrium state. We term these states dissipative because the entropy production does not become zero, as is the case at equilibrium. Entropy production, i.e. dissipation of free energy or economic value, continues to take place. This is a direct consequence of the fact that these structures have a statistical entropy lower than that corresponding to thermodynamic equilibrium. Equilibrium results when the downhill force is no longer operative, e.g. if the consumers are no longer buying the product. In the following, we analyze the stability of such steady states. Fig. 6.2 illustrates the concept of stability.

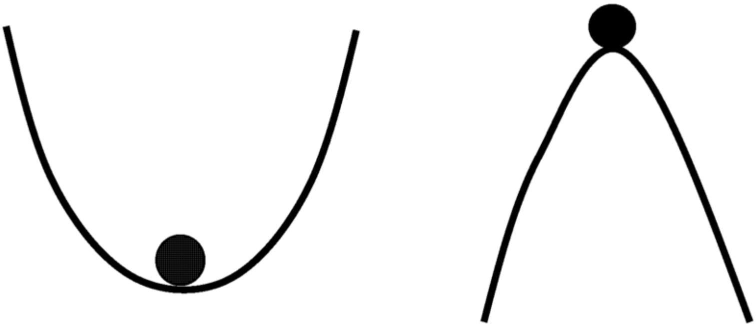

Fig. 6.2. Stable and unstable steady states.

The left hand side of Fig. 6.2 shows a marble at the bottom of a valley. If a disturbance moves the marble slightly to the right or to the left, it returns to its steady state position at the bottom of the valley. In fact, even large fluctuations do not result in a permanent disturbance of the position of the marble. Such a steady state is stable. Thermodynamic equilibrium is a stable state and the slopes of the landscape derive from economic value that reaches a minimum at equilibrium.

The picture at the right hand side of the figure presents an alternative situation. There the marble is at the top of a hill and, although it is in principle possible to balance a perfect marble at that top, such a situation is not likely to be lasting. The slightest disturbance causes the marble to leave the top and it will not spontaneously come back. If the landscape considered in both examples is the system's state space with only one dimension, height, the time change of the state variable would vanish in both cases, i.e. both follow the requirement of a steady state. However, the right hand side of fig. 6.2 represents a steady state that we do not observe in practice, as it is intrinsically unstable. The analysis of the stability of steady states is a branch of science accessible by mathematical methods that we will not pursue here (see e.g. Roels (2010) and the references therein).

In the near equilibrium range, where the linear Onsager relations apply, intrinsically stable steady states lastingly appear. In these steady states, the production of statistical entropy seeks the best alternative to being zero (zero would result if the downhill force vanishes and the system reverts to equilibrium), it reaches the minimum production of statistical entropy compatible with the force keeping it away from equilibrium. It more or less economizes on the dissipation of economic value. This line of reasoning shows that in the realm where the linear phenomenological relations hold, non-equilibrium steady states are intrinsically stable.

Considering the objective of this work to extend the theory to include complex systems in biology and our human society, this conclusion casts doubts on the relevance of our approach so far. Biological evolution leads to a great variety of complex organisms including *Homo sapiens*, our species. In addition, in our society very complex patterns of organization appear, phenomena ranging from language, through science to industrial corporations. Everywhere we see sustained evolution, in apparent conflict with equilibrium but also with the stability of steady states if we remain in the linear region, where further evolution of the steady state does not occur due to its intrinsic stability. Fortunately, developments in thermodynamics in the second half of the 20th century extend the formalism beyond the linear range; this is the subject of non-linear non-equilibrium thermodynamics. This extension of thermodynamics to the systems we want to study is the subject of the following sections.

6.3. Instability of steady states and evolution.

In this section, we relax the limitation to the strictly linear region beyond equilibrium and extend the basic theory of value transduction. Non-linear relations between flows and forces are both a characteristic and a source of these non-equilibrium developments. The remaining part of this chapter focuses on further developing the theory and analyzes its practical consequences to arrive at a systems theory of evolution. We show that theory accommodates evolution of complex organized structures such as human society with its markets and industries.

Fig.6.3 provides a more complex variety of the stability characteristics illustrated in Fig. 6.2.

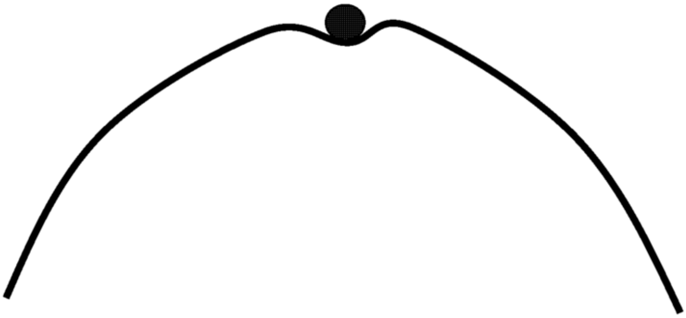

Fig. 6.3. A situation of conditional stability.

In Fig. 6.3, we have our marble in a small dip on a hill. It is in a steady state and on small fluctuations, it returns to that state. However, larger fluctuations drive it “over the top” and its position becomes unstable. Once driven out of the dent, it never returns to its initial steady state. The initial position of the marble is conditionally stable. Small fluctuations do not drive it away from its steady state as it returns to that state, large fluctuations do.

Situations that are more complex exist such as the one in Fig. 6.4. It shows a combination of a conditionally stable and a stable steady state. When small fluctuations occur in the steady state indicated A, it remains stable. A large fluctuation to the left results in instability. Large fluctuations to the right cause the system to shift steady states; it stably ends up at point B. This is indicative of the richness of possibilities that arises if state space becomes more complex, even if we limit our discussion to only one dimension. For the general case, state space has far more dimensions than one, certainly in the complex systems we study in this book. This leads to a very complex landscape with many possibilities for a shift between steady states that are usually only conditionally stable.

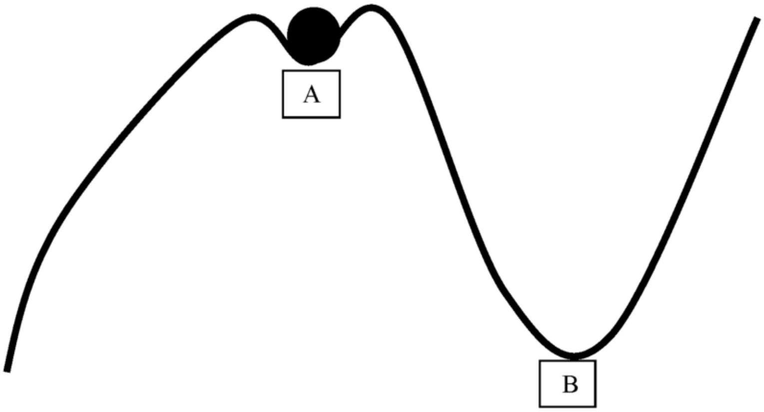

Fig. 6.4. A shift between steady states.

We summarize our findings so far as follows. In the more straightforward situations, a steady state may be intrinsically stable or unstable with respect to all fluctuations of the state variables. One should remember that the macroscopic description only contains information about the average values of the state variable. At the microscopic level, the "state variables" of the system fluctuate around the average. Hence, unstable steady states never appear as these fluctuations challenge such states. Furthermore, we conclude that in the strictly linear region beyond equilibrium, steady states, once established, are not subject to further change as these are stable against all fluctuations. In more involved situations in the non-linear range, more complex behavior results. Steady states show stable against some fluctuations but not to others. The behavior depends on the nature of the fluctuations. This immediately leads to an important conclusion. Information about the fluctuations in the system is not present in the macroscopic description as this approach averages out fluctuations in the process of formulating macroscopic properties. Hence, if future behavior depends on the nature of fluctuations, the macroscopic description provides no definite clues about the behavior of the system. The predictive power of the macroscopic description concerning future behavior evaporates for such systems. As we show in the more involved example depicted in fig. 6.4, this has drastic consequences such as a sudden instability of a steady state if a large fluctuation hits it. In addition, a shift from one conditionally stable steady state to another stable steady state on a fluctuation is possible. There is one final feature that we deduce from the simpleminded examples introduced so far. We refer to fig. 6.4. As long as the fluctuation that hits the system remains within certain limits the system remains in the steady state even if it is unstable with respect to the larger fluctuations. When a, less likely, larger fluctuation occurs the system suddenly changes steady states. The system remains frozen in its steady state for some time and then suddenly changes drastically.

There exists a full-fledged mathematical theory of non-equilibrium steady states and the behavior of systems in the non-linear region beyond equilibrium. We do not introduce this theory here, because the qualitative discussion fully serves the purpose of this book. Interested readers can consult my earlier work (Roels (2010)) and the references therein.

The detailed microscope state of the system is far from known if we only have a macroscopic picture of reality and the richness of the microscopic behavior defines the type of fluctuations that occur. By the very nature of the macroscopic description, that information is not available in such an approach. The development routes the system takes come as a surprise to the reduced information observer. To follow Monod (1971) the system is subject to "Chance and Necessity".

Fluctuations that appear random to the macroscopic observer introduce chance; given the nature of the system, such developments necessarily happen because of the system's detailed microscopic structure.

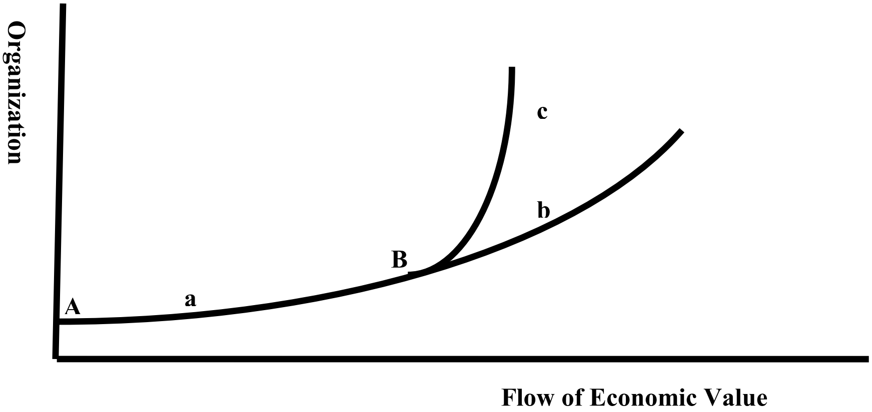

Fig. 6.5. Developing dissipative structures: An elementary bifurcation diagram.

In Fig. 6.5, we introduce an elementary picture of the development of dissipative structures. It shows the level of organization of the system, i.e. the extent to which complex structures exist, in relation the flow of economic value or free energy that drives the system away from equilibrium. This allows us to summarize our discussion as follows. If a system ages in the absence of coupling to an external source of economic value, it finally reaches a state of equilibrium. In that state, the macroscopic appearance of the system changes no longer. However, fluctuations remain in place from the microscopic perspective. At equilibrium, the system's statistical entropy reaches a maximum and it is not able to perform value-adding work. This is point A in the bifurcation diagram in Fig. 6.5. If the system succeeds in coupling to a source of economic value in the environment, a new steady state emerges. Initially, at increasing but moderate forces, the system moves further and further away from equilibrium and the resulting steady states remain stable as long as we remain in the strictly linear region. Developments unexpected from macroscopic reduced information perspective do not take place. The system develops along the so-called macroscopic branch AB of the bifurcation diagram. In terms of economic activity, point A reflects the perfect competition world: Everybody supports his own needs by buying and selling goods at equilibrium prices. The transactions taking place create no economic value. Along the branch AB, some elementary kind of specialization and division of labor occurs and asymmetries in information with respect to certain goods and services start to develop and allow some return to the specialists with more information. At point B, the source of economic value becomes so potent (e.g. caused by the increasing information advantage of some actors) that the macroscopic branch becomes unstable to some of the fluctuations in the system. A new development, indicated by the line c in our bifurcation diagram, characterizes the newly possible steady states. Macroscopically organized states different from the unstable less organized states obtained by extrapolation of the macroscopic branch, become possible. As an example, superior technologies develop and this allows extracting more economic value from resources and needs in the environment. The system continues developing along the macroscopic branch and nothing "unexpected" occurs. However, the macroscopic branch is unstable with respect to some of the fluctuations in the system and the system may suddenly jumps from the macroscopic branch to the branch indicated c.

Note 6.1.
A simple physical example, only limitedly representative, illustrates the phenomenon of instability of the macroscopic branch beyond a certain critical limit. Consider water at an ambient temperature, where it is liquid. In the fluid, the molecules can move around relatively freely and the structure does not show order from the macroscopic perspective. In fact, this absence of order is partly an illusion. At 25 degrees centigrade, there is already appreciable hydrogen bonding, i.e. bonding based on interaction between the hydrogen and oxygen atoms of different water molecules. These bonds are still of a transient nature and do not drastically hinder the water molecules from moving around freely. Different values for the average number of hydrogen bonds per water molecule at 25 degrees appear in the literature. The values range from about 2.4 to 3.2. For liquid water at 0 degrees, the literature states a number of 3.7, i.e. already very close to the number of four representative for ice. At ambient pressure, liquid water is no longer the stable form below 0 degrees. The stable form is ice, an ordered solid structure in which each water molecule engages in four hydrogen bonds with its neighbors. This is an example of a phase transition resembling the appearance of order in the region where the macroscopic branch is no longer stable and structures that are more complex appear. We stressed that the appearance of more complex dissipative structures may not directly occur if we pass the critical point of the bifurcation diagram. The system may continue to develop along the macroscopic branch and suddenly jump to the more stable complex dissipative structure some distance beyond the critical point. This resembles the situation in liquid water where we sometimes observe a phenomenon of super cooling of -40 degrees. The water remains liquid beyond the limit of its normal freezing point. The analogy with the situation of formation of complex dissipative structure beyond the critical point is apparent. The difference is that dissipative structures are non-equilibrium structures whilst ice is an equilibrium structure.

In a more complex system, the situation can be more involved. A bifurcation diagram such as Fig. 6.6 may apply. Again, the system is initially at equilibrium and it starts developing along the macroscopic branch by the creation and exploitation of increasingly potent sources of economic value. At point A, two types of development become possible when the system leaves the macroscopic branch; these developments are AB and AC. The macroscopic observer has no way of knowing what will happen, not only are both development paths unexpected to the observer but, in addition, he has no way of knowing along which branch the system evolves or even being aware of the fact that the system follows a new development route. The behavior of the system becomes unpredictable. Developments become revolutionary to the observer. As the force further increases new bifurcation points B and C appear and the system chooses again between accessible development routes. This may again happen at point D.

The behavior indicated above has important consequences. The first feature is that, once we pass several bifurcation points, the state of the system depends very much on the path the system took at the various bifurcation points. The historical evolution depends on the "choices" made, or alternatively, the fate of the system at the successive bifurcation points. This strongly reminds of features of biological evolution, career paths of people and the historical dimension of the development of firms. From an equal or very similar situation in the past, the evolution often leads to very different situations today. Numerous examples appear in industry, e.g. in the pharmaceutical industry where present-day big pharmaceutical firms have histories ranging from emerging as a drugstore (GlaxoSmithKline) or a producer of citric acid (Pfizer).

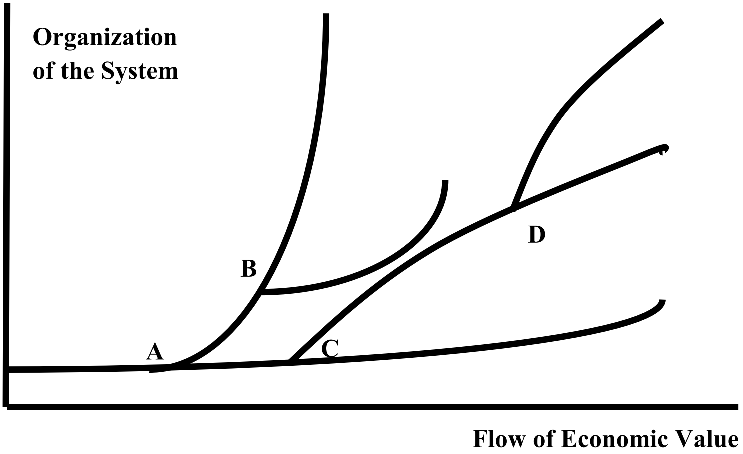

Fig. 6.6. A more complex diagram showing successive primary, secondary and tertiary bifurcations.

The second feature relates to predicting the future. The reduced information picture of reality leads to an unpredictable future. This applies to all actors in the market, even to the management of a firm, although that management generally has more information about the present state of the firm than observers outside the company do. Even for the management full understanding of the competitive environment, the structure of the market in which they operate and the detailed knowledge of the exact value microstate of their firm, requires far more information than the management can economically obtain. The ramifications of this for e.g. strategic management are obvious. We discuss this in somewhat more depth in Sections 8.15 and 15.4. We recover the concept of "bounded rationality" as proposed by, among others, Williamson (1975). Management finds itself in a situation where even apparently completely rational decisions are subject to uncertainty due to the complexity of reality.

So far, we have only discussed the stability characteristics beyond the critical limit where the strictly linear region ends. The system may show intricate transient behavior when a structure becomes unstable and evolves to another stable development route. Furthermore, the return to stable states on a fluctuation may show involved dynamics. This is the subject of evolutionary dynamics of linear and non-linear complex system. It is beyond the scope of this book to treat this important subject in detail. It is, however, to important to the subject matter of this work to leave it completely untouched. In some instances the dynamics of shifts between steady states and the return to steady states after a disturbance by a fluctuation may be smooth, i.e. a gradual development results. However, certainly in complex systems, this is exception rather than rule. The theory of system dynamics introduces, as we discussed in Chapter 2, the concept of so-called relaxation times, characteristic for the dynamics of the damping of fluctuations and the movement to another state, if the initial state becomes unstable. A more or less smooth transition takes place if the relaxation times are real numbers. However, if, some of the relaxation times are complex numbers, wave like phenomena and oscillations may occur. This occurs in patterns of chemical reactions such as in the metabolism of organisms (Glansdorff and Prigogine (1971)) where we observe oscillations with a variety of characteristic times. Also in biological ecosystems (Nicolis and Prigogine (1977)), such phenomena are common. It is not difficult to see the relevance of this behavior for the modeling and analysis of the socioeconomic systems. This framework easily accommodates short and long-term fluctuations in the stock markets and the various cycles that appear in our economies. In Chapter 2, we stated that, for linear systems, the dynamics may be complex but complete predictability remains. As we argue there and further substantiate in this chapter this does not hold beyond the linear region where predictability gets lost due the very nature of the macroscopic description. The macroscopic description only contains average properties of the objects in the system. At bifurcation points, extremes in the properties of the objects govern future behavior and a description in terms of averages becomes a bad guide for predicting that future.

This section develops the foundations for the extension of EVT into the non-linear region. To illustrate this further we discuss some physical examples in the next two sections.

6.4. The Benard problem and instability of the macroscopic branch.

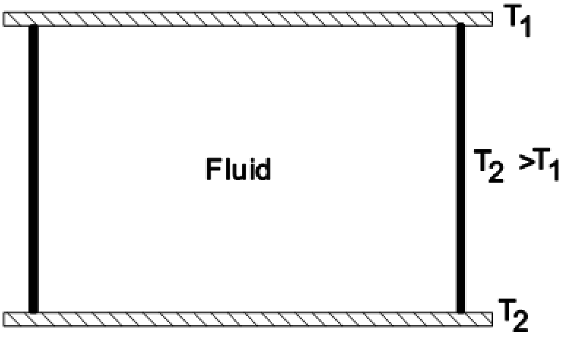

Fig. 6.7. The Benard problem: A fluid is contained between two parallel plates and heated from below.

The Benard problem is an interesting example illustrating instability of the macroscopic branch. We consider the system depicted in Fig. 6.7.

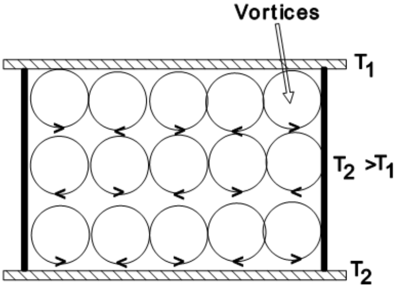

Fig. 6.8. The Benard problem: Structure formation at a temperature difference exceeding the critical level.

We heat a fluid contained between two parallel plates from below. The temperature gradient between the two plates provides a source of free energy. A flow of heat from the lower to the upper plate develops because of the force the temperature difference creates. For a moderate difference in temperature, heat flows through the fluid by conduction, i.e. by interaction between individual molecules. There is no macroscopically observable movement of the fluid.

A simple macroscopic mathematical model, such as the Fourier equation for heat conduction, adequately models this situation.

Fig. 6.8 shows the changes that occur when the temperature difference between the upper and the lower plate exceeds a critical value. The state of heat transport by conduction, without observable macroscopic movement of the fluid, is no longer stable: We leave the conduction regime. The fluid develops more or less regular convection cells or vortices, i.e. macroscopically observable circulation patterns. In these vortices, large numbers of molecules cooperate in a concerted fashion. Clearly, if the free energy gradient exceeds a critical level this leads to "organizing" the system. An interesting feature involves the dissipation of free energy. It is larger than the rate corresponding to a continuation of the macroscopic branch. This implies that conduction is less effective in coupling to a gradient in free energy, than the cooperative efforts of the molecules in the convection cells. These molecules "cooperate" to take better advantage of the free energy gradient in the environment.

Loosely phrased this leads to the following interpretation. As soon as the temperature gradient exceeds a critical value, an opportunity arises to exploit a source of free energy to a wider extent than possible based on the interaction of single entities. The possibility exists to exploit the resource contained in the gradient more exhaustively if groups of entities cooperate in ordered structures.

The foregoing discussion uses a physical example involving free energy. We can also imagine an economic value analogue. If in an economy a gradient of economic value exists, for example due differences in information between a buyer and a seller, a non-equilibrium structure may develop that exploits this gradient in a more effective way than based on an exchange between single actors. At first, in the strictly linear region, no drastic changes in the structure of the system result. In this situation, the gradient in information is still moderate. At a certain critical size of the gradient in economic value, the system drastically changes and ordered cooperative structures appear exploiting the force in a more effective way. This provides an analogy with the nature of the firm. To some authors (Alchian and Demsetz (1972)), a basic feature of the firm rests in team production and division of labor, based on the coordinated effort of many individuals.

The structures that develop may become instrumental for further increasing the gradient in economic value, e.g. by increasing their information advantage. After a while, it is hard to judge whether the structures exist because of the gradient of economic value or whether the structures themselves create that force. A chicken and egg situation arises and the question what came first looses relevance.

We summarize some important features of the elementary example of dissipative structures we discuss in this section. Firstly, a force exceeding a critical limit exists that drives a flow. Only if there are opportunities to harvest sufficient economic value, the development of evolving dissipative structures is possible. Secondly, there is the aspect of cooperation between the entities in the system. To exploit the increasing potential for harvesting economic value effectively, the entities must operate in a concerted fashion. We return to the rationale behind "the firm" (Coase (1937)) and the concept of "team production" suggested by Alchian and Demsetz (1972). Thirdly, a form of a non-linearity, a deviation from a strictly linear relation between flows and forces is in place. In the Benard problem, this non-linearity exists in the inertia phenomena that are inherent to the movement of macroscopic amounts of a fluid. This creates a crude form of memory and a reproduction mechanism stabilizing the convection cells once they emerge. The remainder of this chapter shows that the features we recognize in this elementary example are general aspects of the appearance and development of dissipative structures.

Note 6.2.

Some readers may have difficulty in understanding the relation between inertial forces and a *memory effect. Still the majority of the readers experienced this relation in daily life. If we drive a car, we feel the problem of inertia if we suddenly have to brake. Many of the objects in the car remember their speed before the engagement of the brake, and these objects, including the driver, express this memory by moving towards the windshield of the car during braking.*

6.5. Instability of the macroscopic branch in chemical reaction systems.

In this section, we analyze conditions leading to instabilities of the macroscopic branch in systems where chemical reactions take place. Firstly, we consider a simple transformation process according to an equilibrium reaction. The reaction scheme implies that a resource transforms into a product and that the reverse process also takes place. For readers that are less familiar with the formalism of chemical kinetics, it is also possible to think of the reaction as a transaction in which a product is sold and money is obtained and vice versa.

We analyze this process applying mass action law kinetics (Chapter 2, Roels (2010)). It follows that we observe no instabilities of the type introduced in this chapter for the simple kinetic scheme we describe. In this case the system is stable irrespective the fluctuations that occur.

As a contrast, we introduce a different system; it features a so-called autocatalytic process according to which a chemical substance increases its own rate of synthesis. These cases of positive feedback frequently occur in biological systems and in the socioeconomic system. In autocatalytic processes, fluctuations may destabilize a steady state. This particularly applies if the force that drives the transformation or transaction becomes larger and if the autocatalytic features become more prominent. The change in the kinetic scheme does not change the magnitude of the source of free energy to which the system couples. It just changes the way in which the system interacts with this source of a force. This way of interaction provides the difference between evolving dissipative structures appearing or not.

The cases studied in this section and the foregoing section, identify the following conditions for the development of complex dissipative structures through instability of the macroscopic branch:

- Accessibility of a sufficiently potent source of free energy or economic value.
- Transformation processes characterized by a non-linear dependency of the flows on the forces, particularly also when autocatalysis occurs. This generally also introduces memory effects in which information is stored in some way.
- The existence of fluctuations that continuously test steady states for their stability.

The very evolution of the socioeconomic system on earth illustrates the first point. Solar radiation provides a very potent source of free energy and economic value that drives evolution far beyond equilibrium and all organized structures that appear on earth are a product of this evolution.

In the evolution of the universe, we encounter the second point: Small inhomogeneities and autocatalytic effects of gravity lead to the formation of galaxies assisted by the forces created by the Big-Bang. We discuss this in some depth in Chapter 9.

The third point reflects the inherent complexity of real systems. Fluctuations result from the interactions at the microscopic level, unseen by the macroscopic observer. In fact, unseen may be the wrong expression rather these details are beyond the information the macroscopic observer has. We sometimes experience the effect of fluctuations in a subtle way. If we look at the sky, it appears blue to us, whilst solar radiation is far from blue. The blue results from differences in scattering of the various colors of the light in solar radiation due to density fluctuations in earth's atmosphere.

6.6. Evolution through fluctuation and selection: The general case.

In the preceding sections, we show that a number of quite common conditions cause instability of the macroscopic branch of the evolution of macroscopic systems beyond equilibrium. Complex ordered structures appear unexpected from a reduced information, macroscopic, perspective.

In this section, we introduce some new elements. The first one is competition between the entities present in the system. Competition is an important ingredient of evolution in biological systems and it certainly also characterizes our economies where firms contest markets for goods and services. Competition occurs in an environment where resources, sources of free energy or economic value, are scarce. In that case, interactions between the entities in the system take the shape of a struggle for life in the quest for those sources of economic value or free energy. Situations of pseudo-equilibrium may appear to exist, but certainly in the longer term perspective, the situation is a dynamic one in which new entities emerge and existing ones grow or decay and become extinct. In this section, we follow the discussion of Nicolis and Prigogine (1977) and extend it to economies and markets. Such an approach is not new but a comprehensive discussion of EVT and the role information in relation to the socioeconomic system seems lacking.

As a case in point, we discuss biological evolution in the context of the work of Prigogine and Nicolis (1977). This is relevant to the discussion in this work if we remember that industrial organizations, science, technology, culture, art, economies, nations,... are products of a biological evolution that started eons ago. The source of free energy and economic value latently available in the solar radiation that reaches the earth, fuels this evolution. The literature (e.g. Alchian (1950), Hirshleifer (1977) and Beinhocker (2007)) highlights the analogy between the dynamics of competition for markets and the existence of firms and biological evolution. There also exists an evolutionary approach towards strategic thinking (Nelson and Winter (1982), Dopfer (Ed.) (2005)).

The present-day theory of biological evolution starts from the observation that the information needed to construct, maintain and grow a biological structure rests in a chemical code in the form of large macromolecular chemical entities termed DNA or sometimes RNA. This information represents the genotype of the organism. The code contains four symbols, which in groups of three called codons, provides codes for a limited number (about 20) of structural elements (amino acids) and some editing signals. The amino acids form the monomers for a wide variety of macromolecular structural elements and catalysts. In this way the genetic code, the genotype, of an organism translates into the physical form, the phenotype that allows the functioning of the organism and enables it to compete.

An important process in the multiplication of biological entities rests in copying the genetic information to pass it on to the new generations. This copying process is highly accurate but still imperfect. This is due to the chemical characteristics of the code and the proofreading mechanism that the cell uses to ensure copying fidelity. This imperfect copying is a basic feature of survival value. We return to this later on. Some worthwhile and accessible accounts of the nature of the genetic code provide additional detail (Dawkins (1976, 1987), Eigen (1971)). The above characterization reveals that life is an information-based game and we can extend this to the human organizations encountered in everyday life as we show later.

To highlight the complexity of the information contained in living systems we consider the information contained in the human genome. It equals some 10^{10} bits of information. Construction of all the different DNA-species possible based on this amount of information requires very much more than all matter present in the whole universe. The replication of the genetic code contains errors and hence induces variation in the genotype that passes on to the next generation. This results in changes in the phenotype of the offspring. It leads to sustained evolution of biological systems, an important feature of competitiveness at the species level. Important contributions of Manfred Eigen and Peter Schuster (1971, 1977, 1978a, 1978b) summarize the theory.

We stress that replication errors in the multiplication of DNA are an important driver of sustained evolution. These provide the fluctuations that test the stability of the system. Errors provide, in principle random, variation of the genotype and the resulting phenotype. There is no goal orientation in these variations, i.e. these do not provide a direction to evolution. It is the environment with its scarce sources of economic value shaped by the competition between the actors, which provides an arrow to evolution. A necessary condition for directed evolution is scarcity of resources, scarcity of sources of free energy and/or economic value. This allows selection the most competitive ones out of the variety of phenotypes and associated genotypes. The source of variation exists at the level of the genotype, whilst selection takes place at the level of phenotypes. As soon as there is scarcity of resources, a process of natural selection emerges (Darwin (1859)). In the evolution process, structures develop that, given the resources that exist or develop in the environment and the competition with the other actors, outgrow competing entities and become dominant. This is a dynamic equilibrium with an important historic dimension. Such features also apply to competitive dynamics in industry. An industry structure does not exist but is subject to a constant process of evolution. This is an important consideration in the shaping of adequate approaches to competitive strategy in industry.

As an intermezzo, we return to the limitations of the macroscopic method to account for evolution phenomena. As we discuss later on the copying fidelity is high when we consider organisms containing a large information set, this applies even in simple organism. Hence, the overwhelming majority of the copies of the original code show close resemblance to the original in both genotype and phenotype. Major deviations appear in the tails of the distribution function of the genotype and the corresponding properties of the phenotype. These have a hardly noticeable effect in macroscopic averaging processes. Generally, if a system evolved in a stable environment and a selection pressure in a constant direction exists, the average genotype and phenotype are most likely to survive and macroscopic averaging adequately predicts the evolution of the system. However, if the selection pressure shifts by changes in the environment and/or the appearance of new competitors, the selection pressure may favor the extremes of the distribution function and the average becomes a bad guide for prediction of the further evolution of the system. At these critical points, bifurcation points as introduced earlier, the extremes determine the system's fate and the macroscopic approach breaks down.

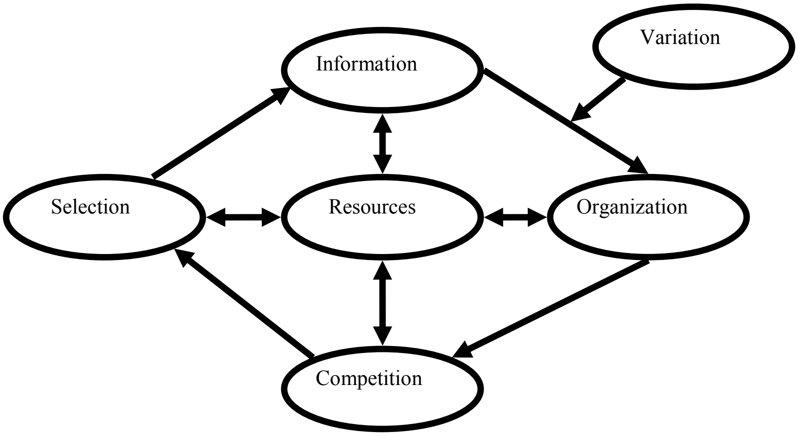

Fig. 6.9. Dissipative structures: Learning systems.

The smooth evolution of the system changes into a revolutionary jump in the properties of the structures appearing in the system.

The picture of evolution, through variation and selection, implies that self-organizing, self-maintaining, reproducing systems already exist. It provides no clues how the first organized structures come into being. The only conclusion the theory allows is that these are the inevitable result of fluctuations that appear at the moment that the macroscopic branch is no longer stable.

Fig. 6.9 again introduces the archetype of the evolution cycle. Dissipative structures exist that are able to store and communicate information about their organization. The information translates into the embodiment of that information, the organization or organism. In the process of copying information, error or variation occurs. For a firm this translation of information results in among others its physical assets, its products and it services and its human resources. The nature of the variation merits some additional remarks. These can be simple errors in the copying or interpretation of the basic information or it can result from deliberate experimentation such as in the research and development based introduction of new products. However, in any instance, reality is too complex to understand it completely and even the best R&D-based new product introduction rests on incomplete information and hence bounded rationality. With its organization, the system couples to and creates sources of economic value and it competes with the other actors in the environment. Selection takes place. This impacts again on the information base of the entities and the process starts all over again. In summary, the essence of the dissipative structure is a developing information set that creates and competes for sources of economic value in the environment.

6.7. The starting point of biological evolution: Prebiotic evolution.

The origin of the first self-reproducing information sets on earth is still subject to scientific debate. Somehow, these appear on primitive earth or these enter out of the universe. One story is the following, whether or not it is completely right is immaterial to the subject matter of this book. The only requirement is that such replicating information sets indeed did emerge.

Some 4 billion years ago, under the conditions on primitive earth, small organic compounds, such as the basic building blocks of living systems, amino acids, nucleic acids and sugars, emerge in small but growing amounts. This view is widely accepted (Miller and Orgel (1974)) and supported by laboratory experiments showing the formation of these compounds in models of the earth's primitive atmosphere. One of the prerequisites is that a source of free energy exists to which the uphill processes for the synthesis of these compounds couple. The abundant source of energy provided by the solar radiation directly or indirectly provides such a resource.

The next step is that these compounds concentrate in small regions of space by adsorption to surfaces or by the evaporation of water from small pools. This process of concentration favors the synthesis of polymeric substances. Some of these polymers exhibit the property of providing a template for their own synthesis; autocatalysis appears on the stage set on early earth. As argued earlier, autocatalysis provides a situation in which the macroscopic branch can become unstable and increasingly complex molecules develop. This provides the precursors for the appearance of the first self-replicating units, units of which today's biosphere provides a multitude of examples. In the subsequent development, the units grow in complexity in the quest for more effective ways to couple to the sources of free energy in the environment and to create new varieties of these resources.

6.8. Competition and sustained evolution.

This section discusses the stage of evolution beyond the point where the first self-replicating units appear. We briefly analyze the behavior of competing populations of self-replicating units.

We assume that the development outlined in the preceding section results in populations of self-replicating units that compete for sources of free energy in the environment and are instrumental in the creation of new sources of free energy. In our system new ways develop to couple to those sources present in the environment or sources of free energy made available by the very action of those self-replicating units. The systems learn to capture free energy and a development away from equilibrium to increasingly complex ordered structures and systems of ordered structures starts to take place. The environment and the structures that appear have the following properties:

- There is an element of scarcity in the available and accessible sources of free energy.
- The structures are metabolized and engage in metabolism, i.e. sources of free energy degrade in downhill processes.
- Some of the information sets in the structures present have the property of self-replication; they provide a mould for their own synthesis. In addition, those information sets develop autocatalytic properties, i.e. they enhance their own rate of synthesis. Increasingly sophisticated ways of storing and communicating information develop. Today's DNA provides a sophisticated and prime example of such an information processing system.
- The replication process is not perfect. The properties of the information sets that derive from the template differ from the mould. New organisms or organizations appear.
- Competition results in the selection of information sets and corresponding structures that more effectively couple to the scarce sources of free energy in the environment.

Eigen and Schuster (1971, 1977, 1978a, 1978b) shows that if the conditions mentioned above apply, only a few types of molecules or structures that directly compete for an identical source of free energy generally survive. This is equivalent to a known feature of biological evolution, where mostly only one organism or a few organisms survive for every so-called "niche". A niche offers one distinct source of free energy. Here we postulate that this also applies to industry structure. We discuss this matter more extensively in Chapter 7. Specifically the molecules, or the structures, with the largest survival value, defined as the largest rate of reproduction, survive.
If a structure is sufficiently complex, i.e. the amount of bits of information underlying the specification of its structure is large enough, sustained evolution results. New, more successful, mutant copies continue to appear. This particularly is the case if the number of possible structures based on variation of the information underlying it, is much larger than the number explored over the lifetime of the evolution. It introduces the concept of sustained evolution and structures of competing entities in a constant process of change (Fig. 6.10).

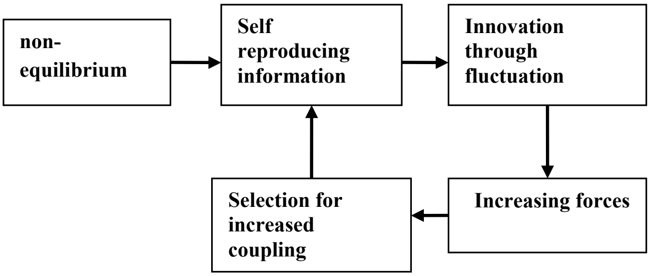

Fig. 6.10. Sustained evolution, evolutionary feedback.

The discussion above makes clear that if sources of free energy exist or can be made available by self-replicating information based structures or organizations, innovations appear that increase the extent of exploitation of the potential value in the system. This leads to increasingly complex and "capable" mutant types of information. This increase in the availability of economic value serves to create more opportunities to fuel the appearance of new generally more complex and more adapted mutant forms of information and their associated phenotypes. The rate of evolution thus tends to increase.

The system evolves further and further away from the unorganized situation characterizing the macroscopic branch. In addition, the minimum dissipation of statistical entropy characteristic for the linear near equilibrium evolution of the system no longer applies. The excess rate of growth of the structures takes its place and the development critically depends on the detailed kinetics at the microscopic level. This is the concept of sustained evolution, also called evolutionary feedback, as introduced by Nicolis and Prigogine (1977).

6.9. The dynamics of competition.

Consider an environment in which competing organized structures evolved because of the appearance of the first self-replicating molecules described earlier.

The next step is to analyze the competition between various organized structures such a supra-molecular structures like organism or, for that matter, organizations such as firms. Clearly, we need to take into account at least the following features:

- Reproduction or communication of information.
- Introduction of variation in information by imperfect copying.
- Translation of the information in the structure or organization corresponding to it.
- Competition between the various organized entities.
- Selection of more competitive entities by competition for scarce resources.

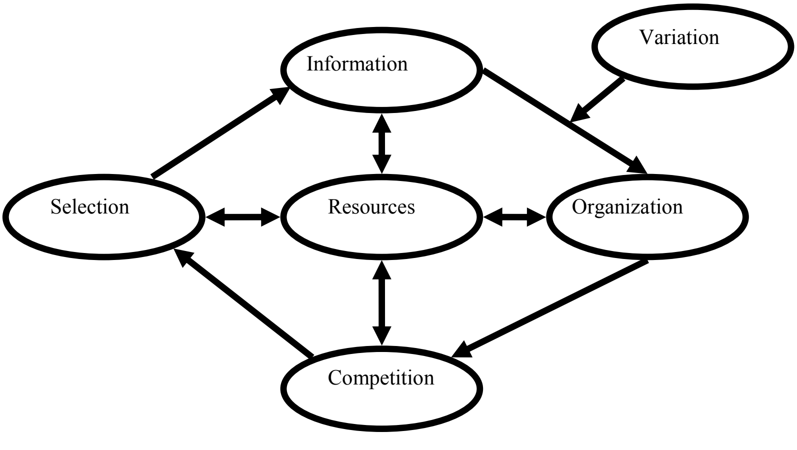

Fig. 6.11. The dynamics of competition.

To illustrate this in more detail we reproduce Fig. 6.8 as fig. 6.11 for the reader's convenience.

The information needed to reconstruct a complex dissipative structure, such as a macromolecule, an organism, a human being, a technology or competence, an industrial organization, an economy, fully characterizes it. An observer of the organized structure, even if he is an observer inside the structure, does not avail of all the information to characterize the detailed microscopic workings of the structure. The statistical entropy of the macroscopic picture quantifies this lack of information. In a living organism, an observer without any additional information would need the amount of information allowing him to select the specific DNA structure from all the possible combinations corresponding to the size of the genome.

Note 6.3

The amount of information needed to understand the details underlying the genetic structure of even a relatively small organism like the enteric organism E. coli is very large indeed. If we take into account a genome size of 5.10^6 *DNA bases, and realize that at each position four nucleic acid "letters" can appear, the amount of information needed would be 10,000,000 bits.*

In the case of an industrial organization, this information is more difficult to grasp. It includes all information needed for the operations of the company, such as the information contained in its products, its captive market knowledge, the blueprints of its tangible assets, the information characterizing its competence and technology base, the information regarding its strategies and future plans. Some of this information exists in a written form or in computer files, some of it is in the heads of its human resources, tacit knowledge, some of it represents cultural aspects of the company.

When considering the human genome we note that it contains 6,000,000,000 bits of information. This corresponds to a choice of one out of the order of $10^{2,000,000,000}$ DNA-base combinations. Even if in a much simpler case, the magnitude of the selection problem becomes apparent. A single hemoglobin molecule, the oxygen carrying protein in blood, consists of four chains of amino acids (polypeptides) twisted together. One of these chains contains 146 amino acids. Each amino acid is one out of the natural 20 that occur in proteins in organisms. The total amount of possible combinations is equivalent to 10^{190}, this compares to the 10^{100} that could simultaneously exist in the universe if the mass present in the universe allows filling it up in a closest packing of molecules.

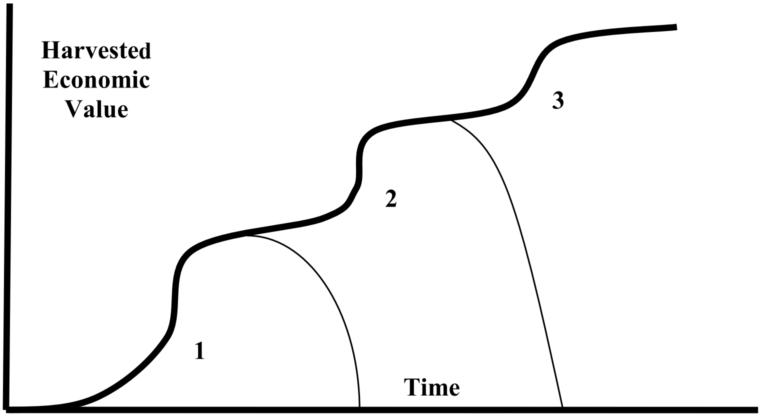

Fig. 6.12. Exploiting a source of economic value by products of increasing sophistication.

Taking the estimated mass present in the universe into account, the number of molecules that can simultaneously exist decreases to about 10^{23} , minute if compared to the total number of combinations possible.

If we extrapolate this discussion to an organism or an enterprise, the significance of the limitations to the magnitude of the amount of information that we can obtain, becomes apparent (Eigen (1971)).

As soon as a potential source of economic value appears in the system, such as the demand for a product, a development takes place by which entities appear that satisfy the demand and feed on the economic value associated with it in an increasingly sophisticated way. Generally, this also leads to increasingly potent sources of economic value to harvest. Typically, a development as depicted in Fig. 6.12 takes place.

Fig. 6.12 depicts a typical life cycle of the way in which the supply of products develops. A single innovative information set triggers a development in which, after it emerges, it goes through a phase of growth into maturity and final decay. This is the trajectory labeled 1 in Fig 6.12. The phases of emergence and growth represent the gradual, evolutionary, phase of the development. The decay is often results from a second innovation going through the same life cycle that replaces the existing one due to its increased sophistication. This development is labeled 2 in the figure. The emergence of the new, more competitive structure, induces an "unexpected" revolutionary change in the evolution of the system. This process repeats again when an additional innovation emerges.

There exist a number of additional complications when we apply the theory to systems of a more involved complexity. The first one stems from the observation that, in almost all advanced systems, the information coding for the dissipative structure and the physical form of the dissipative structure are different entities. The information coding for the dissipative structure acts as a blueprint for the actual form of the structure. This leads to questions of cause and effect. What causes the formation of complex structures? Is it the information or is it the functional structure? This question becomes meaningless once the evolution progresses towards a point where the information carrier and the functional structure divorce. Once the cycle depicted in Fig. 6.11 closes, the evolution becomes truly cyclical and both the information set and the functional structure select simultaneously. Both the information set and the functional structure compete with other information sets and structures. This also changes the environment. Both the environment and the structures become cause and effect. In addition, the concepts of chance and necessity appear in the theory. Evolution of increasingly complex structures is a necessity if the macroscopic branch becomes unstable. If the information coding for the structure becomes large the exact path the evolution takes becomes fundamentally unknown because of the vast number of possibilities. The path becomes unknown to an outside observer that has a reduced information picture; it is also unknown to the actors inside the system as they may have more information but can never obtain complete information.

The divorce of the information and the functional structure is an example of a division of labor type of specialization. If these functions become separate, better-suited structures develop. In fact, division of labor is a characteristic of industries.

A feature that is also of great interest is the balance between stability and complexity. On the one hand, in system of increasing complexity the number of possible innovations that challenge the existing structures increases. On the other hand, it appears that, certainly if the structure has aged, only a large fluctuation in terms of competitive advantage allows displacement of the entrenched structure.

The foregoing discussion highlights the intimate relation between the structures and the environment and vice versa. This introduces a new problem in modeling. One of the first steps in modeling is to decide what the environment is and what the system is. In modeling, we assume the environment given and this assumption becomes dubious if the system and the environment start to interact and become part of a cycle. This results in the need to include an increasing part of the environment in the system, otherwise the modeling exercise becomes futile and leads to erroneous predictions. To the opinion of the author, this complication applies to many of the models used in the analysis of the socioeconomic system. Our discussion also reveals that sustained evolution becomes inevitable in an environment where sufficiently large sources of economic value exist or develop. To complete this chapter we briefly return to biological evolution and its relation to elements of human society such as industries and economies. We also discuss the relevance of exogenic evolution.

6.10. Biological evolution: Dissipative structures.

In the preceding sections, we discuss the origin of life and the appearance of molecules that, through some crude memory and replication capability, sustain their non-equilibrium structure and create and use sources of free energy or economic value in the environment. The analysis highlights autocatalysis, self-replication, information storage and communication, as key drivers in addition to scarcity of resources leading to competition and selection. It shows that the organized entities have two basic functions. Firstly, they contain the information, the genotype, for the construction of their own structure, the phenotype. Secondly, the structures translate the information into the vehicle allowing interaction and competition with other structures and other aspects of the environment.

As an intermezzo, we return to the time when Darwin first publishes his theory of evolution. In those days, the present perspective on the role of the information sets in the mechanism of evolution does not exist. The role of the information carrying genome becomes clearer after the work of Mendel (Mayr (1985)) and finds its culmination in the discovery of the nature of the information carrier DNA. In the days of Darwin, two dominant approaches to biological evolution exist. These are the Darwinian perspective and the Lamarckian approach (Dawkins (1987)). The combination of the theory of Darwin and the genetic evidence on the role of information carrying genes leads to the Neo-Darwinian synthesis (e.g. Lewin and Foley (2004)). Today we know that the information that codes for the phenotype of an organism passes on from the parents to the offspring in the process of reproduction. The genetic code that passes on to the next generation provides a blueprint for the developing offspring. Of course, in higher animals a process of teaching and learning influences the development of the offspring in addition to the information the genome contains. However, these so-called acquired traits do not feed back into the information that transfers to the next generation through the process of reproduction. As far as acquired traits are concerned each new generation starts with a clean sheet. Of course, after the birth of the new generation processes of teaching and learning are instruments by which the collective acquired information conserves in the developing offspring, this represents the contribution of exogenic evolution to the development of the species and for that matter its culture and technology.

From the Lamarckian perspective, acquired information does feed back into the information set the parents directly transfer to their offspring, i.e. the new generation acquires it at conception. In this way, exogenic evolution does not result from other means of communication of information than the process of reproduction. Of course, when we adopt the broader definition of the genotype including the transfer of the exogenic part of the information set, the Lamarckian perspective again enters the picture, albeit involving a way of communication different from the original ideas of Lamarck.

It is important to note that the transfer of information from parent to offspring is the dominant but not the exclusive mechanism for the communication of information in the evolution of life. It is commonly indicated by the term vertical transfer of information. There exist, particularly in bacteria, but also, albeit to a lesser extent, in the more complex eukaryotic organisms, a mechanism of so-called horizontal transfer of information between organisms beyond the transfer in the reproduction cycle. This involves communication of information between organisms of the same species or even differing species. We will not go into detail about the variety of mechanisms that developed in biological evolution, we only want to stress that such mechanisms exist. This matter is relevant when discussing the differences and similarities between the evolution of organisms and organizations like firms, particularly when discussing the relevance of mergers, cooperations and acquisitions as a driver of industry evolution.

The first crude replicator molecules combined the functions of transfer of information and being the vehicle that competes for resources in the environment. Quite early in the evolution of life on earth, bacteria trace back at least 3 billion years, the functional structure and the information carrier become different chemical entities. These molecules successfully discover the beauty of cooperation in the quest for and development of sources of free energy. In biological systems, nucleic acid polymers of the DNA and RNA types take up the role of information carrier and processor.

Microorganisms, plants, animals and humans, i.e. the overwhelming diversity in the biosphere on earth, are products of the versatility of the genetic code and the translation process.

The mechanism that leads to the sustained evolution of the biosphere largely rests in the creative power of the infidelity of the copying process when combined with competition for scarce resources resulting in selection. The imperfect copying of the code leads to a constant exploration of the vast diversity of structures that can derive from the coding mechanism. In this way, new structures constantly appear and challenge the existing structures. The interaction with the environment, both in terms of resources as of other structures that compete, decides whether a mutant copy replicates faster than the mother copy. If it happens to replicate faster, it gradually but inevitably replaces the mother copy and its functional structure. One could say that the codes engage in a gaming or experimentation process in which, by learning by doing, more optimal ways develop to take better advantage of the opportunities in the environment and to create new opportunities. The environment starts to co-evolve with the structures.

An aspect of the process described above is that the coding versatility of even a limited stretch of genetic information allows the creation of far more structures than can be tested in the lifetime of an evolution, even in the case where bacteria appeared on earth earlier than 3 billion years ago. The room for further evolution is therefore endless and new structures, unexpected to the observer, continue to appear.

The fidelity of the reproduction of the code is, although of limited accuracy, still high. This means that the mutant species that develop from the mother copy inhabit only a limited part of the space that contains all possible copies. A mutant copy will be selected and replaces the mother copy if it outperforms the mother copy. The likelihoods that this will happen is, given the rather high copying fidelity, limited in a short time horizon, but it increases in a longer term perspective.

In the way described above many of the species that we observe in the present biosphere and that evolved in the past, largely disappeared or will disappear in the future. In the early biological evolution, the coding function of DNA was the main source of storing and communicating information.

The divorce between the molecules active in storing and transmitting information and those involved in the embodiment of the functional structure proves by no means the only specialization trick the biosphere has up its sleeve. A new approach results with the invention of the brain that, e.g. in mammals, equips organism to store, process and communicate information by other means than the genetic hardware in the nucleic acids. This allows organism to adapt their behavior and to learn beyond the limitations of that genetic hardware.

This innovation further develops with the appearance of the ancestors of humankind. After a while, these develop a much larger brain than the species from which they evolved. Relatively recently *Homo sapiens* appears, with a brain size of about 3 times that of the earlier ancestors, such as *Australopithecus africanus.* The human brain greatly enhances the possibilities to analyze and understand reality. It introduces new ways of storing and communicating information with the emergence of spoken and written language and later on computers. This revolutionizes the so-called exogenic evolution of the human species and its society.

As an intermezzo, we refer to the concept of the "meme" as Dawkins (1976) introduces it in his inspiring work "The Selfish Gene". Dawkins consider the meme as the exogenic analogue of the function of the gene in DNA based evolution. The meme is just as the gene a self-replicating information carrier that competes for a scarce resource allowing its replication and communication. He applies this to explain the evolution of cultural aspects of human society beyond the DNA based communication of information and consider it a next generation of self-replicating information just as we do in this work.

The brain makes it possible to develop tools and machines and is instrumental in the creation of science and technology. In addition, culture, arts and firms and economies, are products of this exogenic evolution. This resulted in new functional entities that are no longer part of the human body. Their contribution, however, to the competitiveness of the species is as real as the clutches and teeth of the large cats. It is a vital of evolution, just as that embodied in the DNA molecules.

The analogy extends further. The further evolution of our culture, including markets and firms and science and technology, follows the same general rules as the early stages of biological evolution. In fact, these are an instrument of further biological evolution. Human culture thrives on a new kind of dissipative structures. In addition to the information stored in our DNA, the information stored in our brains, the information stored in written forms, the information transferred by the spoken word, the information stored in computer systems are all part of the new information sets on which competitiveness relies. This information forms the basis for the creation of new functional structures that create and exploit sources of economic value in the environment. In fact, we learn to harvest economic value, in principle always available, but inaccessible to the more primitive structures of the past. New ways of communicating information develop in teaching and in scientific publications, to mention prime examples. Also for these complex structures, the competitive environment shapes new, more successful, sets of information. Sets of information evolve and increase in sophistication by learning by doing and scientific understanding and the resulting R&D activities that become a hallmark of the academic world and modern industry. Science based research in industry emerges in the 19th century and increases in importance ever since.

Our information about what it takes to compete optimally in economic value space is never complete and it is impossible to specify the required information set with certainty. A considerable amount of information is lacking, as it is possible to access only a limited information set. There definitely exists a large uncertainty, i.e. significant statistical entropy characterizes our knowledge of the relevant reality. This leads to a situation in which taking risks is a necessary element of success. We live in a "no guts no glory" kind of reality.

The important point we reach in our discussion in this section and earlier sections is that biology as well as human culture exist of dissipative structure whose functionality derives from captive information that allows more or less successful competition for potential value in the environment made available as economic value for the more informed actors. In fact, this also holds for industry structure. Our industries are dissipative structures that thrive on and develop information to compete effectively. We further develop this perspective in Chapter 7.

6.11. Conclusion.

In summary, information is the prime resource that allows the creation of economic value from sources of value in the environment. This quantifies in terms of the statistical entropy of the picture of reality of the various actors. It leads to a situation in which forces exist or are created based on asymmetries in information and different perspectives on economic value. This allows transformation and transaction processes to take place that lead to the generation, growth, maintenance and decay of dissipative, information processing, structures. Biological evolution is a prime example and we generalize it to apply to large parts of human society. Limited fidelity copying, or alternatively phrased experimenting with the primary information code, shapes new dissipative structures better equipped to compete. New exogenic ways of storing progressing and communicating information appear and become a prime characteristic of the structures in human culture. Information and its transmission and perfection shows to be the prime competitive tool.

The process of evolution is not, or only limitedly, goal oriented both at the level of the global environment and at the level of the individual actors. This certainly largely applies to pre-human biological evolution, but also applies to entities like firms. Firms generally have explicit goals, but their fundamentally limited information leads to elements of uncertainty and risk that preclude strict goal oriented development. One can “roadmap” a strategy, but a roadmap is of limited use whilst walking in a swamp. However, evolution proceeds in the direction of extracting more of the value globally available as useful economic value. Bounded rationality and the intrinsic characteristic of the dynamics of complex systems, may lead to loss of stability. Ups and downs in the economic value extracting process characterize a system in which interrelated dissipative structures operate in a competitive way in an environment that also is dynamic and in addition co-evolves with the system. An element of crisis and decay is as certain an aspect of sustained evolution as periods of sustained growth at the global level.

We revisit these matters in the next chapter in further developing the theory of competition and selection and further analyzing the nature of the firm.

CHAPTER 7. THE FIRM AND INFORMATION PROCESSING.

7.1. Introduction.

In the preceding chapters, we identified the forces driving maintenance, growth and decay of dissipative structures. Firms and other economic institutions are examples of such structures. These structures take advantage of and create sources of economic value in the environment and couple to the resulting forces in an increasingly effective way. This chapter develops a description of competition and selection. It allows closing this chapter with a discussion of the nature of the firm and the relation, correspondences and differences between biological and economic evolution.

We borrow our approach largely from developments in physics of some 30 to 40 years ago. Our analysis derives from the work of Eigen and Schuster (1971, 1977, 1978a, 1978b). This treatment does not directly apply to firms and markets. We mainly introduce the material to show some general features of evolution under competitive pressure. This leads to observations that also apply to economic systems. Then we introduce an approach based on EVT using the concept of the linear value transducer as discussed in Chapters 5 and 6. We close this chapter by explicitly discussing the nature of the firm from the perspective of EVT. In addition, we highlight correspondences and differences between evolution in the socioeconomic sense and biological evolution.

7.2. The dynamics of competition and selection.

Eigen introduces the notions of metabolism, reproduction, mutability and competition. These features lead to the development of dissipative structures and support their sustained evolution. The work of Eigen refers to the evolution of biological macromolecules and their corresponding biological structures. This does not directly apply to the socioeconomic system. The information storage and processing mechanisms as well as the translation into structures that compete in the environment differ.

Metabolism expresses the need for a source of free energy or, in the socioeconomic system, economic value in the extended sense in which it includes free energy. In Eigen's work, it is a source of the monomers making up DNA or RNA. Eigen describes the competition of the resulting macromolecular species for those monomers if supplied at a limiting rate. This scarcity of resources is a necessary requirement for evolution to take place.

Reproduction involves reproduction of the genomic macromolecules into new copies that build new organisms. As discussed in Chapter 6, in biology there is a clear separation between the information storage and processing function of DNA and RNA and the largely protein based biological structures that engage in competition.

Mutability refers to the copying fidelity of information sets. Copying is not perfect and it changes the information sets in an unpredictable way. Imperfect copying is a requirement for evolution. If the copying is perfect, evolution comes to a halt. We credit Darwin (1859) and Wallace for the development of a theory explaining biological evolution. Of course, in the days of Darwin the role of the information carriers DNA and RNA was not known.

The terminology of Eigen introduces, as said, the notions of metabolisms, self-reproduction and mutability. In discussing Eigen's approach, we keep the mathematics to the minimum required for the purpose of this book. The original literature provides additional detail.

Metabolism expresses the need for a source of free energy, or in our broader interpretation economic value, that the structures can exploit. Darwinian systems connect to value at a low level of statistical entropy and value at a higher level of statistical entropy. The force due to the difference in free energy or economic value, allows the dissipative structures to couple to this opportunity to maintain and grow their structure against the forces of the second law. As we argued earlier, Schrödinger (1945) highlighted this as a characteristic of living systems.

Self-reproduction reflects the fact that dissipative structures in the non-linear range store and communicate the information that underlies their structure that constitutes their ability to compete for sources of economic value. As the structures are intermediate between a source of high economic value and a sink of low economic value the structure specific information is not perfectly stable, it degrades by the action of the very forces that allow its creation and further development. The structures need some form of autocatalysis to fight these degrading forces.

Mutability refers to the fact that perfect fidelity copying of information does not apply. Reproduction errors are a prerequisite for sustained evolution. Errors that provide new information sets are the mechanism that continuously challenges established information sets. This introduces competition as a final element in the treatment. Scarcity of resources leads to a situation where competition drives evolution through selection.

The conditions mentioned above are not only necessary but also sufficient for evolution.

Note 7.1.

A first feeling for the dynamics of this type of evolution results if we analyze an elementary example. We consider a system with an information set that consists of six digits having a value between zero and six. Initially the six numbers are zero. The survival value of the system, i.e. the value that is the object of the selection, is the sum of the six digits. It varies between 0 and 36. After each basic time interval, the information set of the system reproduces. A fraction of the symbols of the information set of the system reproduces exactly but some reproduce erroneously. In case of incorrect reproduction, we insert a random number between 1 and 6.

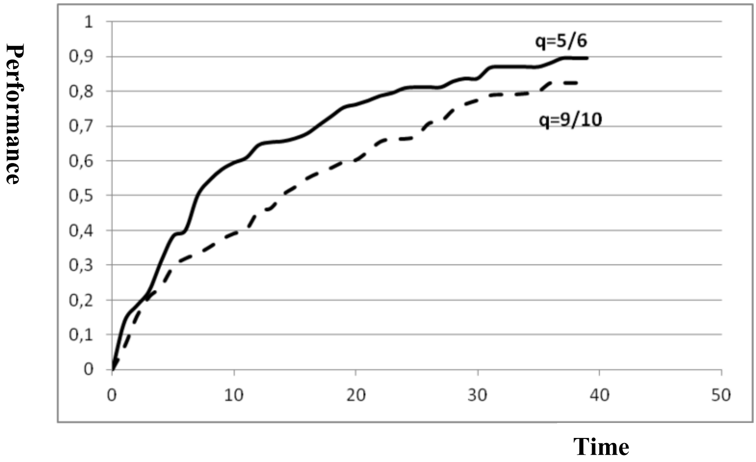

Fig. 7.1. Simulating “Darwinian selection” using an elementary model.

We mimic this example by e.g. the throwing of a die. After each reproduction cycle, we compare the survival value of the new copy with that of the mother copy. We retain the mother copy if its performance is equal to or better than that of the derived copy. We retain the derived copy if it is superior. This is a very sharp selection for improved copies. Fig. 7.1 shows the average of ten experiments for two values of the fraction, q of the symbols that reproduces correctly. We show the performance relative to the maximum of 36 as a function of time.

This admittedly simple model shows interesting features. There exists clear goal orientation of the combination of random mutation and sharp selection. We discuss in Chapter 6 that this is a feature of the evolution of complex information based systems. The performance steadily increases. This is very different from the case in which no selection takes place; then the performance fluctuates randomly around about 0.6. Additionally, the model shows that evolution takes place more quickly if the error rate is higher. This obviously makes sense; if the error rate increases, we try more varieties of the copies per unit time. However, if the error rate increases beyond a defined limit, different features emerge. This is the phenomenon of the error threshold. We discuss this threshold in Section 7.3.

Eigen presents a mathematically rather intricate analysis of this model of Darwinian evolution. We do not reproduce the mathematics here. My earlier work (Roels (2010)) presents a concise summary. The analysis of Eigen leads to the following results. Darwinian selection leads to a final situation in which the reproduction rate of the most competitive copy of information largely defines the reproduction rate of the population present. This reproduction emerges at the level of the phenotype. In biology the information sets do not compete in a direct way, the phenotypic translation of that information set is the competing vehicle. What we finally obtain is not only the most competitive information set. In fact, the total fraction of the most competitive copy in the population finally selected, is relatively small if the information set is large. The population that results exists of a "cloud" of closely related information sets, termed quasi-species by Eigen. However, there exist also situations in which largely different information sets with comparable properties at the level of the phenotype, i.e. information sets with the same selection value, coexist in the final situation. The population can flexibly respond to changes in the environment because information sets better adapted to the new conditions may already exist in the "cloud", the quasi-species, and quickly become more dominant if the environment so requires. If, as an example, the information set contains 4500 symbols and the copying fidelity is .9995, the relative abundance of the most competitive copy can be 10% or less even if the superiority of the ideal copy is relatively high. Thus, 90% or more of the population exists of copies that contain at least one symbol different from the ideal copy.

We can make a few additional remarks regarding the dynamics of Darwinian evolution. If the prerequisites for evolution discussed in this chapter and Chapter 6 apply, evolution towards a maximum reproduction rate is inevitable. However, especially in the case of large information sets, the actual path the evolution takes depends on historical, chance determined, events. Situations develop in which evolution apparently comes to a halt at a level of the reproduction rate that does not correspond to the optimal situation allowed by the characteristics of the information set. Evolution captures the information set in a local optimum and only relatively large fluctuations drive it away from that local optimum to a higher reproduction rate optimum corresponding to a less related superior information set. We regain the concept of sustained evolution. We also have to take into account that for large information sets reaching the final situation may take a long time. Hence, we cannot exclude that in that time the environmental conditions change to favor another information set. For large information sets and a rich environment, sustained evolution prevails and the process of perfection never ends. This does not imply that, once evolution results in a sufficiently sturdy local optimum, innovative information sets easily displace the quasi-species characteristic for the local optimum. There is a definite first mover advantage.

7.3. The error threshold in evolution.

In the preceding section, we discuss the dynamics of evolution of information sets and related phenotypes using the approach of Eigen. The analysis also results in an important constraint to the copying fidelity of information sets. It relates to the amount of information that the set can maintain and communicate. There exists an intuitively logical restriction to the error rate if we want to maintain and progress a stretch of information of a given size. If the error rate is too high, the direction of evolution towards superior copies of the information set disappears. Making too many errors causes the message to lose its coherence. The math shows that the limit weakly depends on the superiority function defined as the excess reproduction rate of the copy of information with the largest selection advantage. It is present as a logarithm. However, it shows strong dependence on the error rate in reproduction. The size of the DNA that shows directed evolution in Darwinian selection proves inversely proportional to the error rate.

For ease of discussion, we assume that the selection advantage of the most superior copy exceeds that of the next best set by a factor of about 3 (or more precisely a factor *e*, being the base of the natural logarithm). In that case, a copying fidelity of .99 allows maintaining a message of maximum 100 symbols. If we add two nines to the copying fidelity, i.e. assume it equal .9999, a message of 10,000 symbols just survives. Thus maintaining the genetic information of *E. coli*, with a genome of $4.\ 10^6$ symbols, requires an error rate below $2.5.\ 10^{-7}$. The human genome of $3.\ 10^9$ DNA bases, requires an extremely high copying fidelity.

This observation results in an important feature of the development of dissipative structures that employ a large set of information. Computer experiments on maintaining a given copy of information under a selective pressure result in the following picture. If the length of a message is beyond the copying fidelity limit, the message quickly disintegrates and the information "melts" away. This is also the case if the target message initially is the only one present. If the copying fidelity exceeds the limit, a very different picture emerges; the mixture evolves quickly to the sequence with the highest competitiveness and its closely related companions in the quasi-species. The optimal quasi-species emerges as dominant even if initially not present at all. The speed of evolution increases with decreasing value of the copying fidelity unto the copying fidelity limit. If the copying fidelity gets lower than the limit, the optimal quasi-species does not emerge or, as said, even melts away if initially dominant.

The process of evolution under a selective pressure shows to be very effective. If we return to the example (Note 7.1) of a six-digit code and a copying fidelity of 5/6, a stringent selective pressure results in a code close the optimal one of 6 sixes after 15-20 attempts on the average. If no selective pressure exists, it takes on the average more than 40,000 attempts before we come close to the optimal sequence. The "learning by doing" behavior clearly pays out.

Analyzing the strategies that evolved in biology leads to the following picture. In nature, the copying fidelity is close to the limit required to maintain the genome given its size. Thus RNA-based viruses (phages), genome size of 1000-10,000 bases, allow an error rate of close to .001 to .0001. The prokaryotic bacteria, typical genome size of 5.10^6, allow an error rate of the order of .0000001. We can formulate this observation in a different way: The copying fidelity of the information set of an organism determines the genome size it can maintain. The trend in evolution to higher genome sizes requires the invention of increasingly sophisticated copying mechanisms. One can speculate that, when reaching the genome size of the higher animals and the ancestors of humankind, the further expansion of the information set on which competitiveness rests requires another strategy. The capabilities of the brain, allowing exogenic evolution, complement the potential of DNA based evolution. Part of the information storage and communication divorces from the DNA molecules and exogenous information, not bound to the physical DNA, appears on the evolutionary stage. Still later, information storage and communication increasingly divorces from the genetic code as mechanism such as language, writing, science and education, academic and industrial research, take their role as information storage and processing mechanisms complementing the role of DNA and the brain. Firms are examples of information processing organizations and the rules of information transmission subject to error (or experimentation) we discuss in this section, apply to these institutions.

7.4. Models of Darwinian systems: The Hypercycle.

This section analyzes more complex systems that replicate and are subject to selection through competition. These more complex systems exhibit features that appear in structures such as organisms, markets and industries.

Evolution on earth results in complex chemical machineries in which a separation develops between information and function. Complex information sets, RNA or DNA molecules, are responsible for storing and communicating information. Protein structures result from the translation of this information into a complex biological molecular machinery. These functional structures compete for scarce resources and the feedback leads to evolution of the information sets. This leads through a process of mutation and selection to increasingly competitive structures. As this discussion reveals, the information set and the functional structure interact to form a closed cycle.

We want to reemphasize the reasons for such separation of information and function. These are manifold but one of the most basic ones is that the coding and reproduction and competition functions pose very different requirements. It is unlikely that one molecular species optimally combines these requirements. Alternatively phrased, separation of these functionalities leads to new sources of competitive advantage and it inevitably appears in the competition for scarce resources.

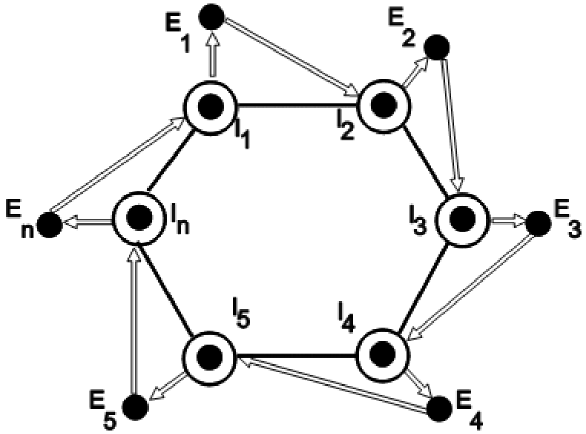

Fig. 7.2. A schematic representation of a Hypercycle.

Of course, separating the functions in different entities leads to coordination problems, e.g. their production and interaction needs to be fine-tuned. In biology, the separation of functions proves necessary when the structures become very complex and need large chunks of information for their instructed creation. In functional biological systems, this leads to types of organization that we can model by "Hypercycles". We will discuss some of the interesting properties of these systems. The work of Eigen presents a detailed discussion.

Fig. 7.2 provides a schematic representation of a simple Hypercycle. The Hypercycle consists of a number of information sets I_i. These entities carry only a part of the total information set of the Hypercycle. The size of these information sets is below the error threshold resulting from the copying fidelity limit. This guarantees their conservation against error copies, of course with the exception of copies that increase the competiveness of the structure. The information sets have a self-reproduction capability indicated by the open circles in Fig. 7.2. The functional structure resulting from the information set directly preceding an information set catalyzes this

reproduction process. An important feature results from closure of the Hypercycle, i.e. the product of the last information set in the system catalyses the synthesis of the first one. If the cycle is not closed the structures in the Hypercycle do not cooperate rather they compete and concerted action fails to result.
Hypercycles show at least the following properties:

1. The overall Hypercycle exhibits autocatalytic growth. The elements of the Hypercycle grow in a concerted fashion if sources of free energy, or economic value in our extended concept, in the environment allow this. This is one of the requirements for sustained evolution. Different Hypercycles engage in competition for scarce resources. They exhibit Darwinian competition and selection.
2. Hypercycles show non-linear kinetics leading to strong selection behavior. A Hypercycle, once established, resists substitution by other emerging cycles. Their substitution requires a substantial selection advantage of the challenger.
3. Its strong selection behavior allows it to evolve quickly and to exploit small differences in selective advantage. It is very effective in improvement through learning by doing once established as a closed loop.
4. The cyclic arrangement allows the system to use more information than consistent with stability in the light of the fidelity of the copying mechanism used (see Section 7.3). The hypercyclic cooperation allows escaping the fidelity limit.
5. The system selects against so-called parasitic branches, i.e. branches that attach to the Hypercycle and replicate with it, but do not contribute to its competitiveness. In addition, parts of the cycle that cease to be functional and do no longer contribute to the competitiveness of the overall cycle, automatically disappear if this results in increased competitiveness.
6. There is an advantage for the system to escape into a closed compartment. In this way, it can evolve and use pieces of information to which its competitors have no access. This also results in protecting itself against pieces of information that evolve elsewhere and pollute the cycle. We recognize a known feature of organisms that have cell walls and membranes. This also is a feature of human organizations, such as firms. Firms generally have defined well-policed interfaces with the environment. In such organizations, restrictions exist to the exchange of materials, resources and information with the environment.
7. Individual Hypercycles do not cooperate but compete. They may link, however, resulting in larger functional entities in which two or several Hypercycles cooperate. Their cooperative rather than competitive behavior critically depends upon the strength of coupling between the Hypercycles. This mimics processes of fusion and alliance in industries. It also resembles critical stages in biological evolution that depend on the merger of separate organisms into one functional biological entity.

Hypercycles, although rather complex from the mathematical point of view, present a highly schematized and simplified picture of the reality of markets in which industries compete. Still, these mathematical abstractions show a rich variety of interesting features.

7.5. Competition and selection: An approach based on EVT.

For simplicity's sake, we discuss competition using the linear rather than the non-linear value transducer. Staying with the linear transducer makes life easier and still leads to relevant conclusions.
In Chapters 5 and 6, we discuss the linear value transducer. Here we introduce some further

features. We define the power output of the value transducer. It is proportional to the product of the flow at the output side and the corresponding conjugated force. We optimize the ratio of the force at the output of the transducer to that at the input of the transducer to result in a maximum power output. This is desirable from the perspective of generating as much economic value as possible per unit time. As shown elsewhere (Roels (2010)) the maximized power output depends strongly on the degree of coupling (Section 5.7) of the output of the linear value transducer to the opportunity provided by the economic value force in the environment. In fact, the power output shows proportionality to the square of the degree of coupling. In addition, it is also proportional to the square of the economic value force to which it couples.

A company can derive a higher profit from a need in the market in two distinct ways. The fist one is straightforward. It develops an information set that allows it to couple more effectively to the source of economic value, i.e. it realizes a higher the degree of coupling. The other situation is slightly less obvious. Companies may differ in the statistical entropy and/or the costs of information and hence create a different economic value from a potential value opportunity of a given size. Also these latter differences lead to increased competitiveness and hence potentially higher profit. This latter source of competitiveness results from an information advantage due to a superior genome, or broader, a superior information set.

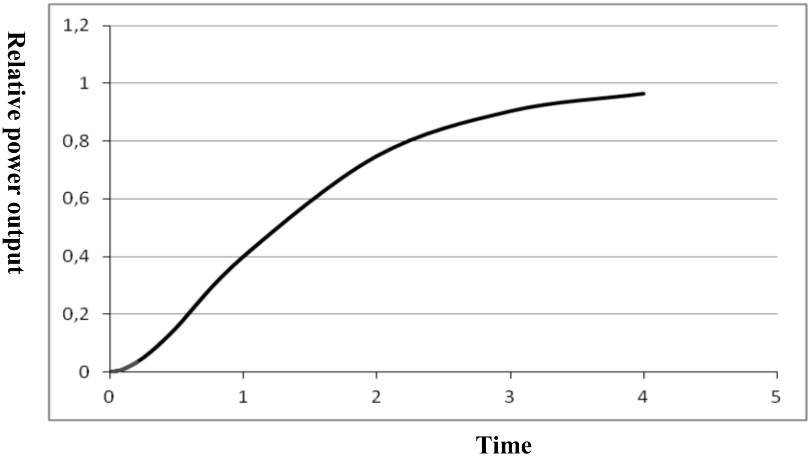

Fig. 7.3. Life cycle characteristics based on EVT approach.

We introduce an elementary learning by doing system. It is the linear value transducer optimized to achieve a maximum power output given its degree of coupling. We use this as a model of a firm competing for a source of economic value. We assume that the nature of the firm’s information set allows it to develop a maximum degree of coupling by further optimizing the set, either purposefully or through learning by doing. Initially, the degree of coupling is zero. We also assume that the rate of increase of the degree of coupling is proportional to the difference between the maximum value of the degree of coupling attainable by the firm and its present value. If we do the mathematics, we obtain an evolution of power output or profit against time. The mathematically inclined reader can consult my earlier work (Roels (2010)). As fig. 7.3 shows, we obtain the familiar life cycle of a dissipative structure (Section 6.9), or for that matter an industry, from this elementary model.

We now consider two value transducers, firms, competing for one source of economic value defined by a force in the environment (Roels (2010). The difference in their initial information sets gives them different values of the maximum degree of coupling they can realize. In addition, their rates of learning by doing are different.

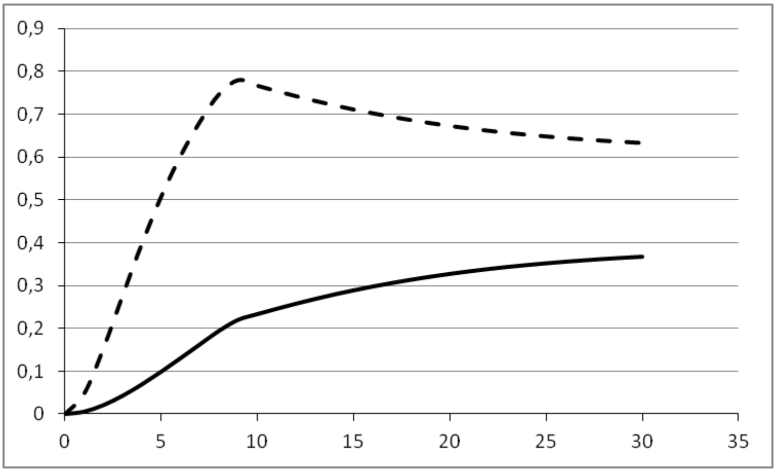

Fig. 7.4. The development of the normalized power outputs (Y-axis) against time of two competing value transducers.

This is an admittedly crude model of competition, but in the author's view, it serves to identify the types of behavior that we can expect.

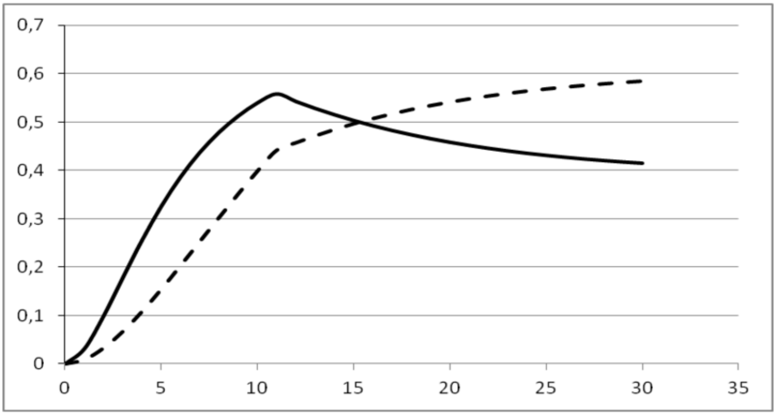

Fig. 7.5. As in 7.4 with different parameter values.

Figs. 7.4 and 7.5 show the results of some simulations based on the model. Fig. 7.4 shows that the faster evolving information set with a higher potential degree of coupling quickly grows its share of the market and starts to feel the pressure of competition after growing through a maximum in profitability. It remains the dominant player.

In fig. 7.5, the faster evolving set with a lower potential degree of coupling is dominant in the early stages of the evolution of the market. Later on, the slower evolving set with a higher potential degree of coupling becomes dominant. We can tentatively explain the picture in fig. 7.5 as follows. The quickly evolving species is a smaller information set that allows a lower copying fidelity and hence more variation in its information set. It learns more quickly. The size of the set, however, does not allow it to reach the maximum degree of coupling of one. The second set is a larger information set with a higher copying fidelity requirement. It evolves slower, but its amount of information allows reaching the maximum achievable degree of coupling of 1.

7.6. The nature of the firm and its evolution.

We conclude that firms are dissipative structures beyond the strictly linear region of value transduction. Firms evolve as both a consequence and a cause of the existence and development of sources of economic value. In this respect, firms are a generation of dissipative structures beyond biological organisms. Firms develop store and process information beyond the DNA macromolecules that drive biological evolution. Organizations are an inevitable product of sustained evolution in the non-equilibrium environment provided by the biosphere. They emerge and grow or decay coupling to economic value forces, which in turn also emerge because of their activities. Competition for scarce sources of economic value feeds back into their information set, which more and more adapts to operate optimally under the conditions in the environment or they decay and more adapted information sets emerge in evolution. Markets and industries, and an industry structure, do not exist but emerge and evolve. Firms result from a non-equilibrium situation they serve to maintain and grow.

A second condition for organization exists in non-linearities in the interaction with the environment and with other firms. Autocatalytic behavior in which an entity enhances its own growth is an example. In addition, there is the need for an information storage, processing and communication system. Firms are information processing structures that contain the information needed to localize and/or create sources of economic value in the biosphere and to produce and sell the products on which their competitiveness rests. Finally, the reproduction of the information set needs to be subject to error or experiment to improve the effectiveness of the operations and the products that characterize the firm. This reminds of the dynamic capabilities approach to the firm summarized in Douma and Schreuder (2008). These authors stress that the operational capabilities of a company change due to goal-orientated changes by the management of the firm. Here we maintain that the distinction between error and deliberate experiment in changing the information set on which the firm's competitiveness rests, is only gradual. Reality is too complex to grasp in detail and there exists only a reduced information picture of reality, characterized by a significant uncertainty or statistical entropy. There is only bounded rationality in the changes the firm makes by, as an example, the R&D based development of new technologies to supply new products and to develop new processes. The competition and the interaction with the environment decide whether the changes that management introduces result in the goals pursued. Bounded rationality applies to the adaptation of the information set and environmental selection provides the arrow to the evolution of firms and industries. This dynamic evolution does not lead to any social optimization of resource use; the system optimizes towards maximum competitiveness of individual firms. This strongly resembles the results of contemporary evolutionary theory of organizational economics (Nelson and Winter (1982), Nelson (1987), Douma and Schreuder (2008), Beinhocker (2007)).

What is the nature of firms? Firstly, we define the market in which it operates. The firm produces products that supply a need in the market. This need translates into a force in EVT. Through its products, the firm couples to that force and derives economic value (profit) to fuel its operations. We can model this in terms of the value transducer discussed in Chapters 5, 6 and the preceding section. There exist other firms that couple to the same need in the market with the same product or a different one that supplies the functionality to satisfy the need. These firms are the competitors in the "niche" of the firm.

The most basic characteristic of the firm is its information set. It consists of information for the production of its products and market information to direct R&D's search for new processes and products. It contains technological competences, organizational procedures, Human Resources policies, the knowledge of the people employed by the firm, its public affairs approach and many things more. The information set form a blueprint for the operational capabilities of the firm. In addition, the information set contains the processes to change the information set, e.g. through business planning and corporate planning or R&D. Changes induced in the information set, can be either purposeful, e.g. as a reaction on the experience with its products in the market, or arise by error or misinterpretation. The interaction with the environment determines the direction of development of the information set of the firm. In the biological metaphor the information set is the "genome" of the firm, although nucleic acid based macromolecules do not store and process it.

The products the firm derives from its information set are instruments to couple to the need in the market and to extract economic value and these are the direct basis of its competitiveness. There are beyond the physical products and the services additional "products" of the information set that engage in the competition in the market, such as the image of the firm, its perceived financial solidity and its marketing activities. These collective aspects of the firm's competitive strength are the "phenotype" in biological terms. These products compete in the market and provide feedback to direct the evolution of the information set. The products in the broad sense defined above, determine the competitiveness of the information set and drive competition and selection.

Not all the elements of the information set need to be internal and captive to the firm. If it has access to other information sets and integrates these with its own set the synergy involved may lead to an increased competitiveness. Even freely available information (e.g. in the public domain through academic publications), often leads to an increased competitiveness on integration with the firm's information set. This latter point, the information set as a whole defines competitiveness, is very important. If we look at the genome of an organism, it consists of, in the perspective of the whole genome, small pieces of information called genes. These code for a certain functionality at the protein level. The collective proteins define the activities and shape of organisms in a complex way; the whole is far more than the addition of the activities of single genes. A gene that decreases competiveness if inserted in one set of genes, may improve competitiveness in another set of genes. The same holds for the information set of an organization such as a firm. The whole information set determines whether the addition of a piece of information, freely available or captive to the firm, is instrumental to increase competitiveness.

In evolutionary approaches to organizations it appears that the resistance to changing information sets is large, there are inertia (Douma and Schreuder (2008), Chapter 10) that result in resisting changes. This resembles the copying fidelity threshold we discussed in Section 7.3. If too frequent and too large changes in the information set take place, it does not develop in the direction of increased competitiveness. Rather a process of meltdown of the information set occurs. There is no longer an arrow of evolution. In an industrial organization, this results in a limit to the rate of change of the information set of the firm. It must be below a certain threshold to prevent "melt-down" of the firm's information set. This also explains why, especially for large firms, it is by no means trivial to cope with a changing environment that requires drastic adaptations to its operational capabilities.

The firm has a boundary with the environment, just as the membrane and the cell wall in organisms. This need not be an identifiable physical boundary, but it can also result from the procedures around communication and publication of firm specific information. In this respect, we stress that the firm's information set loses its value if it diffuses into the environment. The integrated set must be unique to the firm, although, as indicated, elements of the set may be freely accessible and still contribute to competitiveness. Protection of information may also take the form of legal instruments such as patenting law.

The selection theory we discuss, leads to the conclusion that the final situation in the exploiting a given source of economic value, a given niche in the jargon of biological evolution, only supports one dominant information based structure or structures that very closely resemble each other in competitiveness. In general, this leads to increasing concentration of industries that exploit a given niche, certainly if the life cycle of the market develops into maturity. This proves to be a known feature of many industries. An example emerges in the food industry, where in food ingredients such as thickeners and flavors, the number of players per Product Market Combination decreases consistently and strongly over the years.

As an additional element, we discuss the life cycle of the phenotype. Certainly in the higher organisms, the phenotype does not live forever nor do the products of a firm. Through an embryonic phase, subsequent growth and maturity, it ages and finally dies. This does not apply to the genotype. It lives on in a new generation of the organism. Dawkins (1976) introduced the notion that the phenotype serves as a "survival machine" for the genotype. In fact, the genotype is immortal; it survives, albeit as part of an increasing complexity. The same holds for the information set of the firm, it survives the products that the company produces and may increase in complexity over the years. The company's competences survive many generations of products and are precursors of wholly new products. Of course, parts of the company's information set becoming obsolete and no longer contributing to competitiveness, may gradually disappear. This observation also implies that planning of the development of the competence base, the information set, of a company often requires a longer term perspective than the planning of the development of its products and services. This in an important consideration in the development of e.g. the research and development strategy and the strategic planning in industrial corporations. We substantiate this further in Chapter 15 when we discuss the evolution of some of the leading firms today.

As a final characteristic of the firm, we revisit the concept of information work that we identified in Section 5.2, when discussing the overall EVT formalism. There we speculate that economic work, resulting in the creation of economic value out of potential value, is an important aspect of profit creation in an economy. This points to a very important capability that resides in the firm's genome, its information set. This capability allows the firm to engage in various types of information work resulting in a decrease of uncertainty and hence a lowering of the statistical entropy to increase the economic value it can derive from a source of potential value in the demand for products and services in the market. Examples of such information work are R&D activities, market research, business planning and strategic planning.

7.7. Differences and similarities between biological and economic evolution.

We stress that evolutionary theories of firms and markets do not derive by analogy to biological evolution. Evolution is an inevitable feature of complex systems that operate away from equilibrium, are processing and communicating information, and compete for scarce resources in transformations and transactions that exhibit non-linear kinetics. This situation applies to both economic organizations and biological organisms. Both types of organization follow the pattern that derives from the systems theory of evolution. This is the main area of similarity between economic and biological evolution.

Furthermore, there is the aspect of separation between the information set that ultimately defines competiveness and the vehicle with which competition for scarce resources takes place, the firm's products in the broad sense discussed in the preceding section and the organism's protein based phenotype. This distinction between the genotype and the phenotype applies to both economic and biological organizations. In addition, the circumstance that the genotype survives the phenotype applies.

Another important similarity is the concept of the error threshold. If a system "experiments" too much, the information set melts rather than evolves in the direction of increased competitiveness. To the author's opinion and he feels support in the literature (e.g. Douma and Schreuder (2008)) this applies to both biological systems and firms and other economic institutions. The reverse is also true, if an organism perfectly reproduces its DNA, evolution comes to a halt and responding to changes in the environment is no longer possible. Experience shows that organisms seem to operate close to the error threshold and introduce the maximum admissible level of variation in the genome. This also is relevant for economic institutions. If the restrictions to modifying the information set on which the company operates are too strict, it is no longer able to improve the competitiveness of its products and is no longer able to react on competitor moves and other changes in the environment. This becomes particularly important if drastic changes in the environment challenge the position of the firm.

There also are important differences. The present view in biology is that mutations in the genotype that drive evolution are almost wholly random. In firms, there is an element of design, but we indicate that bounded rationality applies due to the complexity of the environment, the competition and the internal processes inside the firm.

A second point is that the DNA of an organism is part of its product, the next generation phenotype. This is not the case in the products of firms. These do not physically contain the complete information set of the company. Furthermore, the information set, the genome of the firm modifies without its physical replication. This is partly different in biology, where the DNA replicates to produce a new generation of the organism. Of course, also in biology mechanisms developed avoiding replication for transmission of information. The brain and its derived products like science and technology lead, also in biological systems, to exogenic evolution and introduce information that not directly replicates in the next generation of the organism.

Another important difference relates to the mechanism of communication of the information set. In biological systems, the communication of the information contained in the DNA is largely restricted to the offspring of the parents (vertical transfer of information). This most certainly applies to the higher organisms, in bacteria mechanisms of transfer other than by reproduction have existed and do exist. We refer to the mechanisms of horizontal gene transfer discussed in Section 6.10. In exogenic evolution communication of information to organisms or organizations other than the direct offspring, takes place. This may shift the driver of evolution from the level of an individual organism or organization to groups of such entities. This of course applies to aspects of the evolution of the socioeconomic system such as the evolution of science, technology and culture and for that matter the socioeconomic system with its markets and firms.

Finally, but with some hesitation, there is the aspect of the speed of evolution. The information set of the firm appears to evolve far more quickly than the pace of biological evolution. This certainly applies to the higher organisms, but less to bacteria and notably viruses.

Additionally, we again address the methodological and philosophical problem highlighted in Chapter 2 and in Chapter 4 when discussing the Capital Asset Pricing Model. An important assumption is that the model and the system are independent. The fact that the model exists should not influence the outcome of the transformations and transactions in the system. If a model becomes part of the system and particularly if the players in the system accept the predictions of the model, the assumption of independence of system behavior and the existence of the model breaks down. This is a significant methodological and philosophical problem. It is akin to the problem in quantum mechanics, where the assumption that observing the system does not influence its behavior breaks down. The problem identified above does of course apply to all theories of socioeconomic systems not just to EVT based approaches.

7.8. Conclusion.

We developed a consistent macroscopic model of the socioeconomic system, EVT. In this chapter, we use it to indentify the nature of the firm. Our model rests on a reduced information picture of a complex reality, as almost all models of relevant phenomena in physics. It identifies the forces that drive transactions and reveals their statistical background in the concepts of statistical entropy and the cost of information. We also argue that the socioeconomic system is a system beyond equilibrium, where significant forces exist and/or result from the activities of the actors in the system. This leads to the conclusion that the socioeconomics system behaves according to a systems theory of evolution that applies to all system beyond equilibrium where the forces exceed a critical limit. This evolutionary concept predicts that, in such systems, organizations appear that extract economic value from the forces present or the forces that can be created, by a process of coupling to these forces. We particularly identify systems that store and process information with a high but not perfect copying fidelity as an outcome of evolution and a source of further, sustained, evolution. The nature of the fluctuations in the system, i.e. deviations from the averaged properties used in the macroscopic description, determines the development of the system at critical branching or bifurcation points. This introduces a historical dimension in the development of such systems. The problem is that, by the very nature of the macroscopic approach to modeling, information about the fluctuations is not present in the model. Hence, the exact future evolution is not predictable, even if we fully characterize the system in a macroscopic sense. The evolution of the system is subject to chance and necessity. In the system evolution to more complex organizations, resulting in the extraction of economic value in increasing amounts, inevitably takes place. We cannot predict, however, the path of the future evolution, although we can say a few things about likely evolutionary patterns based on further assumptions about the system's behavior. In this book, we avoided going into detailed models of the socioeconomic system and this choice should be understandable based on the reasoning we present. This problem does not apply to the socioeconomic system only, but is also present in the engineering sciences. A book of the present author Roels (1983) that describes macroscopic methods for the description of processes involving the activities of microorgapic witnesses this. Also microorganisms are far too complex to describe in detail, but useful predictions result from macroscopic modeling and careful experimentation. Also in that field too complex models, untraceable experimentally, have little practical value. The additional problem in the socioeconomic system is that the freedom of experimentation as in systems in which microorganisms appear, does limitedly apply.

In the foregoing, we stated the limitations and prospects of the macroscopic description. We can predict that evolution will take place, we can generalize some features of the evolution in the direction of increased competitiveness, but we have to remain silent about the path of evolution in an individual case.

The problem becomes clear if we look at the picture of human evolution (Lewin and Foley (2004), Chapter 11). Some 5-7 million years ago, our ancestors abandoned their tree dwelling habit and developed walking on two legs as a new way of earning a living. This allowed them to develop some primitive tools to assist them in earning that living. They developed hunting and got access to high quality food in the form of meat. This allowed and induced the further development and sophistication of the brain, which consumes about 20% of a human's total energy budget at only 2% of body weight. This later on, a few million years ago, leads to an increase in the size of their brains by a factor two and later on three compared to the early ancestors. Animal husbandry and agriculture appear some 12,000 years ago. Language and other forms of communication develop further; exogenic evolution takes place at an increasing rate. Today we witness our socioeconomic system as an outcome of those developments. Was there any chance of predicting these developments when our ancestors appeared?

CHAPTER 8. SOME CONCEPTS OF ECONOMIC THEORY.

8.1. Introduction.

This chapter presents an overview of selected aspects of contemporary economic theory that are relevant to this work. In view of the main theme of this book, we highlight the role of information in economic analysis. We introduce microeconomics and the perfect competition model and selected aspects of the macroeconomic perspective. The discussion of models of economic growth and business cycles introduces aspects of the dynamics of economic development. In addition, we analyze the reigning evolutionary approach to economies, firms and markets in the perspective of this book's main theme. Furthermore, we introduce transaction costs economics and the agency theory of the firm as well as an approach to strategic management.

8.2. The nature of markets and industries.

We propose the model for markets and industries depicted in Fig.8.1. This model serves as a stepping-stone for the discussion of selected aspects of economic theory. Industrial activities derive from the fact that the products needed by society are not directly available in the environment. In addition, the need for specialization to efficiently collect or produce the products the market demands, importantly contributes to the appearance of firms and the markets they support. Finally, there are significant economies of scale and scope that drive the emergence of industrial activity.

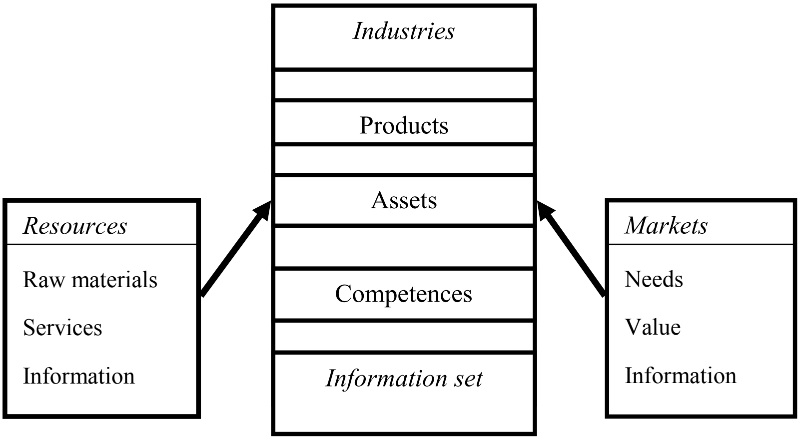

Fig. 8.1. The firm as an information processing entity.

"Hunter-gatherer" strategies, in which individuals obtain the products to supply their needs directly from the environment, support the evolution of humankind for an extended period of time. In the last odd 250 years, industrial activity progressively takes over in the supply of products to support market needs. Conversion of resources to the products needed, initially in an artisanal way, later on increasingly in industry, starts to dominate the economy. In this evolution industrial activity depends increasingly on distinctive assets and competencies (such as refined technological capabilities) to produce and supply in an efficacious way. Industry uses (captive) information and associated assets and competences to source materials, information and services from the environment and to transforms these into products and services that supply a market need. We highlight the central role of (captive) information in this stage. This transformation process adds value as perceived by the buyer. If the firm operates well the buyer pays more than the economic value the firm spends to produce the product. In this way, the firm obtains a profit. It can invest part of this profit to generate new competences or assets and to develop new processes and products. The firm thus extends the information set that supports its operations. This causes the information set of the industry to adapt to the challenges posed by the environment and the market. In this model, industry emerges because of the needs that exist in the market. Industry develops the information set needed to supply products that satisfy those needs efficiently.

Another important concept that shapes industrial activity is competition. Generally, several firms are willing and able to supply products that satisfy a need. These firms compete for the economic value associated with that need. The firms are not equally efficacious and because of competition, some firms grow their share of the market whilst others are less successful. In general, the firms exploit different information sets, i.e. there are asymmetries in information. In addition, shaped by the forces of competition an evolution of the industry takes place, it is driven by changes in the environment (e.g. availability of resources), changes in the needs in the market and the evolution of the collective information sets of the firms operating in the industry. This leads to the familiar life cycle of an industry or a market (Fig. 8.2.) as we first encountered it in Section 6.9 when discussing the systems theory of evolution.

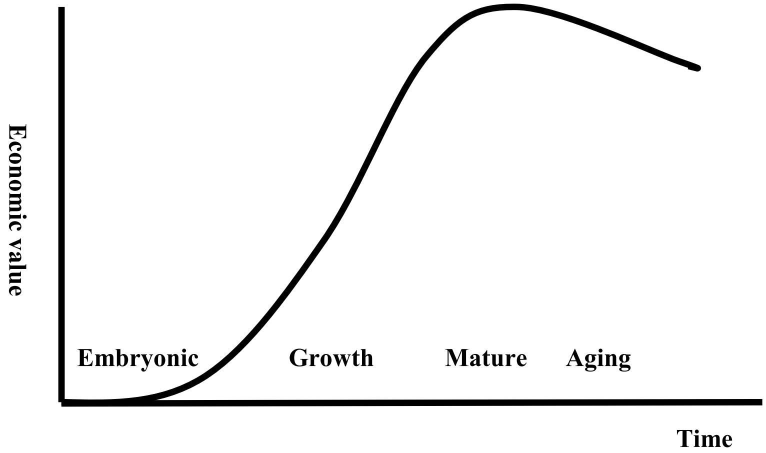

Fig. 8.2. Industry evolution: The life cycle.

An industry supporting a market emerges because of an innovation in either the resource base of the industry, the competences or assets base, the way of supplying the target need, changes in the needs in the market or a combination of these. This triggers the so-called embryonic phase of the industry in which the first firms appear that supply products and services to support market needs. In the subsequent growth phase of the industry, the forces of competition shape an evolution in which a process of learning by doing triggers the development of increasingly efficacious information sets. This results in more sophisticated ways of satisfying the target need. The market grows in terms of economic value. In the growth phase, both the development of better products (product innovation) and the development of better processes to produce a given product (process innovation) take place. Both elements of innovation derive from the creation of competitive advantage through developing increasingly sophisticated information that is, at least partly, captive to the individual firms that compete in the market. In most cases, the number and the diversity of players in the industry increase in the growth phase. A phenomenon called evolutionary radiation in biology.

In the further evolution of the industry, when we enter the mature stage, further optimization of products and processes becomes increasingly difficult. A unit progress needs more and more effort. The growth of the economic value of the market stagnates. The number and the diversity of the players decrease by forcing out the less effective ones. Finally, the industry may enter the decay phase. This often results from an innovation that changes the way the target need is satisfied. Sometimes the target need disappears. Fundamentally better or cheaper concepts, substitute the existing products. An important feature of this model is that an industry structure does not exist in the way Porter discusses this in his influential publications (Porter (1980, 1985)). Industry structure is a dynamic consequence of evolution under the forces of competition.

Throughout our analysis, we use the term industry for a group of products and underlying industrial activities that supply the same need in society. Innovation that leads to the decay of an industry does not necessarily mean that the existing players decay with the industry. One of the leading players in the old industry may pioneer the new approach. However, often a period of fundamental innovation gives opportunities to new entrants and poses a threat to the companies entrenched in the old industry.

8.3. The nature of value and value transaction processes.

We introduce a theory of the creation of economic value (EVT) in Chapter 5. This allows us to identify the forces that drive the evolution of economies, industries and markets. We reiterate the main aspects of EVT to arrive at a self-contained treatment of the economic aspects in this chapter.

Firstly, we note the distinction between potential value and economic value. Potential value is the “real” value of an asset or a product. Economic value is the value to an actor that has only limited information about the asset, i.e. does not have the full information needed to pinpoint the value microstate of that asset. Only a limited amount of the potential value represents economic vale, i.e. value available to perform economic work. The development of asymmetries in information, leading to different perspective on economics value, is the essence of innovation, competition and progress.

The macroscopic description introduces the uncertainty about the exact microscopic state of the system. Hence, also the future evolution is subject to uncertainty. This uncertainty emerges as a direct consequence of the reduced information description of the system. Statistical thermodynamics quantifies the uncertainty as the statistical entropy of the reduced information description. Macroscopic thermodynamics and its extension into a general Economic Value Theory shows that in systems away from equilibrium so-called dissipative structures appear because of self-organization phenomena. This development brings evolution of complex structures in an initially unstructured system within the realm of macroscopic theories. An example of such a system is the biosphere on earth including human society with its markets and organizations such as firms.

8.4. Self organizing systems: Dissipative structures.

Macroscopic thermodynamics leads to the conclusion that an isolated system that does not exchange material or energy with the environment, evolves to a final state of thermodynamic equilibrium. At thermodynamic equilibrium, the state in terms of the reduced information macroscopic description does no longer change. We no longer observe macroscopic changes

in the system.
Equilibrium is a state in which, given our macroscopic information, we reach maximum uncertainty about the microstate of the system. The lack of information of the observer reaches a maximum. The second law highlights this. The second law predicts the direction in which the evolution of a system takes place in the eyes of a macroscopic observer. In the previous chapters (e.g. Chapters 5 and 7), we extend the realm of the second law to include the description of economic transactions and organizations such as firms. We characterize biological systems, markets, economies and industrial corporations as examples of dissipative structures. The existence of galaxies, the earth, life and biological evolution, human civilizations, economies, industrial enterprises and markets shows that evolution proceeds towards increasing complexity. This is in apparent conflict with the second law that predicts an evolution towards maximum disorder or decreasing organization and complexity. Many early investigators assumed that life, human society and evolution belong to a class of phenomena that escapes the second law. The material we highlight so far shows that in systems that are not isolated but exchange energy sources and/or sources of value with the environment, an evolution takes place in which the development of order and organized structures becomes a direct consequence of the second law. What it requires is leaving equilibrium. We developed a systems theory of evolution that shows that, if the distance from equilibrium becomes larger, increasingly sophisticated organizations can and will appear and evolve in a process of sustained evolution. The ordered structures that evolve have, as said, been termed dissipative structures. Dissipative structures can indeed only evolve and persist if these exchange resources with the environment. In addition, dissipative structures are both a product and a source of non-equilibrium and constantly evolve rather than exist in a given state. A book of the Nobel laureate Prigogine (1980) titled "From Being to Becoming" highlights this.

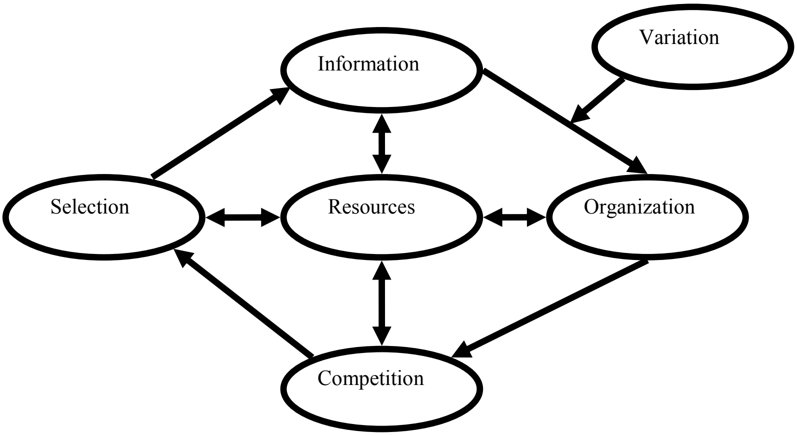

Fig. 8.3. Evolution through change by chance: The learning cycle.

Evolution of ordered structures of increasing complexity is a necessary consequence of the second law. "Order out of Chaos", as it has been termed by Prigogine and Stengers (1984) in their illuminating book. A sufficiently large exchange of energy, economic resources and information that creates a sufficiently large distance from equilibrium is a precursor for evolution. The basic force that feeds industrial firms is the need for products in the market.

Processes taking place in the system characterized by non-linear kinetics e.g. autocatalytic phenomena increase the likelihood of the evolution of organization. Competition, growth and decay are examples of processes with such characteristics. The system contains information about its own structure and its operations. Limited fidelity copying of this information is an important driver of evolution. Copying errors or "experiments" in which the information is deliberately changed, lead to the limitations of the fidelity of copying. The combination of these points leads to "learning by doing" as a characteristic feature of dissipative structures. We arrive at a full description of the drivers of evolution including evolution of the socioeconomic system.

Fig. 8.3 that we introduced earlier summarizes the characteristics discussed above. The central entity that drives the cycle depicted in fig. 8.3 is a (captive) information set. It codes for the organization of the structure in both its tangible and its intangible aspects. The organization contains the assets, again both tangible and intangible that allows it to produce the products and services that allows the entity to contest the market. It also contains the management structures that coordinate the activities in the organization. This process of competition leads to selection at both the level of the organizations and their information sets. This in its turn leads to an evolution of the collective of information sets of the organizations supporting the market need. In addition, the environment in which the markets and industries evolve is very important. It is a source of resources, the source of the market needs that allow organization to appear and subsist and provides the macroeconomic environment in terms of e.g. rules and regulations to which the firms is subject. An important feature is the phenomenon of co-evolution of the organizations and the environment in which they appear. In fact, also policymaking, such as government intervention, is subject to co-evolution with the economic activities they intend to regulate. Such policies are part of a cause and effect cycle in which the distinction between causes and effects is blurred.

8.5. The pivotal role of information.

It is not by chance that information appears at the top in Fig. 8.3, it is our way of expressing the prime importance of information in making our economies, markets and industries work. In Chapter 4, we highlight the role information from the perspective of information theory and derive the concepts of statistical entropy as a measure of uncertainty or lacking information. In addition (Chapter 5), we highlight the crucial role of information asymmetries. Information asymmetries result in the driving force allowing creation of economic value out of potential value. We see asymmetries in information as a creative force rather than a nuisance hampering the adequate working of the market as a coordination mechanism in the economy. We further discus this nuisance approach to informational asymmetries in a while.

The first thing we have note about information is that, as many other macroscopic state variables its definition in an absolute way is not possible, we can only define it with respect to a reference level. We can only define its change. This means that we need a reference level for information. How do we define such a datum level? We refer again to our ultimate resort: The heat engine of Carnot. We know that the functioning of the heat engine depends on having a source of energy that provides heat at a temperature higher than the temperature of its environment. Heat in the environment is a commodity that is both freely available and useless for doing meaningful work. If we consider the conceptual equivalency of heat and information we reach the conclusion that information that is "ambient", i.e. is freely available in the environment, cannot create economic value. Therefore, the information available to all players in an economy is useless regarding its capacity to do economic work and becomes a convenient reference or datum level for our definition of the state variable information. It also follows that information beyond that information is scarce in the definition of scarcity in an economic sense. It has a value and comes at a cost, and there is a definite minimum to that cost, the cost of information in the economy. Scarcity of information allows creation of economic value. Economic value is an example of a quantity subject to intrinsic scarcity. We highlight the intrinsic uselessness, in an economic sense, of information that is available to everybody.

Contemporary economic theory (see e.g. Douma and Schreuder (2008)) defines two model types of coordination. The first one is the market in which prices provide sufficient statistics, i.e. provide all the information necessary to complete transactions between the buyers and sellers in the market. The second one involves coordination as a necessary consequence of the specialization and the division of labour that are some of the root causes of the emergence of firms. In ideal markets, all necessary information is available to all players and prices settle to the economic value that is characteristic for an equilibrium transaction. We discuss this concept in Chapter 5. It is the situation in a perfect competition world where all information is available to all players and all players use that information in an optimal way. Everybody pays a fair price and nobody experiences economic gain. Freely available information is equivalent to heat at ambient temperature in the heat engine metaphor; it has no capacity to do (economic) work.

Organizations like a firm do not use the price mechanism underlying perfect competition as a coordination mechanism. Coordination rests on information that is, at least partly, captive to the organization and forms the very essence of the need for specialization and division of labour that leads to the need for coordination by mechanism other than the price mechanism. Douma and Schreuder (2008) argue that organizations arise as a solution to problems of an informational nature. We agree with this view and complement the remark. Organizations, such as firms, serve to operate optimally in situations in which asymmetries in information exist or can be created. In such case, learning by doing under competitive forces automatically creates organizations by a process that we indicate with the term evolutionary feedback in Chapter 6. This leads to a more or less circular argument. In absence of asymmetries in information, activities that result in economic gain are not possible. The drive for economic gain leads to the need to create informational asymmetries. This necessitates forms of coordination and organization other than markets. Hence, organizations inevitably appear and lead to a further need for, and opportunities to create, additional asymmetries and new organizations. Whether informational asymmetries create organizations or organizations create informational asymmetries is of questionable relevance once the circle between the two closes. Organizations and the associated informational asymmetries constitute a creative power in economies. These are the divers of economic growth through converting potential value into economic value.

We further explain this by introducing a metaphor of the ideal, perfect competition, market: The heat engine. For the heat engine, temperature is the equivalent of information. We note that Carnot argues that the temperatures of the hot and the cold side of the heat engine provide sufficient information, sufficient statistics, to characterize the heat engine. The ideal market is a market in which all players have the same information, equivalent to being at the same temperature. The heat engine equivalent of this situation is a contraption in which the temperatures at the hot and the cool side are equal. Such a heat engine, or such an ideal, perfect competition market, never produces any useful work or in our metaphor economic work. From the engineer's perspective, there is nothing ideal about the perfect competition market or at least its thermodynamic counterpart.

We conclude that asymmetries in information are the source of the vital forces behind economic growth. We handle the concept of these informational asymmetries in that positive way.

Of course, asymmetries in information can also lead to less desirable situations. The book of

Douma and Schreuder (2008) provides examples of this. We cite some examples from these authors' book. One example refers to the fundamental paradox of information. The advantage information offers disappears if fully disclosed. This can lead to economic disadvantages when we need pieces of information of two parties to allow society to reap the benefits of using the information. An example in the pharmaceutical industry occurs when one party, e.g. a small start up company, avails of technology, i.e. owns information that in principle allows the development of a cure for a severe disease. The small company does not avail of the funds and knowledge to progress it to the market through the complex set of regulatory rules characterizing the pharmaceutical market. In that case, legal instruments, like contractual agreements covering cooperations and joint ventures or patent law are helpful. Such sets of regulations are the institutional and legal framework the macroeconomic environment provides.

Another known information asymmetry problem is opportunistic behaviour, also called strategic behaviour. It may lead to business practices judged unfair. Here aspects of the legal framework and the reputational aspects of the workings of the market come into play. In addition, monopolies may prevent adequate clearing at fair prices of free non-regulated markets and a regulatory framework must come to the rescue. In some cases, public interests make relying on the market and organizations with private characteristics less desirable. This is the case for important parts of the socioeconomic infrastructure such as roads and railways. There government can play the role of a public provider. We stress that these regulatory interventions prevent the workings of direct competition that directs evolution in the direction of efficaciousness.

Adverse selection may apply in the health insurance market. Here a company could end up with more than its fair share of risk because of the average health expectation of the people that choose to buy insurance. No doubt, there are many more examples but we close this section with the famous example of the used car market where asymmetries in information between sellers and buyers exist that may lead to excessively low prices for used cars.

8.6. The internal value chain.

If we consider a medium sized or large firm, it is fair to say that the information set on which such firm operates in the market is very large indeed. Classical equilibrium pricing theory suggests that sourcing this information in the open market makes most sense. The transaction costs formalism (Williamson (1975)) (Section 8.11) considers the firm and the market as alternatives for sourcing the vast amount of information the firm needs.

In Section 7.3, we discussed the problems in maintaining and processing of large amounts of information. As the information set becomes progressively larger, the copying fidelity of the information has to increase. This leaves less room for optimization of the operations of the firm by experimentation or error. It also results in inflexibility and little room for adaptation by changing the information set in response to changes in the environment. This problem is clear in many large corporations, particularly also in the development of new products or the development of new businesses. Many large firms observe that venturing firms, small flexible emerging entities, are more successful in developing new businesses in the early stages of development. This is clear in e.g. the pharmaceutical industry, where the advent of genetic engineering in the early seventies of the 20th century, triggers a number of start-ups that pioneer the development of new pharmaceutical products based on this new emerging technology. Successful examples are Genentech, Genzyme and Chiron. Ultimately, large pharmaceutical houses, already entrenched in the pharmaceutical industry, acquire or take a controlling interest in many of these companies. Chiron becomes part of Novartis and Roche takes control of Genentech. This illustrates the difficulties in replacing entrenched players in an advanced stage

of the evolution of an industry.
Many large firms try to resolve such difficulties in coping with radical innovation, by creating internal entities geared towards new business development or, as said, acquire venturing type units. Royal DSM is an example of a company that created an internal venturing unit.

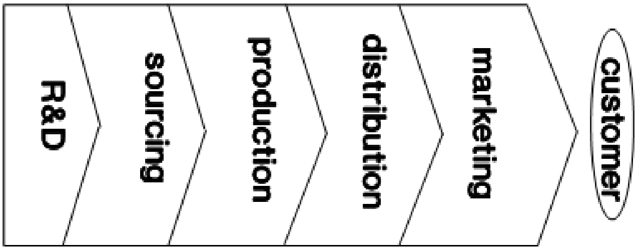

Fig. 8.4. The internal value chain.

We describe a general solution to the problem of handling large sets of information discussing the Hypercycle in Section 7.4. In essence, the Hypercycle arrangement consists of a number of smaller units that cooperate in a larger functional structure. By splitting the overall information set in smaller pieces that allow more effective experimentation with the information set, a higher ability develops to adapt these sets if the environment so demands. One has to realize that the Hypercycle only works if the units within the cycle cooperate rather than compete. As we discussed, this results if the Hypercycle closes, i.e. if the product of the last information set in the cycle catalyzes the production of the first information set.
Fig 8.4 shows a situation where the activities of the firm split up according to functional departments. This arrangement is one of the common approaches in industrial practice. The functional departments work together in developing, producing and delivering the product to the customer and sourcing information and other resources from the environment. This mitigates the difficulties in managing a very large information set. The information sets of each of the functional units are smaller. This allows more possibilities to fine-tune the sets of the various subunits to the requirements of the environment by experimentation. The arrangement leads to at least the following questions. The firm's management faces the question, which functions to organize within the Hypercycle that makes up the firm and which functions to source externally. In the firm, the functional units no longer directly compete with units that have the same function in the outside world. Hence, the inborn optimization of effectiveness that characterizes evolution no longer works. The functional units may be less than optimally effective. Here lies an important task for the overall company management. The advantages of having no transaction costs involved in securing reliable supply of the function must compensate this problem. In addition, the advantages of captive information may be the key differentiating factor. In general, one expects that strategically important functions, of which the information set differentiates the company from its competitors, are not likely to be candidates for outsourcing. Functions necessary to operations of the company, but not differentiating with respect to competing entities, are more likely candidates for outsourcing. However, this distinction between strategic and non-strategic functions is not as straightforward as it seems. As an example, consider the discipline of analysis in industrial R&D. Most of the basic analytical tools are readily available in many institutes or academia. However, using these tools in support of the research activities of a company requires availability of detailed knowledge about the strategic objectives of the research at hand, and requires frequent interaction and sensitive captive knowledge. In such case,

in house facilities can and will sometimes lead to a decisive competitive advantage. It is one of the many examples where, apart from the basic science, the context in which we deploy the science becomes crucial.

A question that needs an answer relates to the type of action that we need to secure cooperation within the Hypercycle in such a way that it serves the best interests of the company. If we assume that the individuals working in the company are to a certain extent seeking the optimization of their own utility function, working in the best interests of the company is not always likely to occur. Somehow tuning activities in the company to optimize the longer-term profitability of the company needs to take place. The company can achieve this, as an example, by developing a remuneration system that rewards the units according to their contribution to the overall profitability of the company. This is not always easy in a way perceived fair. Units may differ in terms of contribution to long-term versus short-term profitability. Another way in which such coordination can result is if we appeal to so-called altruistic behavior. This develops if the individual players consider it in their best interests to adhere to company goals rather than sub optimization of their narrow minded own interests. The literature argues (Simon, in Dopfer (Ed.) (2005)) that such mechanisms of organizational identification, serve as a mechanism to assure effective coordination.

A further potential problem with a split up of information sets over functional departments within a company rests in handling situations where multidisciplinary decision-making requires inputs from the various functional departments to allow identification of the most adequate actions. Different functional departments tend to speak different “languages” and effective communication is far from trivial. This is notably the case in the formulation of effective overall strategies. Leading firms in the industry spend important efforts in time and money to allow multidisciplinary teams to develop a common picture of reality to assist strategic decision-making.

Another way to cope with internal complexity issues of the firm rests in the creation of divisions or business units. The firm splits up in a way considered the most effective way of organization in view of reducing complexity and problems with the optimization and adaptation of increasingly unwieldy information sets. This splitting process can be of a variety of natures. A split up according to the customer group it serves may be instrumental, e.g. in food and personal care directed units. In addition, a split in terms of products groups, such as Tea, Ice cream, Oils and Fats is a possibility. A split up according to position in the external value chain, i.e. base chemicals, fine chemicals and specialties is another approach. Alternatively, the underlying competences, such as enzyme technology and fermentation technology, can be the organizational principle. In all these cases, splitting up results in coordination problems of the types mentioned above. In addition, the question is whether the synergies involved in having these entities as part of an overall firm lead to preferring this arrangement above sourcing these functions outside the firm. The synergies in having the products and the required competences internally have to outweigh the costs of complexity and the problems of coordinating diverse information sets. In this respect, the food industry witnessed an increasing tendency of organizing ingredient supply in firms other than the companies operating in the consumer products markets.

8.7. Macroeconomics.

Macroeconomics considers the economic system at a level of aggregation of a region, a country or even the whole world. Macroeconomics comes close to the subject matter of this book. It is important to the question how the results of analysis at the microeconomic level, i.e. at the level of markets and firms, serve to direct the evolution of economies in a socially beneficial and orderly way. Macroeconomics studies indicators such as the Gross Domestic Product (GDP), unemployment rates, national income, consumption, inflation, interest rates, savings, investments (both public and private), innovation (again also at the aggregated industrial level and public through academia and public institutes), taxes, etc.. Macroeconomics is a broad field of study. A number of aspects that are prominent in the discipline bear a close relation to our field of study. These notably are the business cycles and the drivers of economic growth.

The business cycle theory studies the more or less periodic mid and long-term fluctuations in the economic activity. We analyze some aspects of these issues later in this chapter.

Macroeconomics introduces mathematical models used in economic forecasting by governments and large corporations to facilitate the development of economic policy and (corporate) strategy respectively. Different schools exist that advocate different approaches. The Keynesian tradition focuses on demand and aims to devise policies to influence the level of unemployment and to mitigate business cycle type of fluctuations in the economy. It calls for an active role of governments and their institutions. Main instruments are fiscal policies that allow governments to invest anti-cyclically to stimulate or cool down the economy. The general Keynesian approach, also the more recent New Keynesianism, focuses on demand stimulation to mitigate what they consider imperfections in the workings of the free market. The advent of the financial crisis recently caused a revival of the Keynesian approach with, at least to date, some degree of success. The Neoclassical school bases its approach on attempting a synthesis with less emphasis on government spending and a larger role for monetary policy such as managing the rate of growth of money supply and the interest rate.

8.8. Microeconomics: The interplay between demand and supply.

We approach microeconomics based on the neoclassical theory of economic transactions (for a summary see e.g. Douma and Schreuder (2008)). In a competitive market, laws of supply and demand govern transactions. The logical assumption is that supply goes up if prices increase and that demand decreases with increasing prices. This generally results in equilibrium, a price where supply equals demand. (In fact, to reach a match of supply and demand at a realistic price exceeding zero we have to throw in assumptions about the mathematical properties of the supply and demand price relations; we ignore this complication and assume that the two curves cross at a positive price). As consumers generally buy more than one product and there are limitations to purchasing power, the choice which product the buyers procure and in which quantity depends on the utility curve of the consumer. Their preference for a certain basket of products derives from that utility function. These concepts results in a set of equilibrium prices for the products traded in an economy. We stress that indeed a situation of equilibrium based on balanced supply and demand may result. Without the introduction of additional assumptions, the theory is not predictive regarding the stability of that equilibrium. We discuss this stability problem in Chapters 2 and 6. Furthermore, even if the equilibrium is stable, oscillations around equilibrium may occur before the steady state finally emerges. Again, Chapters 2 and 6 show that even in the simplest situations in linear dynamic systems, oscillations around an intrinsically stable equilibrium may occur. In economic reality, such oscillations are likely to occur. We can easily understand this if we look at the theory of supply. The theory of supply assumes the capacity to produce influenced by two factors, labor and capital. These show different time scales of adaptation to the need for greater supply by increasing demand or the reverse. The traditional approach assumes the factor labor to adapt instantaneously, whilst the deployment and realization of additional capital in production asset takes longer. This difference in time scales is one of the reasons of cycles in the economic system. We return to this feature later in this chapter. It is interesting to study the analogy with the systems analyzed in thermodynamics where we have the same problem. In macroscopic theory, we arrive at the concept of steady states where state variables no longer change. As we showed the approach to these steady states shows characteristics discussed above. Even if the steady state is stable, oscillations around that state may characterize the approach to the steady state. Furthermore, the conditions for guaranteed stability of the steady state involve strong restrictions to the dynamic behavior of the system, i.e. the relation between the flows and the forces. We have extensively discussed these limitations and show that these are often unrealistic. This is the basis of the phenomenon of sustained evolution. In thermodynamics, there exist steady states in which matters behave more orderly. That is when the steady state considered is thermodynamic equilibrium. In that case, if there is more than one force influencing the behavior of the system, the dynamics of the approach of equilibrium may still show oscillatory behavior and generally show such behavior. However, the state of thermodynamic equilibrium always results, as thermodynamic equilibrium is stable by the requirements the second law poses to the production of entropy. The second law proclaims thermodynamic equilibrium stable. If we again invoke the intimate relation between information and the second law, we can rigorously prove stable equilibrium prices if these represent a true equilibrium like thermodynamic equilibrium. This is only the case if gradients in the state variables no longer appear. This would also have to hold for the state variable information. The perfect information assumption should apply to the players in the system: Everybody has the same information set and use it in an optimal way. In this way, we arrive at the conclusion that the perfect competition assumption used in standard microeconomics is sufficient for stable prices in the economic system. Also for economies close to equilibrium, we can prove stability of equilibrium prices. This rests on the analogy with the linear region near equilibrium in thermodynamics (Section 6.2). We also are sure to retain predictability of behavior. If we move further away from equilibrium beyond the stability limit, such simplifications are no longer valid. We insist that many parts of our present-day economic system are well beyond the critical limit characterizing the linear region.

We arrive at a number of consequences of the basic standard microeconomic theory. If firms have profit maximizing as their objective one of the consequences of the perfect competition model results in the paradox of profits. The model rigorously shows that obtaining profits exceeding zero is not possible in the equilibrium situation. We derive this from our EVT formalism in Chapter 5. We also would live in a very efficient world where a so-called Pareto-optimal allocation of resources reigns. This allocation implies that there exists no way to make all people feel better in terms of the satisfaction of their needs. This would be a very ideal and a very dull world because everything would come to a standstill. It is the ideal of a system that is effectively dead as there can be no life without dissipation of value.

Let us conclude this section by citing the assumptions and conclusions of standard neoclassical economic theory and perfect competition and analyze its consequences:

- There are large numbers of sellers and buyers. In theory an infinite number but we settle for large.
- There exist no entry and exit barriers. This somehow already conflicts with the assumptions underlying the dynamics of supply.
- As we stated, perfect information applies as everybody has the same information.
- No transactions costs exist, the settling of transactions does not lead to dissipation of economic value.
- The sole objective of firms is profit maximization.
- Homogeneous products, i.e. no differentiation of products supporting a single market need. This does no justice to the complexity of reality: Perfect reading of market needs would require more information than is realistically available.

- Although not often cited as an assumption (or a consequence) in the relevant literature, all producers should converge on the same optimal production method. Again incomplete information based restrictions make this unlikely.

The consequences of this model also prove enlightening. To cite a few:

- The concept of Pareto-efficiency, as defined above, applies.
- In addition, as already stated, no seller can make a profit.
- It is difficult to accommodate the widespread appearance of large firms and the trend of concentration in most industries.
- In addition, the phenomenon of product differentiation finds no place in the theory. This is an important feature of markets for performance products such as pharmaceuticals and (branded) consumer foods.
- There exists no differentiation in production technology whilst this is a known feature, even for undifferentiated commodities. Especially for undifferentiated commodities, superior production technology is a key requirement for success.

Despite these shortcomings, the standard perfect competition model is widely used in explaining the working of the supply and demand system in the economy. Part of this situation no doubt arises from the fact that not adopting the perfect information assumption leads to far greater difficulties in the analysis of the consequences of the model. We also note this in the development of thermodynamics that, whilst being founded on a practical application, the heat engine, essentially remained an equilibrium theory for some 100 years. This irrespective the fact that non-equilibrium considerations entered physics as early as the publication of Newton's Principia in 1687. It is Lars Onsager who earned a Nobel Prize in 1968, that proposes the first steps into irreversibility and hence gradients in information. The work of Prigogine and his co-workers in the early 1970-ies extends the formalism into the non-linear non-equilibrium area, where, as we argue, our industries and economies reside.

8.9. Models of economic growth, business cycles.

In the context of this chapter on Economic theory, we discuss some of the 20th century developments in economic theory. These elements are the subject of both exogenous and endogenous approaches to economic growth (Aghion and Howitt (1998), Romer (1990) and Solow (1956)). These approaches assume the dynamics of e.g. technology to be an important factor underlying growth. In this book, we highlight the influence of technology development as one of the causes of growth and cycles in the economy. The exogenous approach assumes technological change given, determined by factors outside the economy. In contrast, the endogenous approach seeks its causes within the economic system, i.e. technological change depends on the activities of the actors in the system. Examples are private initiatives such as industrial R&D and government activities to provide stimuli to R&D, such as tax incentives, or public R&D in academia or institutes. The first problem we need to tackle is to decide on the boundaries of the economic system, i.e. we have to define a frame of reference regarding exogenous and endogenous. Are e.g. government activities exogenous or are these endogenous and part of the economic system? We take the latter perspective.

A further problem is again almost philosophical; it refers to the matter of cause and effect. The EVT evolutionary perspective defines cyclical interactions between causes and effects. An example is identifying the driver behind evolution of economies. Is the driver variation of the genome of the firms or is it selection at the level of the phenotype. We argue that closure of the cycle blurs the familiar distinction between causes and effects. Cause and effect merge due to the cyclical interaction. In discussing the concept of sustained evolution, we see that even the forces that drive evolution are both cause and effect. The very activities of the actors in the market determine the size of the forces and hence the drivers of evolution are wholly or to a major extent endogenous to the system. The only almost wholly exogenous factor is the primary source of non-equilibrium on earth, the abundant energy in the solar radiation.

The discussion in this work shows that technological change is part of the essence of the development of the economic system. It is an important aspect of the development of superior information sets and thus drives economic evolution. Theories that do not consider this factor ignore an overwhelmingly important aspect of the evolution of information sets to increased competitiveness. Technological change is a very important aspect of economic growth, certainly in the mid to long-term perspective. We therefore reach the conclusion that technological change is wholly endogenous as the actors in the economic system shape this change. Today this is the dominant view in economic theory. To analyze this further, we turn to the concepts of business cycles and economic cycles as these feature in economic theory.

In 1860 the concept of economic cycles probably first appears in the history of economic thinking. History attributes this idea to the French economist Juglar. His economic cycles had a period of about 10 years; he, however, did not claim rigid regularity (Lee (1955)). Later on, Schumpeter and others propose business cycles or economic cycles of varying periodicity. Examples of cycles are:

- The Kitchin inventory cycle with a period of about 4 years.
- “The” business cycle of Juglar concerning investments in fixed assets, as said with a periodicity of around 10 years.
- The Kuznets infrastructure investment cycle of about 20 years.
- The long-term technological Kondratiev wave of 50 years.

The governing present-day economic theory does not adhere to the belief that cycles of more or less rigid periodicity exist. There are, however, economic fluctuations resulting from imbalances in the supply and demand. Endogenous causes, processes within the economic system, or exogenous causes from outside the economy, lead to these cycles. The present-day dominant Keynesian approach takes the endogenous perspective (Markwell (2006)), although alternatives, such as the Real Business Cycle theory (Plosser (1989)), exist. The proponents of the Keynesian approach assume failure of markets to smoothly clear due to imbalances in supply and demand, with the causes being alternatively at the supply or the demand side.

Let us get back to the approach to linear dynamic systems as introduced in Chapters 2 and 6. There periodicity arises naturally if the eigenvalues resulting from a linear model of the system show an imaginary part. Does this provide a valid vehicle to model economic cycles? If so, we retain predictability as these models predict a number of strictly periodical waves superimposed on the general direction in which the system evolves. This conflicts the present view of economic theoreticians that there are no cycles but rather unpredictable fluctuations. Furthermore, we get into conflict with the EVT approach developed in this work. In our model, the dynamic behavior of systems is unpredictable, as the future is not contained in the present and the past (Chapter 6). This holds for the evolution of the system once a steady state becomes unstable. It also holds for the way in which an intrinsically stable steady state recovers from a fluctuation. What we can predict is that the system returns to the steady state if that state is stable. The exact way in which it returns to the steady state is unpredictable and may follow complex dynamics. We already discussed the limitations of the well-behaved linear systems. Of course, we have to realize that the socioeconomic system is far from linear. Could the linear models still be representative for some features of the behavior of the system in the neighborhood of its steady state, an approach commonly adopted for a first approximation to the analysis of non-linear systems? Here we turn to another problem that adheres to the mainstream approaches to linear and non-linear system dynamics. In system dynamics, we assume the initial conditions describing the state of the system, i.e. the initial state variables, and the equations describing the evolution of those variables in time, the state equations, given. Here the real problem surfaces. We stressed that we only have a macroscopic picture of the system. Hence, only average values of the state variables are accessible. The real value of the state variables constantly fluctuates around the average defined by a macroscopic description. One can argue that this is irrelevant as the fluctuations are very unlikely to be large if the systems is complex and can exist in a large number of microstates. This proves beyond the point in the science of complexity perspective, where the so-called butterfly effect reigns (see also Chapter 2). The butterfly effect (e.g. Gleick (1988)) refers, as stated earlier, to the fact that very small effects, e.g. minute changes in the initial conditions describing the system can, even in the short run, have large effects on the condition in which the system find itself latter on. This comes back to the very essence of the uncertainty of the macroscopic description resulting in unpredictability of the evolution of non-linear complex systems.

What can we say about economic fluctuations based on EVT? Firstly, these will happen and do not have a rigorously specified periodicity. In addition, mainly endogenous mechanisms cause these fluctuations. A relation with exogenous intervention is often an unnecessary assumption. Some exogenous effect will have an impact (we can think of the celestial body presumed to have wiped out the dinosaurs; the prospects of global warming, on the other hand, are endogenous). What we can speculate is that the simpleminded linear systems approach may be relevant for the analysis of short-term effects such as the fluctuations in the stock market.

8.10. The behavioral theory of the firm.

The behavioral approach (see e.g. Douma and Schreuder (2008)) perceives firms as groups of stakeholders, such as shareholders, employees, customers, suppliers, management and customers.

The participants approach the firm starting from their private views on utility. They weigh the value of their contribution to the firm against the value they receive from working with or within the firm and adapt their actions to reflect this balance of value. The behavioral approach furthermore assumes incomplete information and asymmetries in information between the stakeholders of the firm. In our terminology, the stakeholders' pictures of the firm and its operations and products rest on differences in cost of information and statistical entropy.

Furthermore, the concept of bounded rationality emerges. The stakeholders' intend to maximize their utility but due to informational limitations and imperfections in the processing of the information, they behave with bounded rationality. The theory states that the stakeholders do not maximize utility but they "satisfice", i.e. they stay happy with their relation with the firm as long as their utility exceeds a preset limit. Given the nature of the stakeholders, they have a different view on utility with respect to the different aspects of the operations of the firm; this leads to a tradeoff by negotiation and results in the firm's holistic goals. These holistic goals than translate into goals for each of the individual stakeholders that do no longer solely reflect their own utility but also reflect the interests of the collective of stakeholders in the firm.

On comparing the behavioral theory with the standard microeconomic approach, the following differences appear:

- The firm is a coalition of stakeholders, i.e. reflects a team effort rather than interaction between individuals.
- Profit maximization is not the only goal of the firm. Rather it has a multidimensional objective function that reflects the tradeoff between the utilities of its stakeholders.
- The theory introduces limitations to information, asymmetries in information and cost of information and its transmission.
- Bounded rationality appears due to the limitations of information and its processing.

8.11. Transaction cost theory.

Transaction cost theories (Douma and Schreuder (2008)) try to analyze the management of transactions. For which transactions do we use organizations, like firms? Are there transactions for which the market provides the best solution for the arrangement of a transaction? Two basic assumptions about human behavior and human abilities are at the foundation of transaction cost theory. The concept of bounded rationality as discussed in the previous section applies. In addition, human nature results in opportunistic behavior where asymmetries in information exist. The latter point distinguishes transaction cost approaches from the behavioral approach.

The choice between markets and organizations as governance approach depends on three dimensions of transactions: Asset specificity, uncertainty resulting from complexity and the frequency of the transactions. We discuss this more fully in Section 16.2, when analyzing the industry value chain.

Transaction cost approaches ignore aspects of human social behavior such as trust. Many relations within a firm and between suppliers and buyers in industry rely on mutual respect, trust, and building long-term employer-employee or commercial relations.

8.12. Agency theory of the firm.

There seem to be two flavors of economic agency theory (Douma and Schreuder (2008)): The positive theory of agency and the theory of the principal and the agent. Agency theory attributes the existence of firms to the concepts of team production and specialization. These phenomena lead to competitive advantage and hence develop through evolution if the collaboration of people leads to a higher cumulative creation of economic value than if people choose to work alone. We identify a physical analogue of this phenomenon in Section 6.4 when analyzing the Benard problem. It reflects a characteristic of dissipative structures beyond equilibrium. A problem rests in the problem of shirking. Some people in the team may choose to put in less effort whilst still obtaining a share of the advantage of cooperating. It is one of the reasons for the development of management supervision in many, if not all, larger corporations.

The theory of the principal and the agent introduces the risk element in modern agency theory. The problem is how the principal and the agent can arrive at a fair sharing of the economic value that results if the output of a task does not only depend on the effort the agent puts in but also on random influences exogenous to the task. The reward structure has to reflect the risk taken by the principal and the agent. Their contract has to reflect the economic value perception of both parties that differs according to their risk and the associated uncertainty. This relates to the concepts of lacking information and statistical entropy and cost of information.

8.13. Evolutionary theories of economics.

This book tells a tale about change through chance and unveils the mechanism of evolutionary phenomena in reality as we experience it. It comes as no surprise that in contemporary economic theory the evolutionary perspective emerges as one of the approaches. Scholars adopt the approach that the evolutionary theories of economic processes rest on an analogy to biological evolution be it in its Darwinian or in its Lamarckian perspective (Section 6.10). We arrive at a different approach and we argue (Section 7.7) that a general theory of evolution exists that governs the development of systems under specified conditions. Evolution is the result of developing and processing information sets that allow the creation of economic value, in its most general sense, under the forces of competition. We fully discuss this theory in Chapter 6. We argue that the substance of evolution may change but forces of the same nature drive it in a diversity of real life situations. It hence applies to the overall universe, the biosphere on earth, the human species, science, technology, language, culture and last but not least firms and the socioeconomic system. In this section, we summarize the status of evolutionary approaches to economics today.

Nelson and Winter (Nelson and Winter (1982)) pioneer the evolutionary approach. These authors present a largely Lamarckian perspective in which variation and selection direct the change of an information set that is the essence of the firm. It applies to traits that are inherited, such as in the DNA in biology (Darwinian evolution), but in the perspective of Nelson and Winter, inheritance of exogenic traits (a Lamarckian perspective) that result from a process of learning by doing, is more important. We share this perspective concerning the socioeconomic system.

There also exists an organizational ecology approach to economic evolution (Hannan and Freeman (1989)). This is an approach that is, to our opinion, less productive and predictive, and we do not analyze it further here.

We feel that, in the perspective of a systems theory of evolution, the distinction between the Lamarckian and the Darwinian perspective is limitedly important. This difference refers to the type of coding and communication of the information that evolves under the pressure of environmental forces of selection, not to the basic mechanism of variation, selection and retention that lies at the heart of every evolutionary process. Important is that selection forces due to the environment in which socioeconomic systems operate clearly enter the perspective of economic modeling. Some approaches treat the environment exogenous to the system that evolves. This is, in our opinion, a dangerous perspective as the environment and the system are part of a cycle in which distinction between cause and effect becomes futile. We take the position that firms, markets, and the environment co-evolve in the perspective of the overall evolution of the universe.

Nelson and Winter focus on organizational routines as being the information sets that are the substance of evolution. We conveniently call this the "genotype" of the Nelson and Winter evolutionary view. The change of these routines is subject to inertia, i.e. whilst these can change in the process of evolution; there is a tendency of resistance to change. As we show earlier, all evolutionary processes are subject to this feature. Evolution involves a tradeoff between change of the genotype and its conservation to create a direction to evolution. These organizational routines form the so-called dynamic capabilities of the firms. The identification of dynamic capabilities rests on the resource-based view of the firm (Douma and Schreuder (2008)).

In the evolutionary perspective, bounded rationality in the firm's decision-making about the changes in its information set, its dynamic capabilities, is a factor that finally shapes its information set in interaction with environmental selection of successful changes.

We consider it not productive to do an in depth analysis of the reigning approach to evolutionary economics. This is very much the heart of this book as it presents an in depth analysis of the potential of a systems theory of evolution to understand and direct economic evolution.

8.14. Industry structure: The nature of entry barriers.

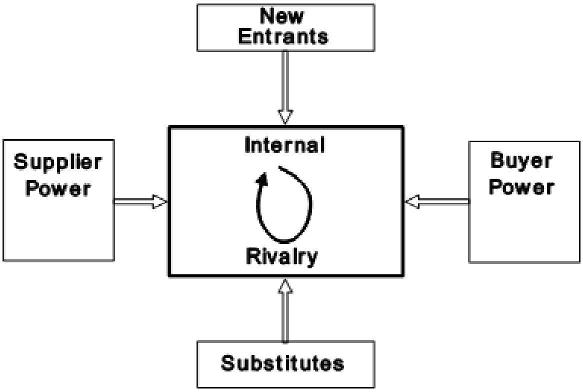

Fig. 8.5. Porter cross for the analysis of industry attractiveness.

This section discusses the concept of the industry value chain. A value chain develops in most industries; it refers to a structure such as depicted in Fig. 8.5. Porter's formalism of analysis of the value chain is widely used in the study of competitive dynamics in industry (Porter (1980, 1981, 1985)). It rests on an analysis of the so-called Porter cross.

The central box in fig.8.5 shows the industry and the internal rivalry, the competition, between the firms that compete for the sources of economic value associated with the need in the market. The figure shows four forces that, in addition to competition in the industry, shape the industry structure. Two forces relate to the power of suppliers of resources and information and the customers of the industry respectively. These forces follow from asymmetries in information between the firms in the industry and suppliers and customers respectively. In general, the attractiveness of the market to the players in it depends on the possibilities for creating, building and maintaining asymmetries in information. These possibilities generally decrease with the progression of the life cycle of the industry, although this is partly or wholly offset by the decrease of the number of players and hence the number of outlets and sources for suppliers and customers.

Another force relates to new entrants contesting the market. This depends on the feasibility of new entrants to match the competitiveness of the already existing players. The difficulty to enter increases with the progression of the life cycle, certainly if new entrants use information sets comparable with those of existing players in the industry. New entrants stand the best chance for a successful entry in the early stages of the life cycle and/or if they compete based on a new information set, e.g. a totally new competence base. Developing the same information set from scratch is extremely difficult and costly, certainly in the late growth and in the mature phases of the life cycle. In fact, acquisition of a leading actor in the industry often is a more effective if not the only tool. On acquisition, the information sets of the new entrant and a player already present merge. A question remains the likely return that such acquisition provides in the long run. There is a buyer and a seller relation and the likelihood that such acquisition is profitable in the eyes of both the buyer and the seller depends on asymmetries in information. As an example, the buyer may see synergies on merging the two information sets that are unknown to the seller. Such leverage can exist if the acquired information set allows the buyer to improve its position in industries in which it is already entrenched. Alternatively, the buyer may see possibilities to create and exploit new market needs based on the merged information sets. In biological terms, the merger creates a new genotype that may result in a wholly or partly new phenotype, i.e. new products and services, with enhanced competitiveness. Such mergers of information sets are also common biological evolution.

A special case of new entrants arises on the introduction of new technologies for the production of existing products. This may provide the turmoil that allows challenging entrenched players in the industry. It may be one of the existing players that pioneers the new approach in an attempt to change the competitive landscape in a mature industry. A new technology often requires a drastically different information set. An example is the development of the “green” routes for the production of β-lactam antibiotics that DSM pioneers the 1980-ies and 1990-ies.These green routes are more environmentally benign than the old classical chemistry based routes. The “green” synthesis strategies avoid the use of potentially dangerous and toxic solvents and reactants, using water based enzymatic technology and fermentation processes. Furthermore, these processes are more efficient, produce less waste and result in lower conversion costs if diligently optimized. These “green routes” had an important impact on the competitive landscape in the antibiotics industry.

The last force in the Porter Cross refers to substitution of the products of the entrenched actors by new products, rendering obsolete the existing products. Such a situation arises on exploitation of totally new information sets, e.g. technologies existing in another industry or totally new emerging technologies, Again the existing players survive if these succeed in acquiring the new information set either organically, through own development, or by acquisition, licensing or joint ventures. Examples of such developments are numerous in e.g. the birth of modern biotechnology after the early seventies of the 20th century.

A pitfall regarding the application of the Porter cross approach rests in treating the industry structure and the forces in a too static way. We highlight that sustained evolution and “inventing” new sets of information is the rule and not the exception in dissipative structures of which industries are examples. The players in the industry invent totally new ways of supplying the need or they create captive access to supplier resources and information, moving upstream in the value chain. In addition, movements downstream are an option. In all these options, alternatives are, in addition to access by organic development, access by acquisition, merger or alliance. The attractiveness of an industry depends on the possibilities to develop, grow and maintain asymmetries in information between players in the industry and the rivaling players identified in the Porter Cross.

An instrument to maintain such forces is R&D to hone the competences of the actors in the industry and legal opportunities to make information captive, such as collaboration agreements and intellectual property strategies. Furthermore, we need to approach an industry from an evolutionary perspective. The attractiveness of an industry depends on the creativeness of the players in the industry and those that represent the other forces. In the end, the battle rests on the ability to develop and maintain a sufficiently discriminating competence base, or, alternatively phrased, information set.

8.15. Strategy development.

In modern business practice strategy formulation and implementation is an accepted tool for business and corporate development. What is strategic planning all about? There exist many approaches to strategic analysis and strategic planning. Most companies follow a more or less formalized approach. We use a version of the procedure outlined in Douma and Schreuder (2008). Fig. 8.6 shows the approach. The process starts with the formulation of goals. Are we happy with what and where the company is today, where do we want to be in the future? Subsequently, we perform two types of analysis. Firstly, we develop a perspective on the environment of the company, including, but not limited to, the competitors. We analyze the external value chain and perform an analysis of the industry structure, for example using the Porter Cross.

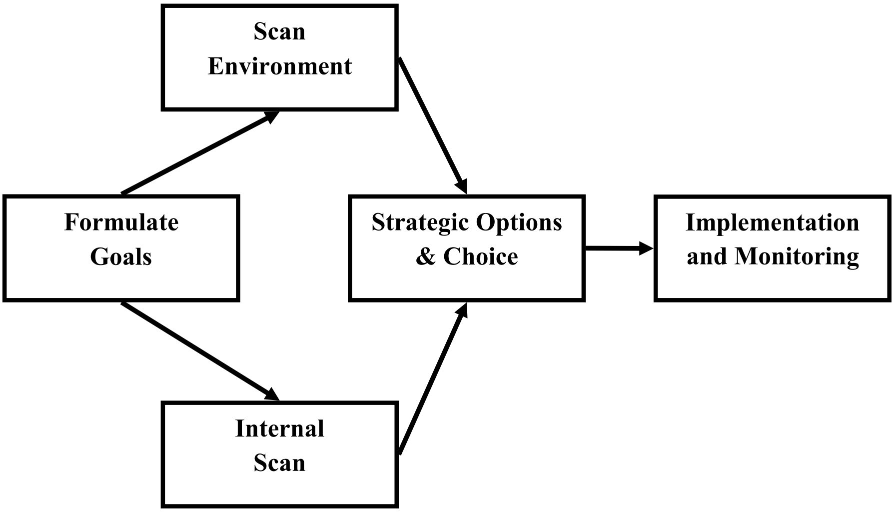

Fig. 8.6. An approach to strategic planning.

Secondly, we scrutinize the internal value chain (Section 8.6). The objective is to build a model of the environment and to get a perspective on the firm's genotype, the information set of the company. We also analyze the competitiveness of the present phenotype, the product portfolio understood in the broad sense we discussed earlier (Section 7.6). We develop a future perspective. Which of the economic value forces we presently exploit to produce profit will grow which ones will decay or disappear? What new forces exist or can we create? Where will the competitors move in terms of genotype and phenotype? Large uncertainties remain even if we do the analysis most carefully. Our picture of the environment and even of the company's own genotype and its possibilities for development, suffers from irreducible uncertainty and hence a significant statistical entropy. Of course, we use the best brains in the company and we develop a truly multidisciplinary perspective, but significant uncertainties remain. Still we have to agree on a model of reality and use it as a shared guideline for future development. From the model, we derive what products we need to develop to optimally couple to the sources of economic value in the future environment. From this, we derive what we need to change in our information set. In which direction do we need to shape it and how do we allow it to evolve? What instruments are available in terms of organic development of the competence set, e.g. by R&D, investments in marketing and new production facilities? Do we need cooperation and what part of the internal value chain needs to be involved? Do we need to acquire? Do we need mergers? In most cases, the uncertainties are large and we end up with a list of strategic options, with different actions and different risk profiles, due to the uncertainties inherent in our picture of the complex reality. We cannot afford not to choose from the options identified, certainly if these result in very different actions with respect to the environment and the information set of the company. We need to make choices. In absence of a choice we do not develop a systematic direction to our actions in the environment and internally. We will not have a consistent plan to direct the evolution of both the environment and our company. Without clear selection, evolution will not proceed in the planned direction. No matter the uncertainties of the mental model, we need to set a clear direction to provide an arrow to internal and external evolution. Given the uncertainties in our mental model we need milestones to measure whether the internal and external evolution develop as we planned. No doubt, there will be deviations inducing changes in the detailed implementation. However, we need to be careful not to deviate too much and too easily from the chosen direction. Also here, we encounter the copying fidelity threshold; if we change direction too quickly and too frequently, evolution will not take place in any direction and we will certainly not end up where we want to be. In addition, we need, for the same reason, to be careful not to overestimate the changes we can organically realize in the company's genotype, its information set. Drastic additions or changes will often require cooperation, merger or acquisition of existing information sets in other entities.

An additional complication rests in the fact that the internal environment in the firm and the external conditions are subject to co-evolution, i.e. the actions of the firm will shape the future environment in a way that is unpredictable in an ex-ante analysis.

8.16. Conclusion.

In this section, we discuss selected aspects of contemporary economic theorizing and analyze these from the EVT perspective. It serves to illustrate the relation to our socioeconomic theory, based on a systems theory of evolution. We showed that many features and limitations of the contemporary view on organizational economics logically follow from our systems theory of evolution.

CHAPTER 9. THE EVOLUTION OF THE UNIVERSE.

9.1. Introduction.

This chapter analyzes the evolution of the universe starting very shortly after the Big-Bang when the universe emerges. With the creation of the universe space-time, as we understand it through the theories of relativity, appears. We return briefly to the relativistic concept of space-time in a while. In The Big-Bang, the energy present in the universe today emerges. We can state a few things more about the initial state of the universe. We assume its statistical entropy close to zero, i.e. almost all energy was available to do work. The Big-Bang also creates all the potential value in our present universe. The universe is the prime example of an isolated system. Nothing exists outside the universe so it has no environment to engage in exchange processes. By virtue of the second law, the statistical entropy of the universe starts increasing at the moment of its creation. The universe is an isolated system in which irreversible processes definitely take place.

In this chapter, we concisely sketch the evolution of the universe, the creation of the chemical elements and of stars and planets. Our sun and the earth and its biosphere, result from this process. The reason we introduce this discussion of the evolution of the universe is to explore the stage on which life and our socioeconomic system evolved and continue evolving in an ever-expanding space-time. This chapter only provides a concise summary of the birth and particularly the evolution of the universe. More detail provides e.g. Glendenning (2004).

9.2. Our present-day picture of the universe.

Cosmology is the field of science that studies the universe. It has a long history and it traces back to at least 3000 BC when the priests of the advanced Babylonian culture try to read the future by studying the celestial bodies. A few hundred years after the beginning of our calendar, the Greeks avail of an accurate though complicated picture of the solar system: The earth centered view of the Ptolemaic system. This remains the reigning picture for several centuries, also due to the vigorous support of the Roman Catholic Church for the creationist earth centered view. Around 1500 BC Copernicus introduces the notion of a sun-centered system. Kepler builds on this view by introducing the laws of planetary motion and the elliptical orbits of the planets around the sun. Galilei uses the telescope to study the solar system and confirms Kepler's model.

In the early second half of the 17th century Newton's law of gravitation and the laws of motion provide the science behind the observed motion of the celestial bodies. Newton's work makes him founder of the modern scientific approach that combines scientific thinking with observations to deduce models, theories and laws (Chapter 2). This marks the beginning of the scientific revolution (Chapter 12). Newton does more; he introduces the notion that light contains a spectrum of colors. This provides the foundation for spectroscopy. Spectroscopy has a large effect on our understanding of the universe as we discuss later. Newton also holds the view that light is of a particulate nature at a time that the reigning paradigm regarding light is that of waves. The birth of quantum mechanics in the early 20th century reveals that light can show both wavelike and particulate properties and that neither view tells the complete story. The way in which we choose to interact with reality decides whether the wavelike or the particulate nature appears to us.

The Herschel's study the universe using increasingly powerful telescopes and chart many so-called nebulae that later prove to be galaxies like the Milky Way, the habitat of our sun and the earth. In the early 20th century, visual observation using telescopes that are even more powerful confirms that the nebulae are systems of stars. The view emerges that we live in a universe that contains billions of stars and that the sun is just an example of such stars. We also realize, following Einstein views on the limit to the speed of light, that in probing into the depths of the universe we look backwards in the past because of the time it takes for the light of distant objects to reach us. We also start studying the age of the universe. It must of course be older than earth. The oldest rocks on earth prove more than 4 billion years as witnessed by the result of studies on the radioactive decay of a uranium isotope.

In 1919, the accepted view is that of a static universe being out there not subject to change. This notwithstanding the fact that Einstein's theory of general relativity allows for the present picture of a universe in accelerating expansion, through the so-called cosmological constant Einstein introduces in the equations for general relativity. The Belgian scientist Lemaitre shows the theoretical possibility of an expanding universe from general relativity in 1929. We cannot overestimate the impact of general relativity on scientific thinking and the philosophy of science. Before Einstein's theory, physics has the character of a play on a stage made up by space and time. The concept of relativity shows that space and time independently are not part of physical reality; the amalgamated space-time takes their place. In addition, relativity shows that space-time is both the cause and the effect of the laws of physics.

The application of spectroscopy in cosmology introduces new possibilities for increasing our understanding of the universe. Light consists of packages of energy, photons. Each package has a wavelength and a corresponding discrete energy. Spectroscopy allows us to analyze light in terms of these photon packages. Furthermore, since the advent of quantum mechanics, one of the big additional revolutions in science in the 20th century, we know that atoms absorb light of a number of characteristic wavelengths. Identification of these wavelengths by spectral analysis provides a fingerprint of an atom. This allows us to study the chemical composition of the stars in the universe.

There is a further important characteristic of the light that reaches us from the distant stars. It results from the so-called Doppler effect. The pitch change of the whistle of a moving train on passing us explains this effect using the analogy of sound. In objects moving towards us, the pitch becomes higher, if these move away, the pitch becomes lower. When a train passes us this reflects itself in a sudden change of pitch. The Doppler effect also applies to the light that reaches us from the star. The spectral lines characteristic for the elements determined in a laboratory setting all shift to the red, i.e. to longer wavelengths, in the light that reaches us from the stars. This indicates that all stars move away from us. Furthermore, we observe that the shift to the red is larger when stars are more distant. This leads to the relation of Hubble. It states that the speed with which a celestial object moves away from us is proportional to the distance of that object. The proportionality constant is the constant of Hubble. Its dimension is the reciprocal of time and it allows us to estimate the age of the universe at over 13.5 billion years.

All objects we see in the universe move away from us with a speed that increases with the distance of the object. We live in a universe that expands and recent research shows that the universe enters a phase of accelerated expansion billions of years ago. By invoking the cosmological principle that states that there is nothing special about earth regarding studying the universe, this observation applies everywhere in the universe. We live in a universe in evolution, its statistical entropy increases and its expansion is subject to acceleration. We see stars at a distance of some 13 billion light years, i.e. their light originated 13 billion years ago. Beyond that limit, no stars seem to appear as the universe only contains clouds of gas of some 24% of Helium the remainder being largely Hydrogen. Gravitation, the reflection of energy, that is the source of the creation of stars and the evolution of the universe, has not yet made matter collapse to kindle the nuclear reactions that provide light in the darkness several hundreds of millions of years after the birth of the universe.

9.3. Understanding the small and the large.

The human brain, a product of evolution, allows us to understand reality and the way in which it evolved and evolves. We indicate reality as we perceive it with our senses, as the intermediate size level. In this way, we distinguish it from the very small, the world of atoms, nucleons, electrons and molecules, and the very large, the stellar systems in the universe. Although unforeseen in the evolutionary design of the brain, it allows the exogenous evolution of the scientific method. This methodology allows our collective brains to get a grasp on the part of reality that we not directly observe with our senses. Science in the 20th century starts presenting a picture of reality that is counterintuitive. Particularly the very small, the very fast and the very large in terms of mass and energy, show properties that challenge our intuitive intermediate level picture of the world.

Important to the understanding of our world are the four fundamental forces that drive the processes in the universe. These are the weak force, the strong force, the electromagnetic force and gravitation respectively. The weak force affects the building blocks of the atoms and it drives, among others, the fusion processes in the stars. The weak force affects the more elementary constituents of the nucleus of atoms, such as quarks. The strong force is responsible for the existence of the nucleons, the charged nuclei of the chemical elements. It holds the quarks in protons and neutrons together and creates the nuclei. Both forces act on a relatively short range. The electromagnetic force is responsible for the interaction between positively and negatively charged particles. It determines the attraction and repulsion of charged entities. Entities with different charges attract each other and particles with the same charge exert a repulsive force. Finally, there is the universal attractive force of gravitation to which all mass is subject; this includes radiation through its equivalent mass defined by relativity.

Einstein introduces the concepts of special and general relatively. Relativity implies that space and time in physics have no absolute meaning; only the combination in space-time is absolute. This theory rests on the fact that the speed of light is constant and it involves the concepts of space contraction and time dilation. If two observers move relative to each other, they both experience space contraction and time dilation in their frames of reference. The fact that time dilation requires space contraction is a direct consequence of the velocity of light being independent on the speed of movement of a system. The extent of time dilation and space contraction increases with increasing speed of movement of the observer. The work of Eddington during the 1919 eclipse beautifully substantiates Einstein's theory. It predicts that the gravitational force of energy condensed into mass results in the bending of light. This proves the case and the observations quantitatively agree with the predictions of general relativity. By the way, the laws of Newton prove to be a valid approximation at normal speeds and normal masses, i.e. in our intermediate level world.

Another important feature of the theory of relativity regards the equivalency of mass and energy. Mass can transform in energy and vice versa. The famous formula $E = mc^2$ expresses this equivalency. The energy, E, equivalent to a mass, m, is equal to that mass multiplied by the square of the velocity of light, c.

Einstein also analyzes the photoelectric effect, i.e. the fact that light of different wavelengths when interacting with a metal surface results in the emission of electrons with a different well-defined energy. This leads Einstein to introduce the concept of photons, packages of energy with an energy content increasing with decreasing wavelength of the light. Through the work of among others de Broglie, Planck, Pauli, Schrödinger, Heisenberg and Dirac this leads to the full-fledged quantum mechanical approach to the very small. The world of the very small is highly counterintuitive. Small particles sometimes behave as waves, sometimes as particles. The same holds for electromagnetic radiation, such as visible light, that sometimes behaves as wave, e.g. in diffraction experiments, sometimes as a particle as in the photoelectric effect. This is the well-known particle-wave duality. Quantum mechanics also introduces the uncertainty relation, i.e. we cannot accurately define place and velocity of a particle simultaneously, the product of their uncertainties must be positive above a certain limit. We generalize this uncertainty principle to apply to all dynamic system as information enters the macroscopic description through statistical entropy and that there is a cost of obtaining information. As we never have complete certainty about a system's microstate, the initial state of the system is never completely certain.

Another consequence of quantum mechanics defies our intuition. It implies that there exists no objective physical reality as this depends on the way we choose to interact with it. We cannot assume observation detached from reality as it is an integral part of reality. This is also true for all macroscopic descriptions where we gain information only if we interact with the system and hence perturb the system. It rests on the fact that gathering information always comes at a certain minimum expenditure of value. There is no such thing as a free lunch in the business of gathering information.

Light is as said an electromagnetic phenomenon. The changes in movement of charged particles leads to the emission of energy packages called photons. It is a way in which subatomic matter interacts with radiation. The reverse is also true; i.e. photons can transform in moving charged particles.

Electromagnetic radiation comes in a broad range of wavelengths and photons with an energy increasing with decreasing wavelength (Fig. 9.1).

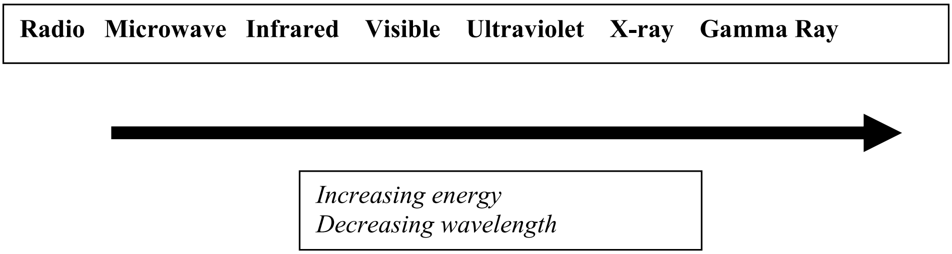

Fig. 9.1. Features of electromagnetic radiation.

Electromagnetic radiation is the prime means of communication, apart from sound, in both daily life and with the very small and the very big. It is the source of information about the distant objects in the universe.

What does initial universe, directly after the Big-Bang look like? It is immensely hot, heat contained in the motion of particulate matter in the form of particle-antiparticle pairs. The particles are electrons, neutrons, protons, the constituents of atoms, and their mirror images, their anti-particles. There were many more elementary particles in addition to the present constituents of the atoms and their antiparticles. A particle-antiparticle pair can convert in a photon and also the reverse process exists. In the early universe, these processes are so fast that thermal equilibrium exists. The universe is extremely dense at this stage with a sun mass condensed in a volume with a radius of 10 km. The primordial universe is expanding rapidly and it is more or less homogeneous and isotropic although fluctuations induce constant changes in local mass densities and intensities of radiation.

Matter and radiation interact in other ways. In the electrically neutral atoms, the electrons move around the nucleus in different orbits. The electrons can exist in a number of discrete energy states with increasing energies. The ground state is the state with the lowest energy allowed by quantum mechanics. The electron can move from the ground state to an excited state by adsorption of a photon of the energy that equals the difference of the energies of the ground state and the excited state. This adsorption process leads to the spectral lines that fingerprint the atoms as discussed in the preceding section. The electron can also drop back to a lower energy state and emit a photon with a wavelength and an energy corresponding to the energy difference between the states**.** These quantized phenomena lead to the typical adsorption lines in the spectra of light from distant stars and allow us to judge the composition of the stars.

We summarize the picture of the very small. Radiation comes in packages; massless photons characterized by their wavelengths and associated energy. The energy is inversely proportional to the wavelength. Photons, travelling at the speed of light, can also change into particles in which case these particles have a mass determined by $E = mc^2$ the famous expression of the equivalency between mass and energy. The proportionality constant, the square of the velocity of light, is a very large number indeed. A minute quantity of mass creates an enormous amount of energy in this conversion process. This is the principle underlying the destructive power of the atomic bomb.

The assumption that atoms of substance exist, i.e. that we cannot divide substances an indefinite number of times, is old. We owe it to the Geeks and the word atom still reminds of that heritage. It derives from the Greek word “atomos” that stands for “uncuttable”. The Greeks Leucippus and Democritus introduce this concept in the 5^{th} century BC.

Near the end of the 18th century, two observations concerning chemical reactions emerge. Explaining these observations does not lead to the explicit introduction of the atomic theory at that stage. The first observation is a conservation law with respect to mass in chemical conversions. Of course, this later proves an approximation. Also a chemical reaction involves a mass change and a corresponding energy gain or loss. However, for chemical reactions the approximation is good enough. Serious problems arise if nuclear reactions take place, i.e. if atoms split or fuse. Lavoisier formulates the mass conservation law in 1789. On chemical reactions, the mass of the products equals that of the starting products. In 1799, the French chemist Proust formulates the so-called law of definite proportions. The observation is that if a compound is broken down into elements, the masses of the elements always have the same proportions irrespective the amount or source of the compound.

The two laws inspire Dalton to propose an atomic theory in which he postulates that each chemical element exists of atoms of a single, unique type. Chemical reactions cannot destroy or alter these atoms. Elements combine to form more complex chemical compounds, molecules. Dalton thus introduces the first truly scientific theory of the atom. He reaches his conclusion by considering the results of observations on reality and the hypothesis of the concept of the atom. This involves a process of both deduction and induction to explain phenomena observed in nature. It is unclear to what extent earlier ideas inspire Dalton in formulating the atomic theory.

Based on his theory, Dalton derives a list of relative atomic weights for a number of substances in 1803. He estimates the atomic weights from the mass ratios in which they combine, with hydrogen being the basic unit. Unfortunately, he does not take into account that in some “elements”, the natural form is already a molecule, e.g. nitrogen and oxygen exist as N_2 and O_2. He also assumes that a simple compound of two elements always contains one atom of each element; he considers water to be HO, not H_2O. Avogadro corrects these flaws in Dalton’s theory in 1811. He introduces the famous Avogadro number based on the assumption that equal volumes of two gases at equal temperature and pressure; contain the same numbers of molecules. This leads him to conclude that some elementary gases are of a diatomic nature.

In 1827, the British botanist Brown observes that small particles, such as pollen, suspended in water, show random movement now known as Brownian motion. In 1905, Einstein theorizes that water molecules continuously knocking the pollen, cause this Brownian motion. He develops a mathematical model to describe this phenomenon. Experimental validation of his model in 1908 and 1909 by French physicists provides additional indications for the particulate nature of matter and supports atomic theory. This marks the period in which the atomic theory becomes firmly established. The story, however, does not end there because the “uncuttable” character of the atom still needs challenge.

Until 1897 when Thomson discovers the electron through his work on cathode rays, the reigning belief is that atoms are the smallest possible division of matter. Thomson discovers that an electric field deflects cathode rays. He concludes that the rays, rather than being waves, exist of negatively charged particles, electrons. Thomson correctly assumes the particles to emerge from the atoms of the electrode. He concludes that atoms are divisible, and that the “electrons” are one of the building blocks. To accommodate the overall neutral charge of the atom, he proposes that a uniform sea of positive charge surrounds the electrons. This is the plum pudding model of the atom as the electrons reside in the positive charge like plums in a pudding.

In 1909, one of his former students, Rutherford, refutes Thomson’s model. He discovers that most of the mass and positive charge of an atom resides in a very small volume at the center of the atom. Based on these observations Rutherford proposes a model in which a cloud of electrons circles around a small, compact nucleus of positive charge. This model of the atom involves two problems. The first is that electrons are charged particles and an accelerating electric charge emits photons involving a loss of the electrons kinetic energy: An orbiting charge should steadily loose energy and spiral towards the nucleus. The second problem is that the planetary model cannot explain the discrete lines in the emission and adsorption spectra of the elements.

At the beginning of the 20th century, quantum mechanics revolutionizes physics. Max Planck and Albert Einstein introduce the concept of the photon. Light can only be absorbed or emitted in packages containing a discrete amount of energy. Light cannot divide beyond the energy contained in one photon. In 1913, Bohr incorporates this concept into a model of the atom in which an electron can only orbit the nucleus in discrete circular orbits with a fixed distance from the nucleus and a corresponding energy. In this model, the electron cannot lose energy in a continuous way. It can only jump between discrete orbits under emission of a discrete amount of energy in a photon. Bohr's model is not perfect. It does not accurately describe the spectra of the atoms and elements.

A further refinement of the concept of the atoms involves the discovery of the isotopes and the unit of positive charge, the proton. Elements with a given number of protons can have different atomic weights. In 1918, Rutherford observes that nitrogen gas emits hydrogen nuclei when bombarded with alpha particles (alpha particles are nuclei of the element Helium). Rutherford concludes that the hydrogen nuclei emerge from the nuclei of the nitrogen atoms. He later observes that the positive charge of any atomic nucleus is always that of a discrete number of hydrogen nuclei. This, coupled with the facts that hydrogen is the lightest element and that the atomic mass of every other element is roughly equivalent to an integer number of hydrogen atoms, leads him to conclude that hydrogen nuclei are singular particles and a basic constituent of all atomic nuclei: The proton. Further experimentation by Rutherford establishes that the nuclear mass of most atoms exceeds that of the protons it possesses; he speculates that this surplus mass is due to unknown neutral particles, tentatively called “neutrons”.

In 1924, de Broglie proposes the wave-particle duality: All moving particles, such as electrons, can exhibit wave-like behavior. This inspires Schrödinger to explore if an electron in an atom exhibits a wave like rather than a particulate behavior. Schrödinger's equation, published in 1926, introduces electrons as a wave function rather than a discrete particle. His approach predicts many spectral features that Bohr's model does not accommodate. This concept is mathematically convenient but difficult to visualize and it faces opposition. One of its critics, Born, proposes that Schrödinger's wave function does not describe the electron but rather all its possible states. It predicts the probability of finding an electron at any given location near the nucleus.

Description of electrons as waveforms leads to the uncertainty principle of quantum mechanics. This involves that we cannot simultaneously know the position and the velocity of an electron with absolute precision. The theory invalidates Bohr's model, with its neat, clearly defined circular orbits. Electrons can exist at any distance from the nucleus but, depending on their energy level, are more likely found at certain distances from that nucleus. This introduces the probabilistic view in atomic theory.

The quantum mechanics based model of the atom presents one of the triumphs of modern science. It forms the basis for explaining chemical reactions between the elements and tells the full story of the spectral lines. It allows the understanding of all physical phenomena, including evolution phenomena like the evolution of the universe and life on earth.

In fact, the story about the atom does not end with the matter introduced so far. We discussed the elementary particles like the proton, the neutron and the electron as the main constituents of the so-called baryonic matter in the universe. However, these particles are, as initially supposed, not elementary. The protons and the neutrons exist of more elementary particles such as quarks and leptons. For our purpose, it is not necessary to introduce the myriad of particles that appear in the theory. We make an exception for two types of particles as these play a crucial role in the story about the evolution of the universe. Firstly, every particle has its corresponding antiparticle, a kind of mirror image. The electron has the positron as antiparticle with equal mass but a positive rather than a negative charge. If a particle and its antiparticle meet these annihilate and a photon forms that carries a way the joint energies of the particles. In addition, there exists a neutrino and its antineutrino. Neutrinos are uncharged small mass particles that are difficult to detect as these hardly interact with other forms of energy, including matter. They play a crucial role in the early stages of the evolution of the nucleons and the light elements.

9.4. Evolution of the structures in the universe.

In our present-day picture, the universe originates some 13.5 billion years ago when space-time emerges. We do not know from what the universe emerges but the law of physics, particularly quantum physics and general relativity, allow us to trace its history back to shortly after its creation. The present understanding of quantum physics dictates that we remain silent about the very early beginning. The early universe is extremely dense and of an extremely high temperature. From this point, the universe starts expanding and creates space-time in an increasingly larger volume. In this expansion, its temperature decreases. From the point where we can apply the laws of physics, we can paint a picture of what happens starting 13.5 billion years ago. We do this by considering a number of time intervals that where crucial to the universe reaching today's state.

Firstly, we consider the radiation-dominated era. In this era most of the energy in the universe consists of radiation, i.e. high energy, short wavelength, photons. Both matter and radiation are subject to the laws of gravity, i.e. are both subject to and a source of the gravitational force. We begin the detailed story of the evolution of the universe at a very early stage, one hundred billionth of a second after the beginning. The temperature is very high, about 10^{15} K. Apart from very high-energy photons there are quarks and antiquarks, electrons and positrons and neutrinos and antineutrinos. Equilibrium exists and the distinction between mass and radiation is fleeting. Photons produce particles and particle-antiparticle pairs and also the reverse processes take place. There is a small excess of matter over antimatter and this forms the basis for the very existence of mass in our present universe. The expanding universe cools down at a high rate. Initially the protons and neutrons that make up baryonic matter together with electrons are not stable for a long period of time. The stability of nuclei in which protons and neutrons bind together depends on their binding energy. The binding of nucleons together depends on a loss of mass on binding; the loss of mass results in the emission of a high-energy photon that guarantees energy conservation in the process of binding. Without emission of a photon, the equivalent of the mass lost on the creation of the nucleus cannot transform into energy. In the early high temperature stage of the evolution, such photons exist abundantly and the nuclei have a fleeting life.

In this early stage, the density of mass and radiation show fluctuations. Regions of increased density are likely to grow by the force of gravitation but tend to disappear due to the conversion of matter into radiation. One could say that the density fluctuations provide a blueprint, like DNA, for the development of a universe based on the non-linearity introduced by gravitation. This is a kind of autocatalysis that prefers growth of regions with densities in the high extreme of the probability distribution of mass density. This is one of the ingredients of evolution we identify in Chapter 6 when discussing the systems theory of evolution. The process of equilibration between mass and radiation is, early in the evolution of the universe, so fast that this prevents building sufficiently large gravitational forces. The distance to equilibrium is not sufficient to allow sustained evolution to occur. Such a point emerges when, later in the evolution of the universe, the interaction between radiation and matter decreases to allow creation of meaningful lasting density differences. This allows the blueprint in the pattern created by fluctuations to grow into the present-day large-scale structures in the universe. The decreased interaction of matter and radiation results in the universe's "DNA" to move beyond a copying fidelity limit allowing directed evolution through selection by the gravitational force.

When the temperature of the primeval universe drops to a level at which the energy contained in the radiation photons is no longer able to overcome the "strong force" that binds the nucleons together in the nucleus, the nuclei of light elements, such as Deuterium, Helium and Lithium form and stably exist. The supply of these elements for the lifetime of the universe emerges almost wholly in the first few minutes of the evolution of the universe. All other elements, such as carbon and oxygen that are vital to the evolution of life, start appearing hundreds of millions of years later by processes taking place in the stars, particularly in short lived very large stars of 10 sun masses or more. In the early phase of the universe's evolution radiation and matter move together as a single cosmic fluid bound together by the gravitational forces that apply to both radiation and matter. In this phase, the energy of the photons is too large to allow electrons to bind to the positively charged nuclei by the electromagnetic force. This behavior of radiation and matter as a single cosmic fluid stops, i.e. radiation becomes detached from matter, when the temperature drops to a level that allows electrons to bind to the nuclei to form charge neutral entities. Transformation of matter in radiation energy takes place by the acceleration of charged particles. When atoms become neutral entities, this mode of coupling disappears. This causes thermal equilibrium to subside and radiation starts to move freely, largely undisturbed by matter. In fact it kept moving almost freely ever since and this is the source of the cosmic background radiation that Panzias and Wilson discover as one of the strong evidences of the way in which cosmic evolution took place. This decoupling starts after some 300,000 years when the temperature of the universe reaches 3000 K.

The mass of the neutron is higher than that of the combined masses of an electron and a proton, hence free neutrons are unstable and spontaneously degrade into electrons and protons. In this process an antineutrino is produced that accounts for the excess energy (or mass) of the neutron over that of the combined electron and proton. As neutrons are very unstable, only a small window of time exists for the formation of deuterium and helium, i.e. the sources of these elements for the remaining life of the universe. In fact, antineutrinos and their interaction with matter were crucial in allowing helium to become about one quarter of all the mass in the present-day universe. All other elements together present only a minor, albeit a literally vital, fraction of all the mass in the universe. Without these traces of other elements, life as we know it can never evolve. In a complicated way the evolution of the light elements, particularly deuterium and lithium allow us to calculate the baryon mass density in the present universe. When the temperature drops to a level where the radiation decouples from matter, i.e. when the transformation of photons and nuclei and vice versa stops, the nature and number of photons and the amount of baryonic mass no longer changes. This happens some 300,000 years after the emergence of the universe. Today's amount of baryonic mass, after 13.5 billion years of cosmic evolution, equals the amount present at that stage.

A relic of the evolution of the universe is, as said, the cosmic background radiation. When the interaction between radiation and matter subsides, the photons, blackbody radiation at 3000 K, move at the speed of light in their creation of space-time as the evolution of the universe progresses. The expansion of the universe causes the photon's wavelengths to increase and their black body radiation temperature to decrease. Today after 13.5 billion years of expansion, we observe a cosmic background radiation with an equivalent temperature of about 2.7 K. The expansion of the universe causes this decrease in temperature. Because of this expansion, the energy and the equivalent mass density of the photons in the universe strongly decreases and their contribution to gravitational forces dwindles into negligibility. This is not where the story of the cosmic background radiation ends. Recent observations on the cosmic background radiation provide a missing piece that we discuss shortly.

We already discussed the cosmological principle: The universe is homogeneous and isotropic and appears the same to every observer anywhere in the universe, at least if we consider the large-scale characteristics. It always appears as an expanding universe with a constant average density of mass and radiation. If we consider gravitational forces in a finite universe with evenly distributed mass, the net gravitational forces at every point in space are zero. This also means that the laws of gravitation do, in itself, not lead to an explanation of the structure of the universe in which large voids and occasional concentrations of matter in galaxies, suns and planets appear. Somehow, probably in the way we describe above, matter becomes organized, a feature we indicated as a general aspect of evolution. The universe is a non-equilibrium system, the force of all forces that created the non-equilibrium being the Big-Bang. Furthermore, the laws of gravitation be they purely Newtonian or relativistic introduce autocatalysis. A fluctuation to a larger mass density enhances itself by the very laws of gravitation, resulting in increasing concentration of mass by gravitational attraction. The detailed analysis of the background radiation in the early universe shows that, albeit small, variations in density exist in the early universe when radiation decouples from matter. These fluctuations freeze and grow in size when the interaction of matter and radiation disappears in the decoupling phase of the evolution of the universe. Measurements show that small fluctuations in the temperature of the background radiation reveal small inhomogeneities about 300,000 years after the emergence of the universe. These small inhomogeneities are the seeds of the stellar structures in the universe and for that matter the source of other structures, including economies and industries.

As we indicated we can calculate the density of so-called baryonic mass, i.e. the mass of all the structures, whether visible or not, in the whole universe. The amount is equal to the amount present when radiation and mass decouple. It comes out to be 5 baryons per cubic centimeter of space or $3.5.10^{-31}$ g/cm^3. This is about 10 times the amount of mass we derive from the radiation it emits.

When matter and radiation separate by the fact that the temperature in the universe becomes too low to allow the energy of the photons to fight the Coulomb force and the strong force that binds nuclei together, matter and energy start to go their own separate ways apart from the force of gravitation that keeps their fates related for some time. However, the gravitational force due to radiation decreases more quickly than that of baryonic matter as a consequence of the expansion of the universe that causes a decrease of the energy and hence the gravitational mass equivalency of radiation. This process ends when gravitational effects of baryonic matter and the mysterious dark energy and dark matter start dominating the further evolution of the universe. The universe enters the matter-dominated phase of its evolution.

In the matter dominated era the galaxy clusters, galaxies and the stars and planets start to form. These large structures that appear in the universe, large in terms of the length scales that dominate our everyday existence but small compared to the length scale of the overall universe, nucleate by the inhomogeneities that preexist. These nuclei start to collect matter and increase in size due to gravitational forces. At a certain size, dependent on the temperature and the density in the universe at the moment under consideration, the structures do no longer grow in size but start to collapse under the force of gravitation. In this process of collapsing the history of the emergence of the universe reverses. Locally, the density and the temperature increase. The process of collapse cannot take place infinitely as compression leads to an increase of the pressure that starts to resists collapse by gravitational forces. This may, after some oscillations, lead to a situation where a temporary equilibrium exists between the gravitational forces and the pressure. A further element is relevant in the process of collapsing. The compression of the photons present shortens their wavelength and their energy increases, leading to a corresponding increase in temperature. At a certain point, this leads to thermonuclear reactions in the gas clouds that have the composition of the elements formed in the initial universe, i.e. the initial mass ratio of hydrogen to helium. Fusing four nuclei of hydrogen to form one atom of the next element in the periodic table (Helium) produces energy in the form of photon packages. This is due to the helium nucleus being lighter than the combined four protons used in its creation. The mass difference defines the energy released in photons by the equivalence of mass and energy according to Einstein's theory of relativity. This process leads to a balance between the internal pressure in the ignited star and the force of gravitation. It results in the production of radiation for an extended period of time that depends on the mass of the collapsed matter. The radiation may last for billions of years as in the case of our sun. The process of fusion proceeds beyond Helium and heavier elements unto Iron form. This fusion process also produces Nitrogen, Carbon and Oxygen that, in addition to Hydrogen, are the main elements that constitute living matter. Fusing to heavier elements beyond iron does not proceed with an energy gain. Iron is the final element that amasses in the centre of the stars.

For heavier elements, we depend on the processes in supernova and nova explosions. Heavy stars, i.e. much heavier than our sun, are relatively short lived and explode into a nova or a supernova in which large amounts of the material of the star, including the higher elements, disappear into space, where it becomes the material from which new stars form. The explosion leaves behind objects like white dwarfs, with a density equivalent to one sun mass in a 100 km radius object, or neutron stars with a sun mass in a radius of 10 km. Finally, the remaining object may become a so-called black hole, so heavy that no light can escape from it due to the gravitational force of the object.

In principle, the equations governing the evolution of the universe on a large scale derive from the theory of general relativity. These equations leave some space for different possible evolutions as a number of constant not given by the theory appear. These are the curvature of space and the so-called cosmological constant. To understand the curvature of space we use the analogy of the world of Euclidean geometry with its Cartesian space and the geometry on earth where we have a curved plane. Experimental evidence seems to point in the direction of a Euclidean universe. The cosmological constant reflects gravitational effects due to so-called dark energy. Experiments show that about 75% of all energy in the universe must be dark energy. This also leads to the conclusion that some 5 billion years ago the universe enters a stage of accelerated expansion.

Furthermore, in addition to baryonic mass, i.e. mass we see because of its interaction with radiation, so-called dark mass exists that does not interact with radiation and that we cannot detect by this interaction. It participates, however, in the effects of gravitation. The present picture is that only a small fraction of the total energy of the universe is visible baryonic mass.

9.5. Conclusion.

The universe starts with the Big-Bang and it is initially very dense and hot. It starts expanding and its density and temperature decrease. The early history of the universe is only traceable by the known laws of physics and particularly general relativity and quantum mechanics are instrumental to the understanding of its evolution. In fact, the theory of general relativity allows the prediction of a limited number of evolutionary scenarios for the universe. The expanding universe is one of them. The Belgian scientist Lemaitre derived it using general relativity. Accelerated expansion, the phase that apparently presently reigns, follows from the theory by a judicious choice of the cosmological constant that reflects dark energy or vacuum energy.

A few minutes after the emergence of the universe, the light elements form and their relative amounts freeze at the level we now observe in the stars with the aid of the spectroscopic fingerprints of these elements. This reflects itself in the ratio of helium and hydrogen, the almost exclusive constituents of all matter in the universe. This mass ratio of about 25% is in agreement with the theoretical predictions. In the beginning, radiation dominates the evolutionary fate of the universe. No light escapes in this period, hence we have no radiation evidence about that period. At about 300,000 years of age of the universe, radiation starts decoupling from matter and ever since light moves freely. We observe the remains of the radiation that originates in that period in the cosmic background radiation at a black body temperature of about 2.7 K. This reflects the 3000K at which it originally emerges 300,000 years after the universe's birth. The changed temperature reflects the expansion of the universe since that time. The relation of Hubble that relates the speed of stellar objects to their distance also provides evidence for an expanding universe.

After the radiation-dominated era of the universe's evolution, it enters a phase of matter-dominated expansion as the contribution of radiation to gravity decreases more quickly than the mass related part when the universe expands.

A detailed analysis of the cosmic background radiation reveals that 300,000 years after the birth of the universe it is almost homogeneous and isotropic, certainly at a large scale. On a small scale, inhomogeneities exist, and in the course of the evolution of the universe, these form the seeds of the structures we see: clusters of galaxies, galaxies and stars. These structures result from the non-linear amplification of the initial "lumpiness" of space-time by gravity. If large amounts of gas amass, the resulting clouds collapse under the gravitational force and the density and temperature increase to a level where fusion of light elements becomes possible. This fusion process provides the energy for the light we receive from e.g. our sun. Particularly for large stars the life may be relatively short as fusion reactions cannot go on forever (beyond the atom number of iron fusion does no longer result in energy generation). This causes the stars to explode in a supernova or nova. This process expels matter into space, including all the elements that we now observe on earth. All the material on earth results from the processes just described including elements like Carbon, Hydrogen, Oxygen and Nitrogen, vital ingredients of life and our socioeconomic system. In the next chapter, we further trace these latter evolutionary processes.

CHAPTER 10. THE CO-EVOLUTION OF EARTH AND LIFE.

10.1. Introduction.

In this chapter, we discuss the evolution of earth and the emergence of life on earth. In fact, we argue that we have to consider the co-evolution of earth and life. The conditions on early earth are such that replicator molecules emerge or enter from outer space. However, once the first replicators appear the pattern prescribed by the systems theory of evolution takes over (Chapter 6) and earth and life start to co-evolve in a process of sustained evolution. The evolution of life is not a play that enrolls on the stage the environment on earth provides. In fact, life and earth both create the stage and are players in the play. The appearance of life drastically changes the path of earth's evolution.

This result is the typical pattern of evolution of complex system where the cycle between system and environment closes and we have to shift the boundaries of the system to perform a meaningful analysis.

In addition, critical historical events (bifurcation points from the perspective of the systems theory of evolution) are important, these, among others, concern sudden often catastrophic changes in the conditions on earth of an exogenous or an endogenous nature. Critical bifurcation points change evolutionary patterns in an unpredictable, revolutionary, way. In this way, periods of gradual evolution and events that cause revolutionary changes, characterize the direction of the developments.

This Chapter contains only a brief summary of the subject. For more detail, we refer to e.g. Gould (1984, 2001), Cockell and Carfield (2008).

10.2. The evolution of the solar system.

We do not discuss the emergence of the Milky Way, as this galaxy is not special from the evolutionary perspective. It emerges along the lines of the general pattern of evolution of the large structures in the universe. We briefly touch on the solar system as a typical example of stellar systems harboring planets like earth.

Our solar system evolves from a rotating concentration of interstellar dust and gas, the solar nebula. It largely consists of hydrogen and helium produced shortly after the Big-Bang. It contains minor amounts of heavier elements produced in the stars or by (super) novas in the evolution of the universe. Roughly 5 billion years ago, the solar nebula begins to contract. The cloud accelerates its rotation and gravity and inertia flatten it into a disk. Most of the mass concentrates in the middle and heats up, but perturbations create concentrations of mass that rotate around this centre.

The attraction of mass, the increasing rotation and the effect of gravity create an enormous amount of energy at the center. The disk's center starts heating up. Nuclear fusion of hydrogen into helium starts and the Sun ignites. Meanwhile, gravity causes matter to condense around objects outside the new sun due to the balance between gravitation and the centrifugal force. Larger fragments grow into larger objects and ultimately form precursors for planets. This includes an object about 150 million kilometers from the center of the solar system: Earth. It forms about 4.5 billion years ago.

10.3. Origin of earth.

As a consequence of gravitation, the emerging earth grows and the inner part becomes hot. Metals start to melt and a separation of a mantel and a metallic core results. This is instrumental in creating the layered structure of earth and is responsible for its magnetic field.

The early atmosphere surrounding earth consists of light elements, mainly hydrogen and helium. The solar wind and earth's heat clear away this early atmosphere. When earth becomes larger, gravitational attraction retains an atmosphere that includes water.
Our planet has a large satellite, the moon. The giant impact hypothesis assumes that the moon forms during a collision of early earth with another object. The impact releases a large amount of energy and both the earth and moon melt completely. The enormous amount of energy released, again blows of the planet's atmosphere completely. The impact also changes earth's axis of rotation; this is responsible for the seasons.

10.4. The formation of the continents, oceans and atmosphere.

Earth lacks an atmosphere and cools quickly. Within 150 million years a solid crust forms. It is of a different composition than the silicate minerals based crust of today. Within the earth, further differentiation begins when the mantle partly solidifies again. During the early period of earth's evolution the mantle is much hotter than today.
Steam escapes from the crust and volcanoes release gases to create a new atmosphere. Collisions with other bodies result in import of additional water. As the planet cools, clouds form and rain creates the oceans.
Earth's atmosphere now contains mainly ammonia, methane, water, carbon dioxide, and nitrogen. Free oxygen is absent as it reacts with minerals on the surface. There is intense volcanic activity and, in absence of an ozone layer, strong ultraviolet radiation reaches the surface.
Convection drives the process of plate tectonics. It results from a heat flow from the earth's core to the surface. This creates rigid plates that evolve to the precursors of today's continents in a complex way.

10.5. The origin of life.

We only know the broad principles of the origin of life. There are at least two hypotheses about the origin of life. One suggestion is that organic components or replicators arrive on earth from space, whilst the other hypothesis argues these form on earth. If life arises on earth, the timing of this event is around 4 billion years ago. The chemistry of early earth and the influx of free energy from the solar radiation create sources of free energy and allow the emergence of the first replicators. The conditions for the start of sustained evolution exist and it inevitably starts. After the initial emergence of the early replicators, their diversity increases due to the limited fidelity of the copying mechanism these employ. In the beginning, few replicators and relatively abundant free energy sources exist and a process of evolutionary radiation sets in to result in molecular species growing in number and diversity. Soon, however, scarcity of resources triggers increasing competition. In addition, this leads to the search for new sources of free energy; in part, the evolving life forms themselves become free energy resources.
The nature of the first replicator is unknown because life's current replicators, DNA and RNA, successfully challenged their function. Most present-day life forms use DNA as their replicator. RNA is a likely candidate for an early replicator, because it both stores information and catalyzes reactions. At some point, DNA largely takes over the coding role from RNA and enzymes become the catalysts.
Modern life forms have their genetic material packaged inside a cellular membrane. The prevailing theory is that the membrane appears after the replicators. The combination of membranes and replicators leads to the first cells. The last universal common ancestor (LUCA) emerges roughly 3.5 billion years ago. It is the ancestor of all life on earth. It was

probably a prokaryote, possessing a cell membrane and RNA and protein based molecules (ribosomes) as the units responsible for the translation of the genetic code into proteins. Prokaryotes lack a nucleus and membrane bound organelles such as mitochondria, the energy factories of cells, or chloroplasts, units responsible for photosynthesis. These more complex structures appear in the eukaryotes. Like all modern living cells, LUCA uses DNA as its genetic code, RNA for information transfer and protein synthesis, and proteins as structural elements and catalysts.

10.6. The proliferation of advanced life forms.

Life forms using surrounding organic molecules as raw material and energy resources are candidates for early cellular life forms. As this food supply becomes limiting a new strategy evolves in some cells. These cells directly couple to solar radiation as a free energy source. About 3 billion years ago, the precursor of modern photosynthesis develops. This makes the sun's energy available not only to autotrophs directly using this radiation but also to heterotrophs that consume them. Photosynthesis uses the plentiful carbon dioxide and water as raw materials and produces energy-rich organic molecules (carbohydrates) using the energy of sunlight.

Moreover, varieties of photosynthesis evolve that result in the production of oxygen. At first, oxygen binds with minerals and dissolves in water. When the oxidation of most of the reacting minerals completes, oxygen accumulates in the atmosphere. The photosynthetic production of oxygen starts transforming earth's atmosphere to its current state. Some of the oxygen transforms into ozone under the influence of ultraviolet radiation and forms the ozone layer in the upper part of the atmosphere. This ozone layer absorbs a significant part of the ultraviolet radiation. It allows life to grow at the surface of the ocean and ultimately on the land. Without the ozone layer, ultraviolet radiation leads to unsustainable levels of mutation in the exposed cells. As we know from the general theory of evolution, too low copying fidelity caused by excessive mutation, precludes directed evolution of the genetic code.

Photosynthesis changed the evolution of earth in another major way. Oxygen was toxic to the early anaerobic life forms. Most of the anaerobic life forms die as levels of oxygen rise. Some survive and thrive in the evolutionary space created by this mass extinction. Some cells succeed in using oxygen to enhance their metabolism and derive more energy from the same quantity of food. Anaerobic growth on glucose, e.g. the fermentation of glucose to ethanol, produces 2 moles of the energy currency ATP per mole of glucose. Aerobic metabolism of that same compound produces almost 40 moles of ATP. This proves a major bifurcation point in the development of life on earth. An oxygen rich atmosphere allows respiration, a far more effective use of resources than anaerobic growth.

About 2 billion years ago, eukaryotic cells appear. Eukaryotic cells are larger and more complex than prokaryotic cells, and we only recently understand the origin of that complexity. Around this time, the precursor of the mitochondria, the energy factories of aerobic cells, form. An aerobic bacterial cell related to today's bacteria enters a larger anaerobic prokaryotic cell and survives inside the larger cell. Using oxygen, it metabolizes the larger anaerobic cell's waste products to produce energy. Part of this energy returns to the host. A stable, mutually beneficial symbiosis develops between the two cells and they become dependent. The two cells develop into a single organism, a eukaryote, and the smaller cells become organelles now called mitochondria. These mitochondria are the energy factories of the eukaryotic cell.

An analogous mechanism results in photosynthesis in eukaryotes when cyanobacteria enter larger heterotrophic cells and become chloroplasts, the organelles that allow photosynthesis. This leads to photosynthesis in eukaryotes less than 2 billion years ago. Both the bacteria and the eukaryotes continue to evolve and this leads to further sophistication of the exploitation of the energy that is potentially available in the environment.

The first precursors for plants, animals and fungi arise but still live as solitary cells. Some of these start to live in colonies, and gradually division of labor takes place. Around 1 billion years ago, green algae emerge as the first multi cellular plants.

An increasing pace of development of new multicellular life forms characterizes the period that follows. The diversity of new, larger, forms of life strongly increases. These are the precursors of radically new life forms. An important development is the emergence of muscular and neural cells. The neural cells are the precursors for the development of the brain, a vital element in the appearance of exogenic evolution.

Some 500 to 600 million years ago, many modern life forms emerge. Life colonizes the land. The evolution of life is sometimes gradual, but sometimes shows explosive radiation of new species or sudden mass extinctions. These sudden events often result from rapid changes in the environment by natural disasters such as volcanic activity, meteorite impacts or climate changes. Important bifurcation points in earth's evolution, caused by both endogenous and exogenous factors, result. Gould (2004) terms this phenomenon "punctuated equilibrium".

The rate of the evolution of life accelerates as the biological proliferation increases. Some 500 million years ago, most modern groups of species are present. The development of hard body parts facilitates the preservation and fossilization of such life forms. The first animals appear.

As said, oxygen accumulation from photosynthesis results in the formation of an ozone layer that absorbs much of Sun's ultraviolet radiation. This allows survival of organisms on land and life becomes better adapted to survival out of the water. Several hundred million years ago, plants (probably resembling algae) and fungi start growing at the edges of the water, and then colonize land. The oldest evidence for animals on land appears around 450 million years ago. These feed on the vast food sources provided by the terrestrial plants. The first four-legged animals evolve. This is the origin of the amphibians. Some 200 million years ago the mammals appear.

A severe extinction event takes place 250 million years ago; most of the life on earth dies. However, some life persists and dinosaurs evolve from their reptilian ancestors. The dinosaurs survive another extinction event, 200 million years ago, and they soon become dominant. Some of the mammalian species further establish themselves during this period. Competition with birds drives many of the flying dinosaurs to extinction and the dinosaurs are probably already in decline when, 65 million years ago, a large meteorite strikes earth. The resulting ejection of vapor and dust in the atmosphere obscures sunlight. This inhibits photosynthesis. Most large animals, including the dinosaurs, become extinct. This creates the evolutionary space that allows mammals to become the dominant group. Perhaps a couple of million years later, the last common ancestor of the primates lives. We leave the story of the co-evolution of the earth and life on earth here, as the appearance of the primates leads to a recognizable ancestor of the Hominins and finally *Homo sapiens*. We continue the story in the next chapter where we discuss the evolution of humankind and its society.

10.7. The heat balance of earth.

Another interesting feature in discussing the interaction between life and the conditions on earth is the heat balance on earth today (Fig.10.1) (Dam (2009)). The solar radiation provides energy at a rate of 235 W/m^2. Some of the radiation absorbs in the atmosphere at a rate of 67 W/m^2. Earth absorbs 168 W/m^2. Energy and heat from the atmosphere transfer to earth and space at rates of 324 and 195 W/m^2 respectively. Vaporization of water leads to a loss of heat of the earth at a rate of 102 W/m^2. Heat radiates to the atmosphere at a rate of 390 W/m^2, this largely absorbs in the atmosphere under the influence of small amounts of greenhouse gasses like carbon dioxide and methane.

The radiation of the sun passes through the atmosphere without much absorption whilst the atmosphere almost wholly absorbs radiation back to space. This difference results from the different wavelengths of those two types of radiation. Small amounts of greenhouse gasses, such as carbon dioxide thus act as a selective filter and the scenario of an increasing concentration of carbon dioxide leads to the expectation of a warmer earth in the future. On the other hand, without the selective filter the earth's atmosphere provides, the temperature on earth decreases to an inhospitable level of -18 degrees centigrade. In fact, the changes in the composition of the earth's atmosphere and fluctuations and evolution of the intensity of incoming radiation led to drastic changes in the temperature on earth due to changes in the heat balance.

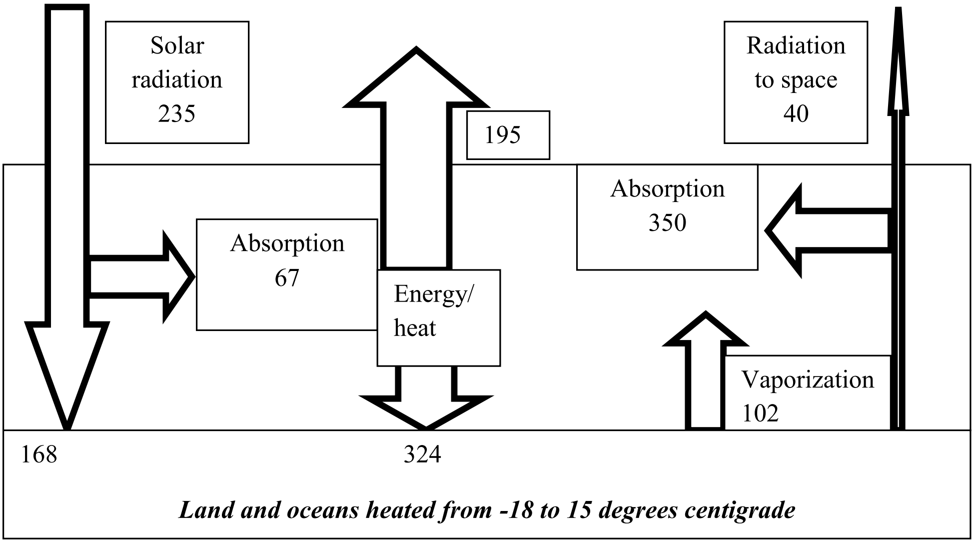

Fig.10.1. Heat balance for earth and its atmosphere. All figures W/m^2

10.8. Conclusion.

In this chapter, we traced the evolution of the solar system, earth and the various life forms on earth. The first life forms are heterotrophic bacteria feeding on chemical substances present on earth. These evolve out of replicator molecules of which DNA survives to data as the information set underlying most living systems. Autotrophic bacteria develop the capability to use sunlight to synthesize energy sources out of water and carbon dioxide thus tapping the vast free energy source in the form of the solar radiation. This allows life to proliferate beyond the level possible based on free energy sources on earth itself. The autotrophic life forms induce the emergence of heterotrophic life that feed on the source of free energy the autotrophic biomass provides. In this way the tree of life with its bacteria, eukaryotes, plants, insects, birds, fish and land colonizing animals forms.

An important event is the emergence of photosynthesis that results in the production of oxygen. This shifts the composition of the earth's atmosphere to the around 20 % oxygen we observe today. This has a number of effects. Firstly, it made aerobic growth, using oxygen, possible. This results in far more efficient ways for heterotrophic growth, thus increasing the free energy available to sustain the biosphere on earth. Secondly, it results in the formation of the ozone layer that shields the earth's surface from the ultraviolet radiation of the sun. This lowers the rate of mutation of DNA molecules to a level that sustains evolution, i.e. below the copying fidelity requirements for developing information sets of sufficient size.

What we also see in this story is the phenomenon of closing cause and effect cycles leading to evolutionary feedback as the earth and life together create an environment allowing life to evolve. The emergence of life on earth shapes an environment in which sustained evolution of life becomes possible. Neither life nor its environment on earth preexists; rather these concertedly come into existence. We recognize the general features of the systems theory of evolution.

We end this chapter when the mass extension of the dinosaurs creates the evolutionary space for the primates, ancestors of the apes from which the human species evolves. We continue the story from this point in the next chapter.

CHAPTER 11. HUMAN EVOLUTION.

11.1. Introduction.

In the preceding chapter, we analyze the co-evolution of earth and life unto the point where the first primates appear. In this evolution, we recognize a number of features of the systems theory of evolution we delineate in Chapter 6. In this chapter, we trace the evolution of humankind, as we know it today. We reemphasize that, with the appearance of humans and the further evolution of the functions of the brain, exogenous information increasingly complements and augments the coding function of DNA, the reigning information carrier in early evolution. Of course, primitive forms of exogenous evolution appear much earlier but the dominance of DNA based evolution remains in place when the primates first appear.

This chapter complements the story unto the precursors of our present-day socioeconomic system. We provide only a sketchy picture of human evolution. Lewin and Foley (2004) provides far more detail.

11.2. History of ideas about human evolution.

The reconciliation of the concept of evolution and considering humans as its product challenges the scientific community for an extended period. Linnaeus and others of his time, consider the great apes to be the closest relatives of human beings. This idea rests on the similarities in morphological and anatomical characteristics. The possibility of linking humans with apes by descent obtains a deepened scientific basis after 1859 when Darwin (1859) publishes his On the Origin of Species. This work establishes the idea of a continuous line of evolution of new species from earlier ones. In the days of Darwin's publication, some people continue to cling to the existence of a discontinuous gap between humankind and the other biological species. Is the continuous force of evolution responsible for closing this apparent gap?

Thomas Huxley’s (1863) arguments about the nature of human evolution from apes by illustrating the similarities and differences between humans and apes, trigger the debate about the origin of our species. Particularly his 1863 book contributes to the debate. However, many of Darwin’s early supporters fail to agree that the more advanced aspects of human culture derive from natural selection. Darwin (1871) applies the theory of evolution to humankind when he publishes The Descent of Man.

A major problem is the lack of fossil intermediaries between the apes and humankind. In 1925, the discovery of *Australopithecus africanus* remedies part of this problem. Although its brain is small (410 cm^3 compared to around 1350 cm^3 in modern man), its rounded shape, unlike that of chimpanzees and gorillas, resembling a modern human brain, makes the species a candidate intermediary. In addition, these Hominins have different teeth, and the position of the arms indicates walking on two legs. These traits lead to the conviction that the fossil is an early transitional form between apes and humans. This supports the present ideas that the Hominins share a common ancestor with the chimpanzees.

The classification of humans and their relatives changes considerably over time. The Australopithecines are ancestors of the genus *Homo*, the group to which modern humans belong. These both belong to the Hominins. Recent material suggests Australopithecines to be a diverse group and that *A. africanus* may not be a direct ancestor of modern humans.

11.3. Evolution of the apes.

The primates, early common ancestors of the great apes and the humans, trace back at least 65 million years. There is considerable evolutionary radiation in the primates and a variety of ape like creatures develops, including a wide variety of species adapted to life in trees. These live in East Africa. These ape-like species trace back to at least 20 million years ago.

11.4. The relation between humans and the great apes.

Species close to the last common ancestor of gorillas, chimpanzees and humans appear in the fossil record. Molecular evidence suggests that more than 5 million years ago the gorillas, and later the chimpanzees, split off from the line leading to the Hominins from which the human species evolves. Human DNA is approximately for 98 % homologous, or loosely phrased identical, to that of chimpanzees (Lewin and Foley (2004)).

Other Hominins appear in drier environments in Africa. Fossils of these species in the human lineage following divergence from the chimpanzees are relatively abundant. These earliest Hominins are the last common ancestors of modern humans. The early savannah apes, the Australalophitecus species, follow these. Subsequently, about 2-2.5 million years ago, the early intelligent *Homo habilis* marks the appearance of the genus *Homo*.

11.5. Genus *Homo.*

Fig. 11.1 presents a simplified evolutionary tree of the genus *Homo.* There is still considerable debate about the exact shape of this tree and it is only one of the more likely ways in which the present *Homo sapiens* appeared on the evolutionary stage.

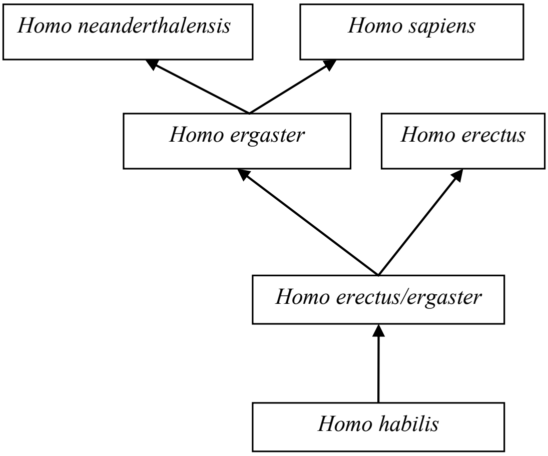

Fig. 11.1. A simplified evolutionary tree of *Homo*.

Homo sapiens is the only non-extinct species of its genus. Other *Homo* species are now extinct, some of these are ancestors of *H. sapiens*. Some diverge from our ancestral line and are no direct ancestors.

H. habilis lives from about 2.5 to 1.5 million years ago. The first species of the genus *Homo* evolves from the Australopithecines in South and East Africa. *H. habilis* has a larger brain than the Australopithecines and makes tools. The identification of *H. habilis* as the first member of the species *Homo* is subject to debate.

H. erectus lives from about 1.8 million years ago to about 500,000 years ago. Often scientists use the name *H. ergaster* for species in the early phase, from roughly 2 to 1 million years ago. Some populations of *Homo habilis* evolve larger brains and make more elaborate stone tools. These differences result in classification as a new species, *H. ergaster/erectus*. In addition, *H. erectus/ergaster* is the first human ancestor to walk truly upright. Probably, he uses fire to cook meat.

H. heidelbergensis lives from about 800,000 to about 300,000 years ago. Most current experts believe *H. rhodesiensis*, estimated to be 300,000–125,000 years old, a relative of *Homo heidelbergensis. H. neanderthalensis* lived from 400,000 or about 250,000 years ago to as recent as 30,000 years ago. Competition from *Homo sapiens* probably contributes to Neanderthaler's extinction. *H. sapiens* (the "intelligent" *Homo*) lives from about 200,000 years ago to the present.

Between 400,000 and 200,000 years ago, the trend in skull expansion and the elaboration of stone tool technologies increases, providing evidence for a transition to *H. sapiens*. As we argue the relation between tool making and brain expansion, reflects a co-evolution phenomenon. The direct evidence suggests a migration of *H. erectus* out of Africa followed by further speciation of *H. sapiens* from *H. ergaster* in Africa. A subsequent migration within and out of Africa eventually replaces the earlier dispersed *H. erectus*. Out of Africa is the term for this migration and origin theory.

Molecular DNA evidence shows that human beings are genetically highly homogenous. This may be the result of their relatively recent evolution. Most probably, a population bottleneck resulting from natural events, results in the appearance of *H. sapiens* some 200,000 years ago or even later, 150,000 years ago.

The modern approach to the evolution of modern man considers four key events:

- Adaptation to terrestrial live, i.e. leaving the trees as a habitat.
- Adaptation to walking on two legs.
- Growth of the brain relative to body size.
- Development of culture and civilization.

The early ancestors of our species are most probably adapted to living in trees and one of the first events in the evolution of the ancestors of present-day man involves leaving the trees and living at the surface of the earth. This induces adaptation to walking on two legs. Walking on two feet is one of the earlier traits that distinguish our ancestors. The morphology of the arms of the earliest species indicates that those were still climbing trees, and spent a significant proportion of their time in trees. The possibility to use the hand for advanced manipulation of objects is an advantage of walking on two legs. Genetic evidence shows that the chimpanzee human divergence takes place more than 5 million years ago. Early humans start making tools 2.5 million years ago. Stone tools allow early humans to access new resources.

The appearance of *Homo erectus* about 2 million years ago shows the origins of three important human traits. *Homo erectus* exhibits a much larger brain size than earlier Hominins, i.e. about 900 cc or more than two times that of the early Australopithecines. *Homo erectus* children took a relatively long time to grow and thus relied on parental investment for a longer period of development. This extra time allows for and encourages cultural behavior. No doubt, the impact of exogenous evolution, the transfer of traits beyond those hardwired into DNA, starts to increase. Unlike earlier Hominins, *Homo erectus* has short arms and long legs and appears more adapted to a terrestrial life. The earliest fossils exhibiting all the characteristics of modern humans date 200 thousand years ago or less. Mitochondrial DNA sequences support the hypothesis that all modern humans originate from a small population in Africa within the last two hundred thousand years. This small population probably results of a mass extinction event that creates the evolutionary space for the early *Homo sapiens*.

The earliest evidence for the use of symbols and artistic expression appear in Africa about 80-100 thousand years ago. We stress that the first steps into thinking in symbols indicates the development of more abstract thinking, allowing the construction of mental models of reality. Such mental models (Chapter 2) are important prerequisites for scientific thinking and a catalyst of further increasing the importance of exogenous evolution. These aspects of exogenous evolution seem to precede and likely trigger a rapid population expansion, migration and dispersal event 60-70 thousand years ago.

11.6. Use of tools.

Using tools is a sign of intelligence, and the theory is that tool use and most notable the continuing expansion of the human brain are an important co-evolution event. The expansion of the brain allows the development and use of more sophisticated tools whilst the evolution of larger brains results from the need to develop and use tools that are more sophisticated. The brain demands a high-energy consumption; hence, the evolutionary advantage must be large. As indicted earlier, the brain, at 2% of total body weight in a modern human, consumes about 20% of the total energy budget. Increased tool use allows hunting for energy-rich meat. This suggests that early human ancestors were under evolutionary pressure to increase their capacity to create and use tools and thus needed an increase in the capacities of their brains. Many species make and use tools, but the humans dominate the making and use of more complex tools. What species makes and uses the first tools? The oldest known tools are the "Oldowan" stone tools from Ethiopia; these tools are about 2.5 million years old. Since the first appearance of stone tools 2.5 million years ago, significant changes take place in the ways that humans manufacture and use stone tools.

Some authors distinguish five technological modes. The mode 1 technology represents an elementary approach strategy to creating sharp edges on a stone. Production of these tools involves simple reduction strategies and results in choppers, radial cores, polyhedral cores and retouched flakes. Their use includes cutting, scraping and chopping and these appear as early as 2.5 million years ago in East Africa, 1.8 million years ago in Eurasia and South-West Asia and 800 thousand years ago in Europe. *Homo habilis* produces and uses these tools.

Large bifacial tools such as hand axes and cleavers are representative of mode 2 technology. This involves transmitting preconceived designs and relies on increased intellectual capabilities. The earliest examples of this technology appear 1.6 million years ago in East Africa, quickly spread to China but appear as late as 500 thousand years ago in Europe. Users of these tools are *Homo ergaster/erectus* and later *H. heidelbergiensis*.

Mode 3 technology focuses on producing flakes of predetermined shape and size and requires abilities for abstract planning. The variation in types of tools increases, such as prepared cores, retouched flakes and projectile points to use in the manufacturing of spears for hunting. The technology is 300-400 thousand years old and associates with *H. neanderthalensis* in Europe*, H. heidelbergensis* and *H. sapiens* in Africa.

Mode 4 technologies include blade production using a special type of core reduction and retouching these into very standardized forms. The earliest appearance is roughly 50 thousand years ago, and is likely associated with the first appearance of anatomically modern *H. sapiens* in Europe.

In mode 5 technologies microliths appear. Microliths are very small flakes and blades. These types of tools first appear in Africa 30 thousand years ago.

11.7. Modern humans: The "Great Leap Forward".

Until about 40,000 years ago, the use of stone tools progresses stepwise. Each phase starts at a higher level than the previous one, but once that phase starts further development is slow. These *Homo* species are culturally conservative, but about 40 thousand years ago the pace of development increases. The term "Great Leap Forward" indicates this development. Modern humans start burying their dead, making clothing out of hides, developing sophisticated hunting techniques and engage in cave painting. Strong innovations in e.g. hunting technologies and human culture first start to appear some 50 thousand years ago. Modern human behavior includes aspects as abstract thinking, planning, innovation and symbolic behavior.

11.8. Out of Africa model of human evolution.

According to the "Out of Africa" model, the presently surviving *Homo* is not the first species of its genus: *Homo habilis* evolves in East Africa at least 2 million years ago. *Homo erectus* evolves more than 1.8 million years ago, and by 1.5 million years ago, this species appears throughout the Old World. Studies of DNA largely support a relatively recent African origin of our present species of *Homo*.

According to the Out of Africa model, modern *H. sapiens* evolves in Africa less than 200,000 years ago, begins migrating from Africa between 70,000 – 50,000 years ago and eventually replaces existing species in Europe and Asia.

11.9. Conclusion.

In this chapter, we trace the evolution of humankind. We show its emergence from the primates and particularly the great apes. The apes closest to modern man that are presently living and have a common ancestor with humans, are the chimpanzees. We stress the increasing importance of exogenous evolution that starts dominating the further development of the species. We left the story with primitive and more advanced tool making and the appearance of symbolic thinking and mental modeling of reality.

In the following, we focus on exogenous evolution and the changing characteristics of storing and processing information to allow further development to our present-day socioeconomic system. We discuss the histories of science and technology and highlight the industrial revolution and the emergence of industries and markets, as we know them today.

CHAPTER 12. HISTORY OF SCIENCE.

12.1. Introduction.

Science is the collection of empirical and theoretical knowledge about reality as we experience it. It represents an information set that reflects the results of the collective human efforts to obtain an understanding of the environment. Science evolves according to the laws that govern the evolution of information sets, as science is an example of such an (exogenously) evolving information set. It uses scientific methods that emphasize observation, explanation by inductive and deductive reasoning, and prediction of real world phenomena. An important element of the scientific method is the formulation of mental models of reality and associated mathematical models. This results in laws, theories and models. Science provides a very important competitive advantage as it, among others, enhances the discovery and use of sources of economic value. It represents part of the information that is required to transform potential value into economic value. Science is instrumental in reducing the statistical entropy of our picture of reality. The Scientific Revolution in 16th and 17th century Western Europe shapes modern science.

Here we present a concise summary; additional material provides e.g. Bunch and Hellerman (2004).

12.2. Pattern of the evolution of science.

Science is an information set that evolves according to the rules of the systems theory of evolution. This information set is present in the mind of people engaging in scientific thinking, in a written form in publications and textbooks and increasingly also in electronic form. This information represents the “genome” of science. The “phenotype” of science is its ability to model reality as it appears to us and to predict the outcome of new events. This phenotype is subject to a selection process that rests on comparing scientific predictions with observations. In addition, the usefulness of these predictions to the harvesting of economic value can be a selection criterion, particularly if we consider applied science. This selection process leads to changing the information set of science if conflicts arise between scientific predictions and the behavior of reality. Here we summarize a view on the nature of the evolution of science following the enlightening work of Thomas Kuhn.

Thomas Kuhn (Kuhn (1962)) indicates in his famous book “The Structure of Scientific Revolutions” that scientific progress follows a life cycle of development analogous to that of other evolving information sets. In the so-called pre-paradigm phase an established view on explaining a scientific problem does not exist, several related theories compete in explaining aspects of reality. This is akin to the embryonic phase in the general life cycle. Many new approaches develop by evolutionary radiation in the quest for explanation of a phenomenon. This is a period in which the scientific community tries to develop a common scientific approach that materializes in a so-called paradigm. After this period a consensus on the preferred paradigm develops. The paradigm enters a growth phase in which its application to a broad range of observations increases in sophistication in a kind of scientific “puzzle solving”. The scientific community locks in on a certain approach to the interpretation and explanation of the aspects of reality that are subject to the scientific investigation. This process continues until the paradigm enters the mature phase in which its explanatory power gradually becomes exhausted. This generally also is a phase in which conflicting evidence appears that challenges the validity of the paradigm that the scientific community adheres to. Generally, the scientific community resists the challenge of the reigning paradigm and continues adhering to the existing beliefs. This process continues until the evidence against the existing paradigm becomes overwhelming and a revolution results in a new paradigm that opens new avenues to explaining reality. In many instances, the innovation is due to outsiders in the existing paradigm. The resistance to change inherent in the scientific method is an important feature that we should consider also positive rather than only negative. If we allow too frequent changes, the scientific approach looses direction as it becomes subject to the restrictions of the copying fidelity limit that apply to all evolving systems. Resistance to change is essential to the realization of scientific progress.

To many people science is the ultimate temple of rationality. This results in the belief that scientific knowledge represents some absolute truth. It reflects itself in the advertisements for e.g. food ingredients, where suppliers claim that a beneficial health effect has been scientifically substantiated and proven. The reader that has carefully followed the discussion about the pattern of development of science presented in this section will by now suspect that this is definitely not the case. All scientific interpretations of the behavior of aspects of reality rest on assumptions and simplifications that may prove invalid when new experimental evidence, or a new way of interpretation of existing experimental evidence, emerges in the scientific community. Scientific interpretations and theories involve a creative inductive approach that does not follow strict logic. Science contains some aspects of an art. An example exists in the law of gravitation as proposed by Newton. This law is subject to revision when the theory of general relativity emerges more than 200 years after the publications of Newton. New experiments substantiate general relativity and the theory of Newton proves to be only an approximation of the complexity of reality.

This observation is of great significance in e.g. industrial research when the existing scientific state of the art predicts the impossibility of the development of products that would be superior in satisfying a need in the market. It sometimes pays to challenge the existing scientific state of the art by insisting on the development of such product. It is the way in which market needs drive the evolution of the scientific state of the art and the information set that characterizes science. In chapter 10, when discussing human evolution we saw such a need driven development in the co-evolution of the need to use more advanced tools and the capacities of the human brain.

12.3. Early developments in science.

In prehistoric times, knowledge passes orally from generation to generation. The development of writing enables storing knowledge and its communication with greater fidelity. When agriculture increases in sophistication a surplus of food results and more time for systematically progressing knowledge becomes available.

Many ancient civilizations collect astronomical information through observation in a systematic manner. In absence of knowledge about the nature of the stars and planets, many theoretical speculations emerge. Basic facts about human physiology exist and several civilizations practice alchemy, the precursor of chemistry.

Advances in Ancient Egypt include astronomy, mathematics, medicine and alchemy. Geometry is an outgrowth of the development of land ownership in support of agriculture. Significant other advances in geometry result from challenges in architecture.

Early classical studies of the celestial bodies take place, both for establishing a reliable calendar and in investigations that are more abstract.

The early Greek philosophers probe into question about the origin of the cosmos. The school around Pythagoras investigates mathematics in its own right. Pythagoras postulates the earth a sphere rather than a flat plane. In addition, Plato and Aristotle contribute to systematic discussions of natural philosophy. They introduce deductive reasoning in scientific inquiry. This inspires the efforts of others in later periods.

Factual observations, especially in anatomy, zoology, botany, mineralogy, geography and astronomy, are part of the heritage from this period. Furthermore, philosophical problems regarding change and its causes become the subject of scientific inquiry. Crucial aspects of the modern scientific method emerge, such as the awareness of the methodological importance of applying mathematics to natural phenomena. In addition, the practice of undertaking empirical research emerges. In the Hellenistic age, scholars frequently employ these principles from earlier Greek thought. Thus, clear lines lead from ancient philosophers to present-day science.

The Hellenistic astronomy develops a heliocentric model of the solar system. In a modified form, this is still the reigning paradigm. In addition, geographers accurately calculate the circumference of the Earth. The first systematic star catalog emerges. In medicine, the knowledge of human anatomy shows significant progress.

In the influential textbook Elements, the mathematician Euclid lays down the foundations of the concepts of definition, axiom, theorem and proof still in use today. Archimedes provides an accurate approximation of π. He also lays the foundations of physics in e.g. hydrostatics. Descriptions of plants and animals, establishing the first taxonomy, and categorizing properties of minerals develop. These developments take place between 500 BC and 100 AD.

India invents the production of steel and widely exports it by the year 1000 AD. Indian astronomers and mathematicians develop an accurate heliocentric model including elliptical orbits of the planets. Trigonometric functions, trigonometric tables, and techniques and algorithms of algebra develop. In the 7th century, scientists identify gravity as a force of attraction. An important development is the use of zero and the numeral system now universally in use.

China has a long history of contributions. Great inventions develop in navigation and in paper making and printing. These discoveries importantly shape the development of Chinese civilization and have a broad global impact. The Jesuit China missions of the 16th and 17th centuries lead to the diffusion of China's scientific achievements to Europe. However, factors of a philosophical and religious nature prevent accepting the notion of regularities that lead to laws of nature. Hence, the precursor of modern science fails to develop.

12.4. Medieval Science.

With the disruption of the Roman Empire, the Western Empire looses contact with much of its past. The Byzantine Empire still holds scientific centers such as Constantinople. In Western Europe, the monasteries take the lead as centers of knowledge until the medieval universities take over in the 12th and 13th centuries. These are the catalysts of an important intellectual revitalization in Western Europe.

Under the newly created Arab Empire, the Greek philosophy finds great support. With the spread of Islam in the 7th and 8th centuries, great advances in science result in the Islamic Golden Age that starts with the diffusion of the Islam in the 7th and 8th century and lasts until the 16th century. This development rests on several factors. The use of a single language, Arabic, facilitates communication. Access to Greek and Latin texts from the Byzantine Empire and Indian sources of learning provide a knowledge base to build upon. An early scientific method develops with important developments in the emphasis on experiments to distinguish between competing theories. Arabic numerals originate in India, but Muslim mathematicians refine the number system, such as by the introduction of the decimal point notation. Apart from mathematics, contributions to astronomy, chemistry and medicine materialize. In astronomy, the corrections made to the geocentric model later influence the Copernican heliocentric model. The foundations of modern chemistry emerge in this period. The works of Arabic scientists influence Bacon and Newton.

The contact with the Islamic world in Spain and Sicily, allows Europeans access to scientific Greek and Arabic texts. The European universities contribute to the translation and diffusion of these texts and are the seed of the new scientific infrastructure. Furthermore, contact with Indian and Chinese science influences European science.

At the beginning of the 13th century, Latin translations of the crucial works of almost all the ancient authors are available. This allows diffusion of scientific ideas via universities and the monasteries. Several important science practitioners extend and challenge the natural philosophy in these texts. Precursors of the modern scientific method appear in Grosseteste's mathematical approach towards understanding nature and in Bacon's empirical approach.

The first half of the 14th century witnesses important scientific work, within the framework of commentaries on Aristotle's scientific writings. Ockham introduces the principle of parsimony that involves avoiding the postulation of unnecessary entities. He introduces the well-known Ockham razor. The first seeds of the laws of motion appear.

A delay in philosophic and scientific developments occurs due to a number of disasters such as the plague. However, after the Fall of Constantinople in 1453, the pace of these developments increases again when many Byzantine scientists move to the West. The introduction of printing on the diffusion and development of knowledge in European society is also an important catalyst to progress. These developments herald the Scientific Revolution. This landmarks a decisive bifurcation point in the evolution of science.

12.5. Impact of science in Europe.

The Northern Renaissance introduces a decisive shift in focus from Aristotelian natural philosophy to more practical approaches, such as in chemistry and the biological sciences. Modern science in Europe enters a period of accelerated progress. Important factors are the Protestant Reformation and the Catholic Counter-Reformation, the discovery of the Americas, the fall of Constantinople and the rediscovery of ancient approaches. This creates an environment that stimulates questioning existing scientific beliefs. We argue that this is one of the prerequisites for progress of science through evolution. A suitable environment develops that stimulates questioning scientific doctrines. This coincides with the questioning of religious beliefs by the contributions of, among others, Luther and Calvin. Scientists become receptive to accepting that earlier scientific work sometimes conflicts with experience. In the terminology of Kuhn (1962), the scientific community opens the search for new paradigms.

This attitude of questioning paradigms and the associated search for new answers triggers a period of major scientific advancements and the Scientific Revolution emerges. This is a critical bifurcation point in the advancement of science that changes its further evolution. The Scientific Revolution begins in mid 16th century when books on anatomy (Vesalius) and astronomy (Copernicus) appear in print. Copernicus' book introduces the concept of the earth orbiting the Sun. A landmark publication is Newton's Principia. It appears in 1687. Other significant scientific advances emerge during this time by the works of several other important scientist of the period. The scientific method evolves to the modern approach that emphasizes experimentation and deductive and inductive reasoning and the formulation of mental abstractions and mathematical models. This culminates in decisive steps towards modern science during the 18th century. The works of Newton, Descartes, Pascal and Leibniz, pave the way to modern mathematics, physics and technology. In electricity (Franklin), mathematics (Euler, d'Alembert), biology, cosmology and mechanics important progress materializes. The impact is not limited to science and technology, but affects philosophy (Kant, Hume), society and politics (Smith, Voltaire).

12.6. Modern science.

The Scientific Revolution firmly establishes science as a precursor for the growth of knowledge and provides an impetus for the industrial revolution and economic growth. During the 19th century, the practice of science professionalizes and institutionalizes. This continues to date. The first industrial research laboratories emerge. The role of scientific knowledge in society grows and pervades many aspects of the functioning of the socioeconomic system.

A cycle of advances in technology and science develops. This allows the sustained and enhanced evolution of both. We again recognize the results of the workings of the general theory of evolution (Chapters 6). Whether advancement of science drives technological innovations or whether new technological achievements drive science is ambiguous because of the inherent cyclic nature of their sustained co-evolution. Technological innovations bring about new scientific discoveries and new technological possibilities derive from scientific discoveries.

The Scientific Revolution is a bifurcation point for the transformation of ancient science into modern science. Copernicus reintroduces a heliocentric model of the solar system. This inspires the model of planetary motion of Kepler in the early 17th century. Galileo further refines the model by the introduction of a key element of the scientific method in using experiments (observations of Brahe) to validate physical theories.

In 1687, Newton publishes the Principia and this introduces two important and successful bodies of theory in Newton's laws of motion, which lead to classical mechanics and Newton's Law of Gravitation, which describes the fundamental force of gravity. As noted earlier, this provides the scientific basis for the model of Copernicus. Several other scientists and practitioners study the behavior of electricity and magnetism during the early 19th century. Their efforts trigger a unification of these phenomena into a single landmark theory in the concept of electromagnetism that Maxwell introduces.

The beginning of the 20th century brings the start of a revolution in physics (see also Section 9.3). The established theories of Newton show not the absolute truth, although these agree with large chunks of experience with reality at the intermediate level we observe with our senses. We also sketched the "quantum revolution" in the beginning of the 20th century; it introduces the concept of the photon and invalidates the concept of an objective and deterministic reality. It shows that the laws of motion fail to apply on small length scales. In addition, the theory of general relativity shows that the fixed background of space and time, on which both Newtonian mechanics and special relativity depend, is an illusion.

The observation by Hubble in 1929 that the speed at which galaxies recede increases with their distance, leads to the understanding that the universe is expanding and induces the formulation of the Big-Bang theory (Chapter 9).

Developments during World War II lead to the practical application of radar and the development and use of the atomic bomb. Particle physics enters the phase of "Big Science", requiring massive machines, budgets, and large laboratories to test theories and move to new frontiers. State governments become the primary patron of physics. They recognize that the cycle of fundamental or strategic research and the progression of technologies is useful for both military and industrial applications.

The unification of general relativity and quantum mechanics into one theory is definitely still "Work in Progress".

The scientific base of modern chemistry begins with the distinction of chemistry from alchemy by Boyle in 1661. Lavoisier recognizes the role of oxygen and the law of conservation of mass. This eliminates the mystical concept of phlogiston, the reigning paradigm in those days. The theory of Dalton states that all matter is made of small constituents of matter, atoms that loose the basic chemical and physical properties of that matter on further division. After Dalton, substantiation and acceptance of atomic theory still takes 100 years, i.e. at the start of the 20th century. The synthesis of urea by Wöhler opens a new research field, organic chemistry. In fact, organic chemistry is an important precursor for the development of the chemical industry following the work on synthetic dyes by Perkins. By the end of the 19th century synthesis of hundreds of organic compounds results. By the twentieth century, systematic production of advanced compounds and materials from fossils provides a supply of many products, providing not only energy resources, but also synthetic materials for clothing and medicines. Large-scale industries and their industrial research start to drive progress.

Biochemistry results from applying organic chemistry to living systems. In the 20th century, science succeeds in explaining chemical properties from the electronic structure of the atom. Pauling's book on The Nature of the Chemical Bond (Pauling (1960) uses the principles of quantum mechanics to describe the molecular properties of increasingly complex molecules. This also is instrumental to the understanding of the nature of life's information carrier DNA and unravels a secret of life (Watson and Crick, 1953).

Further revolutions occur with the advent of genetic engineering in the early 1970-ies. This allows the editing of the genetic code and expands the possibilities of using microorganism to produce new products such as pharmaceutical proteins. These new medicines closely mimic the activity of regulator molecules that naturally appear in the human body. Science, a product of the evolution of the brain based on the DNA information set, leads through exogenous evolution to methodology to edit DNA beyond the blind force of mutation. This widespread use of genetic engineering, DNA sequencing and modeling techniques improves the prospects of directing the metabolism of microorganisms and leads to important opportunities for technological and industrial applications.

Midway through the 19th century, the focus of geology shifts from description and classification to attempts to understand how the surface of the earth changed over time. After the discovery of radioactivity, radiometric dating methods develop, starting in the 1900s. The concept of continental drift becomes accepted in the 1950s and 1960s and leads to the theory of plate tectonics. Since 1970, it is the unifying principle in geology.

Scientists speculate about evidence for the Big-Bang in the background radiation in the universe. In 1964, Penzias and Wilson discover an about 3 Kelvin background radiation as a further substantiation of that idea (Chapter 9). We also witness the first observation of the neutrino.

In the mid 19th century, the work of Jenner, Semmelweis and Pasteur leads to the germ theory of disease. Further research establishes this theory. This revolutionizes medicine. It also results in the discovery of vaccination, one of the most important methods in preventive medicine today. In the period after the Second World War, the availability of powerful antibiotic substances causes an additional revolution in the treatment of infectious diseases.

Pasteur invents the process of pasteurization, to help prevent the spread of disease through milk and other foods. His discoveries are also instrumental in enabling the widespread development of fermentation as an industrial technology.

A prominent theoretical development, part of the basis of the present book, is formulation of the evolution theory by Darwin in his book On the Origin of Species in 1859 (Darwin (1859)). We discuss elements of this theory in Chapters 6, 7 and 11. Darwinian evolutionary theory initially develops independent from knowledge of genetics. The study of heredity revives after the rediscovery of the laws Mendel. Mendel's laws provide the beginnings of the systematic study of genetics that becomes a major field for both scientific and industrial research. By 1953, Watson, Crick and Franklin clarify the role of the immortal coils of DNA. They identify it as the genetic material that codes for life in all its diversity. This leads to the Neo-

Darwinian synthesis in the approach to evolution.

12.7. Social sciences.

Successful use of the scientific method in the physical sciences leads to attempts to apply the same methodology to other fields of science. From this effort the social sciences develop. Although these developments are certainly important, particularly to the understanding of the behavioral aspects of exogenous evolution, these are outside the scope of this work.

Classical economics rests on Adam Smith's "An Inquiry into the Nature and Causes of the Wealth of Nations", published in 1776. He postulates an "Invisible Hand" that regulates economic systems through enlightened self-interest. Marxian economics rests on the labor theory of value and assumes the value of a good to depend on the value of the labor required to produce it. Under this assumption, capitalism based creation of profit rests on failing to pay the full value of the workers' labor. The Austrian school responds to Marxian economics by viewing entrepreneurship as driving force of economic development. This replaces the labor theory of value by a theory of supply and demand. For a more in depth discussion of relevant aspects of economic theory, we refer to Chapter 8.

12.8. Conclusion.

In this chapter, we present a broad-brush picture of the development of science, particularly the natural sciences. We argue that science follows the laws of the general theory of evolution as described in Chapter 6.

This observation leads to an important conclusion about the nature of scientific knowledge. We use observations and deductive and inductive reasoning to arrive at laws, models and theories about the nature of reality as we observe it when we interact with reality. This constitutes the information set of science, the "genome" of science. It is subject to selection in an evolutionary process and the selection criterion rests in sciences' ability to understand and predict the phenomena we observe in reality. Of course, reality is far too complex to develop a complete picture of reality. This already reflects itself in some of the methods of scientific analysis, e.g. the use of macroscopic models to reduce complexity. Therefore, science is subject to change, to mutation. This reflects itself in the revolutions in science in the beginning of the 20th century, when relativity and quantum mechanics invalidate long held laws such as Newton's law of gravitation and change the picture of time and space and a reality not influenced by our interaction with that reality. This phenomenon of incompleteness and mutability also holds for sciences' present picture of reality that suffers from problems with dark energy and dark matter and the recent picture of a universe in accelerated expansion.

This does not mean that science is unable to contribute its function as an instrument to increase the survival value and competitiveness of humankind. The incompleteness of science does not mean that it is not instrumental to develop tools, instruments, equipment and to obtain the products and services that we need. A prime example is the development of the heat engine, originally based on a misconception on the nature of heat.

We highlight some of these general characteristics of science and its evolution in this chapter and we return to these matters in the remaining chapters of this work.

CHAPTER 13. THE EVOLUTION OF TECHNOLOGY.

13.1. Introduction.

The history of technology concerns the invention of tools, techniques and equipment. It thus interacts in a cyclic way with the evolution of humanity and society. There clearly is an evolutionary pressure that drives said cyclic interaction. Technology allows the harvesting of potential value as economic value and thus improves the competitiveness and the standard of living of our species.

Technology also enables people to create new scientific tools allowing new scientific endeavors by investigation of aspects of reality not directly accessible to our senses and to probe environments we cannot physically visit. Examples are the electron microscope and the satellites that allow us to probe the universe. This is one of the mechanism by which science and technology co-evolve in a cyclical way. On the one hand, scientific progress may lead to new technologies as it may lead to inventions that translate in technology innovations. The distinction between inventions and innovations is that inventions represent an idea for a new technology, innovation results if this idea is reduced to practice in applications in industry or broader in society. In addition, technology or innovations may open new approaches to scientific investigation and may result in new challenges for the scientific community. There are many examples of both interactions as we have seen before and illustrate further in the remainder of this work. Technology is a product and a precursor of science. It is also a force driving economic growth in terms of the growth of the Gross Domestic Product. Technology profoundly influences everyday life and the competitiveness of our species.

13.2. Drivers of technological progress.

Technology, be it the use of stone tools, the use of the heat engine or internet technology, is one of the primary drivers of the evolution of humankind and its civilization. Technological milestones characterize major stages of social evolution, like fire, the bow, and pottery in the early era, domestication of animals, agriculture, and metalworking in the intermediate era and the alphabet, writing and information technology in the later era.

White (Peace (2004)) considers energy intensity the measure to judge the evolution of culture. This agrees with some of the concepts we discussed so far. He argues that human culture serves to get access to energy resources and to use these in such a way that these effectively support the needs of humankind. We term this control over statistical entropy to create free energy out of energy or economic value out of potential value. We develop this concept further in Chapter 16 where we show the strong correlation between GDP and the use of energy resources. White distinguishes five stages of human development. In the first stage, people use the energy of their own muscles. Subsequently, in the second phase, they use power of domesticated animals. The third phase introduces the use of the energy of plants (agricultural revolution). In the fourth phase the use of the energy in fossil resources, such as coal, oil and gas, develops. The fifth involves harnessing nuclear energy. Recent experience makes it questionable whether this fifth period provides a lasting solution to the need for energy resources in the economic system. We add a phase we just enter. It introduces the challenge of the direct use of solar radiation and derived renewable sources of power, such as wind, water or biomass. This must result in the ultimate solution for realizing a sustainable supply of energy resources. We further discuss this sustainability issue in Chapter 16.

Lenski (Lenski and Nolan (2005)) uses a different approach and focuses on information. When the information and knowledge in human society increase and more effective ways of their communication develop, a more advanced cultural stage results. Again, we clearly recognize the theory this book develops. In our approach information and its communication is the pivotal factor in transforming potential value into economic value. Economic value being potential value that information makes available to do useful work. Lenski identifies four stages of human development, based on advances in the technology of communication of information. In the first stage, genes pass information. In the second phase humans learn from experience and pass on information. In this phase, the first signs of an ever-increasing importance of exogenic evolution appear (Chapters 6 and 7). In the third phase humans use signs and develop logic. In the fourth, they create symbols and develop language and writing. Again, we add a period in which humankind develops mental models and mathematical models of reality to progress science as a means of accessing potential value in the environment. This is a key factor in technology innovation and it drives the advancement of human society. Advancements in the technology of communication and processing information translate into advancements in the economic and the social system.

Alvin Toffler (author of Future Shock) approaches the theories of post-industrial societies, arguing that the current era of industrial society ends. Services and information become more important than material goods. We argue that this distinction is artificial if we use a broader definition of the products of an industry and include, as we do, services and information as economically valuable products.

13.3. Geographical and temporal aspects of the evolution of technology.

During the Stone Age, human life involves limited use of tools and permanent settlements remain rare. In this period, the major technologies develop from the need to survive and to support activities as hunting, food preparation and warfare. As we argue, evolutionary pressure drives their development. Fire, stone tools, weapons and clothing are technological achievements of major importance during this period. In the later Stone Age, the Neolithic period, the seeds of a primitive agricultural technology develop. In this period, new approaches towards design and manufacturing of stone tools appear (Chapter 11). Advanced stone axes allow forest clearance and the establishment of crop farming.

The development of agriculture coincides with a shift from nomadic life to settlements. We can understand this by the fact that living in groups in settlements increases the pace of learning by doing as an important driver of exogenic evolution. In fact, we see evidence of an increased rate of exogenic and cultural evolution in the fossil record.

The Neolithic period involves radical changes in agricultural technology, including animal domestication and the increased appearance of permanent settlements. These developments create room for the evolution of technological innovations by their positive influence on improved learning by doing as we just mentioned. Examples are the developments in the use of metal technology in the production of tools, with copper and later bronze being the materials of choice. In this period new means of communication in the form of language develop. The belief is that the development of agriculture preceded the development of more sophisticated means of communication.

The introduction of iron as a material makes it possible to produce stronger tools. In addition, iron is more readily available than bronze. These developments take place in the so-called Iron Age. In the Iron Age, the last major step occurs towards the development of written language as a crucial catalyst of exogenous evolution.

Originally, the technology for mass production of steel is not available. However, mechanical processing at high temperature is in use. Using more effective iron axes increases the pace of land clearance and provides more farmland to support the growing population. This in its turn reduces the amount of time humans need to get access to food. Thus, more time remains for efforts to enhance the progression of technologies. The pace of exogenic evolution increases.

In the Hellenistic period, we observe a sharp rise in technological inventiveness, both with respect to existing and new technologies. Openness to new ideas, the blossoming of a mechanistic philosophy and the establishment of the Library of Alexandria contribute importantly. Innovations of the Ancient Greeks include one off the first step towards using energy resources beyond the power of muscles by using watermills and wind power. As we argue, improved access to energy resources is an important precursor of growth of the standard of living and the pace of exogenic evolution. The full potential of such increased access to energy resources develops much later in the industrial revolutions and shows indeed dependent on increased access to sources of energy. Developments in irrigation technology enhance the productivity in agriculture.

The Romans sophisticate agriculture, expand iron-working technology, create laws for individual ownership and advance architecture and infrastructural technology (e.g. road building). They also progress military engineering. Early applications of concrete develop. Achievements of the Inca and the Maya are in, among others, construction and engineering. Their technology also includes an efficient agriculture using irrigation canals and drainage systems.

European technology in the Middle Ages provides a synthesis of tradition and innovation. New approaches to architecture allow construction of cathedrals and fortifications.

The Islamic world witnesses a “Green Revolution” involving a fundamental transformation in agriculture. Due to the global economy, the Muslim traders establish across the Old World, many crops and farming techniques migrate between the different parts of the Islamic world. The first steps to urbanization become visible. They also create new innovative approaches towards increasing the availability of energy (tidal power and wind power). This induces the appearance of factories on a larger scale than ever before.

The Renaissance era shows profound technical advancements, e.g. the printing press and patent law. As we argue, the communication by printing enhances the evolution of science and technology. Patent law has, as is witnessed by the developments during the industrial revolution in the 18t and 19th century, an enhancing effect on the pace of the innovation in technology. In the Renaissance, science and technology enter the familiar cycle of co-evolution. This increases the pace of advancement in both science and technology. Ultimately, this triggers the Scientific Revolution. This provides one of the seeds of the industrial revolutions in the 18th and 19th centuries.

In the Age of Exploration that begins in the early 15th century and lasts until the 17th century, we witness the discovery and the colonization of the Americas. This proves to be one of the additional seeds of the industrial revolution later on.

13.4. The industrial revolution.

The development of the steam engine is one of the important drivers of the British Industrial Revolution. It induces developments in the areas of textile manufacturing, mining, metallurgy and transport. The revolution roots in accessing new energy resources in the form of coal, produced in ever-increasing amounts from the vast resources in Britain. Coal converted to coke allows cast iron availability in much larger amounts than before. Coal removes the constraints of waterpower driven mills. The steam engine helps draining the mines providing access to more coal. The steam engine technology and its resources thus become part of a self-propelling cycle driving the further evolution of industry. The development of the high-pressure steam engine makes locomotives possible and another necessary revolution, a transport revolution, follows. We further discuss the industrial revolution in the next chapter.

The 19th century sees strong developments in transportation, construction, and communication technologies mainly originating in Britain. Other technologies, such as the light bulb, emerge. In the late 18th and the early 19th century machine tools to manufacture other machines appear. Steamships allow opening up Japan and China for trade with the West.

The Second Industrial Revolution in the mid 19th century sees rapid development of chemical, electrical, crude oil and steel technologies. This is coincides with and results from a strong growth in industrial research.

20th Century technology develops rapidly. Communication technology, transportation technology, (university) teaching and implementation of scientific methods and increased (industrial) research spending, propel the advancement of modern science and technology. The National Academy of Engineering (www.greatachievements.org), by expert vote, establishes the following ranking of the most important technological developments of the 20th century: Electrification, automobile, airplane, water supply and distribution, electronics, radio and television, mechanized agriculture, computers, telephone, air conditioning and refrigeration, highways, spacecrafts, internet, imaging, household appliances, health technologies, petroleum and petro chemistry, laser and fiber optics, nuclear technologies and materials science.

In the 21st century, technology develops even more rapidly, especially in informatics, electronics and biotechnology.

13.5. Conclusion.

In this chapter, we highlight the development of technology and its role in the evolution of society. We discover that the pace of technology increases when the scientific revolution allows the closure of a cycle in which science and technology co-evolve in a process of sustained evolution. With the closure of this cycle, the question whether science drives technology or vice versa, becomes a debate about the beginning and the end of a circle. A meaningless exploration once the circle becomes closed. We note that technology is just as science an information set that is subject to the systems theory of evolution discussed in Chapter 6. In fig. 13.1, we illustrate the co-evolution of science and technology.

Fig. 13.1. The co-evolution of science and technology

We see the impact of developments in the harnessing of energy sources such as coal by the introduction of the steam engine. The steam engine is not only the father of the second law. It also is an important factor in triggering the industrial revolution that increases the pace of evolution and is a catalyst of the growth of the per capita Gross Domestic Product. The industrial revolution also catalyzes the advancement of science both by being a driver of the science-technology cycle and the emergence of large-scale industrial research. We recognize a clear case of autocatalysis in both the development of science and technology and their co-evolution.

We analyzed how technological progress reflects itself in making sources of free energy and economic value accessible and in the development of evolving information sets to fight statistical entropy and to increase the share of the potential value accessible as economic value.

The industrial revolution is too important to the subject of this book to leave our discussion of this phenomenon here. The next chapter focuses on the industrial revolution.

CHAPTER 14. THE INDUSTRIAL REVOLUTION.

14.1. Introduction.

The Industrial Revolution marks a period in the 18th and the 19th century when major changes in agriculture, manufacturing, chemistry, mining, communication, scientific knowledge and transport, have a profound effect on socioeconomic, cultural and institutional changes (e.g. More (2000), Ross (2008)). It starts in Britain, and then spreads throughout the world. The onset of the Industrial Revolution is a major bifurcation point in human history. It influences many aspects of human life in a decisive way. We see a dramatic increase in the standards of living, particularly in the second industrial revolution that starts around 1850. In addition, a period of increased population growth marks the beginning of the industrial revolution.

Several factors have been proposed for explaining the onset and the deployment of the industrial revolution. Growing demand for goods and services results from an increase in the population by roughly a factor 3 between 1750 and 1850. Furthermore, there is a growth in demand and in the nature of demand by an increase in per capita GDP of roughly a factor 2. Also important is the agricultural revolution that started at the end of the 17th century. It leads to a strongly increased effectiveness of agriculture and decreasing food prices. This induces a shift of demand from basic needs, such as food, to more advanced products and services. In addition, the revolution in agriculture leads to an increased wealth of landowners and agriculturalists in a broader sense. The demand this upper class creates leads to a spin off to other activities related to the supply and distribution of goods and services. This creates a middle class with an increasing standard of living and spending power. In this way, a self-propelling supply and demand cycle develops. Furthermore, the revenues out of agricultural activities lead to availability of capital. In addition, an effective capital market develops. In addition, human capital plays a role as a sound scientific infrastructure exists as well as a skilled labor force. Another important issue is the development of an efficient transport infrastructure. As we will indicate later, this is a prerequisite for the emergence of industrial activities. The existence of a stable government and a reliable legislative environment stimulate the creation of economic activities. In the opinion of this author, also the scientific revolution of the 17th century plays an important role. Perhaps initially not so much by the new possibilities scientific progress offers for the advancement of technology by driving the technology science cycle, but rather by creating a climate that stimulates challenging existing beliefs and approaches. This creates an environment that induces inventions and innovations. In much the same way, Protestantism and openness to new ideas are important facilitating factors. In a longer-term perspective, this author feels that the need for increased competitiveness, i.e. a selection induced evolution of technological possibilities, triggers the overall process and science further develops as an instrument to increase the effectiveness of this development. Science and technology were and are instrumental to reduce the statistical entropy of our picture of reality and this allows humankind to transform potential value into economic value in an increasingly effective way.

The First Industrial Revolution, which starts in the 18th century, evolves into the Second Industrial Revolution in the mid 19th century, when technological and economic progress gain momentum. Later in the 19th and in the 20th century, the internal combustion engine and electrical power generation become additional sources of progress. Increasing scientific knowledge catalyzes the first and particularly the second Industrial Revolution and triggers an enhancement of scientific progress by increasing institutional research and the emergence of industrial research, again notably in the second industrial revolution. Some twentieth century historians argue that the process of economic, technological and social change took place gradually and the term revolution is not an adequate descriptor. However, Gross Domestic Product (GDP) per capita, broadly stable before the Industrial Revolution, increases drastically (Fig 14.1) and the modern capitalist economy emerges (Maddison (2007)). The Industrial Revolution triggers an era of unprecedented per capita economic growth in capitalist economies. We will discuss these developments more extensively in Chapter 16. There seems broad agreement that the Industrial Revolution is one of the most important bifurcation events in the history of the evolution of humankind.

Fig. 14.1. Per capita GDP (1990 Int. $/capita) versus time for W-Europe and the USA.

14.2. Causes of the industrial revolution in Britain.

The causes of the Industrial Revolution in Britain remain a topic for debate. A convenient starting point for the analysis of the causes rests in the studies of the determinants of economic growth as provided by contemporary economic theory (More (2000)). We summarize some of these approaches in Section 8.9. Particularly the endogenous theories of economic growth that consider the triggers for economic growth to arise from the economic activities themselves rather than being caused by factors in the environment, seem to be fruitful in such analyses.

There are two important approaches in endogenous theories of economic growth that we highlight. To present these approaches, we first summarize the theory of production. In the theory of production, we start out identifying the main variables that influence production in industry. In the traditional approach, the theory considers labor and capital in production assets as the determinants of production. In our approach, we would like to include intellectual assets as a production factor. The established approach to production theory does include this factor only in a limited way. In our approach, intellectual assets include, among others, human capital, the skills of the workers in the company, the competence base of the company and other tangible and intangible assets such as knowhow and intellectual property. These intellectual assets serve to improve the efficiency of the use the production factors capital and labor. This increased efficiency leads to a higher return on these primary assets and this results in increased profits and the accumulation of capital.

One of the factors that triggers the industrial revolution in Britain is the availability of capital that is in search for profitable investment opportunities. This “surplus” capital derives from an agricultural revolution that takes place at the end of the 17th century. As stated before, it leads to increasing profits and capital accumulation for landowners and other agriculturalists. In addition, it results in lower prices for basic foods. These combined factors lead to increased demand for products and services beyond the basic needs, such as food. Industries develop to supply products that serve these needs.

Another crucial factor is inventions and innovations. As we indicated in the preceding section, the cultural climate in the period in which the industrial revolution kindles in Britain fosters invention and innovation activity. In this way, the increasing demand is supported by a developing industry that supports those needs as it can supply products and services in a cost effective way. Some authors consider inventions and innovations as a crucial catalyst for the industrial revolution and we agree with this view. The investments in innovation are supported by a patent system that allows innovators reaping the benefits of their investments in innovation. It is important to stress that invention and innovation is a competence that once it evolves in one sector of industry colonizes other industrial sectors. This leads to an autocatalytic process in which the industrial revolution affects a broader range of industries. More (2000) distinguishes two broad categories of innovations. Macro innovations are of a revolutionary nature. These constitute the embryonic phase of the development of the life cycle of a new industrial activity (section 8.2). Subsequently, learning by doing, leading to micro innovations, leads to the growth phase of the industry. After a while, the potential of the initial macro innovation gets exhausted and the industry enters the mature phase. This leads to diminishing returns. This triggers, analogous to the mechanisms governing biological evolution, a new macro invention and a new more effective approach to harvesting the opportunities in the market develops.

Also the developments in transport need mentioning. The existence and the development of an efficient transport infrastructure leads to the possibility to concentrate industry in larger scale activities. This leads to increased possibilities for division of labor across industries and improves the effectiveness of industry by specialization and enhanced learning by doing. The concentration of industry also provides the necessary market size to allow profitable investment in innovations leading to increased efficiency.

A final factor that triggers the industrial revolution is the availability of new abundant energy resources in the form of coal. This leads to the introduction of the steam engine that in a co-evolution cycle with the availability of coal, results in increasing availability of sources of useful work beyond those in waterpower and the power provided by the muscles of animals and humans. Following the view of White (Section 13.1) this leads to and advancement of the standard of living. We provide additional substantiation of the importance of this factor in Chapter 16.

To the author’s opinion, this brief summary adequately covers the factors that lead to the industrial revolution in the mid 18th century in Britain.

14.3. Role of inventions and innovations.

In the preceding sections, we discuss the general framework of the factors that trigger the industrial revolution in Britain. In this section, we identify the specific industrial sectors in which innovations cause growing industrial activity in the second half of the 18th century:

- In cotton spinning, new technology develops rapidly followed by the erection of many cotton mills. Similar technology proves to apply to other textiles. In addition, patenting possibilities contribute.

- The steam engine develops, based on the availability of energy resources in the form of e.g. coal. Initially it is mainly in use for pumping water from mines. As stated earlier, this triggers an autocatalytic co-evolution cycle that leads to increasing availability of coal and increasing availability of steam engine based capacity to perform useful work. In addition, the steam engine proves also useful to power machines. This enables rapid development of efficient semi-automated factories on a large scale in places where waterpower is not available.
- Another sector that leads to industry innovation concerns the production of iron by e.g. the use of coke instead of charcoal in its production.

These represent three “leading sectors” of key innovations that increase the creation of economic value and in its turn, this fuels the further deployment of the Industrial Revolution. In Chapter 6 discussing the General System theory of evolution, we identified communication and further refinement of information as an important factor allowing evolutionary progress. In the age in which the industrial revolution kindles, knowledge transfer takes place by various means. One mechanism involves transfer by workers trained in the technique moving to another employer. To date this aspect of Human Resource Management remains an important factor in industry innovation. Another common method exists in study tours to gather information. Further, the network of informal societies, in which members meet to discuss scientific issues and their application to manufacturing, contributes to the diffusion of knowledge. To date specialized congresses and workshops are important to the disseminating of knowledge to complement a company’s information set and to further scientific and technological progress. Such societies start publishing proceedings and transactions. Over time, scientific publications in an increasing number of journals contribute importantly to scientific progress and the application of science in industry. On line publishing, which evolves with the emergence of the internet, increases the pace of dissemination and its timeliness. Periodical publications about manufacturing and technology start to appear in the last decade of the 18th century and regularly include notice of the latest patents. What also contributes is the increasing interaction between academic and industrial research. In addition, governments contribute in this respect by an active innovation policy by e.g. funding research in universities, industries and academia-industry collaborations.

14.4. Continental Europe.

The Industrial Revolution on Continental Europe starts later than in Great Britain. In many industries, it involves the application of technology developed in Britain in new places. In addition, purchasing the technology from Britain and British engineers and entrepreneurs moving abroad in search of new opportunities, cause penetration in continental Europe. The German, Russian and Belgian governments provide state funding to the new industries. In some cases (such as iron), the different availability of resources locally, means that only some aspects of the British technology apply. This pattern of the first industrial revolution mimics as we said that in Britain, albeit with differing accents and a bit later. Therefore, to avoid repetition, we omit details as far as the industrial revolution in the rest of the world is concerned.

14.5. The second Industrial Revolution and later evolution.

The strong demand of the railways for more durable rail leads to the development of the means to cheaply mass produce steel. Some consider steel as the first of several new areas for industrial mass production, characterizing a “Second Industrial Revolution”, beginning around the mid 19th century. The technology to mass produce steel becomes widely available in the 1870s.

This second Industrial Revolution gradually starts to include the chemical industries, petroleum refining and distribution, electrical industries, and, in the twentieth century, the automotive industry, the pharmaceutical industry and the computer industry. It involves a transition of technological leadership from Britain to the United States and Continental Europe, particularly Germany. Many of our present-day multinational corporations emerge wholly or in part at the beginning of this second revolution. We discuss some notable examples in the Chapter 15.

The chemical industry no doubt is a prime example of the development of a science based industry. The birth of the chemical industry traces back to the discovery of the synthetic dyestuff mauve by Perkin in 1856. It triggers an unprecedented growth of the synthetic dyestuffs industry in the late 19th and the early 20th century. Before the discovery of Perkin inorganic chemicals, more or less directly derived from mining minerals to produce acid and alkaline chemicals, dominate the emerging chemical industry. After Perkin's discovery, organic chemistry starts to dominate the chemical industry and remains leading ever since.

The birth of the chemical industry also marks a shift in the resource base of the industry. Before the science based chemical industry, minerals, obtained by mining and agriculture derived chemicals, constitute the raw material base of the industry. The dyestuff industry marks a first shift to fossil resources. These initially derive from coal, later on increasingly from crude oil and natural gas.

One cannot overestimate the impact of the chemical industry on society. It revolutionizes agriculture and the supply of foods and processed foods, it provides new materials through the development of polymers, and it makes it possible to develop new pharmaceutical products based on synthetic chemistry. Many of the products we routinely use today are unavailable without the activities of the chemical industry. In addition, many of today's large multinational industries emerge in the revolution that creates the chemical industry. No doubt, the chemical industry importantly contributes to the strong growth in per capita GDP that characterizes the period following the birth of the Second Industrial revolution.

The increasing availability of economical petroleum products reduces the importance of coal and further widens the potential for industrialization in the energy and chemical industries.

By the end of the 19th and the beginning of the 20th century, industrialization in these areas results in the creation of the first large industrial corporations with global interests, such as companies like U.S. Steel, General Electric, Bayer, Royal Dutch Shell, Unilever, Pfizer and Roche.

14.6. Conclusion.

In this chapter, we briefly analyze the causes and effects of the two industrial revolutions and recognize some features of the systems theory of evolution in the industrialization of society. We highlight the crucial role of availability of capital and innovations.

The industrial revolution triggers a strong growth in per capita Gross Domestic Product. These revolutions provide a decisive development in the evolution of our present-day socioeconomic system. The focus of our discussion refers to the first wave of industrialization. We say more about the second wave when we analyze some cases of the evolution of individual multinationals and discuss some general features of present-day industries and economies in Chapter 15. Other than in the study of mankind, where we largely have to rely on fossil records as written accounts of early human evolution do not exist, written accounts of the history of multinational are available and the task becomes a little easier than that of the scholars that trace the emergence of present-day humankind.

CHAPTER 15. THE DEVELOPMENT OF SOME OF THE LEADING FIRMS.

15.1. Introduction.

In this chapter, we analyze the evolution of some leading industrial corporations from the perspective of the theory of the firm we developed in this book and earlier work (Roels (2010)). We follow the development path of leading corporations since the time they emerge as a product of the industrial revolution. We discuss several instruments by which corporations evolve over time and we particularly highlight the factors that induce organic growth as well as the circumstances that lead to growth by cooperation, acquisition and merger. In section 15.2, we perform an in depth analysis of the development of Gist-brocades and try to derive general features of the development of industrial corporations from this example. Section 15.3 analyzes the development of selected leading industries. We close this chapter by revisiting some features of corporate strategy. In addition, we try to formulate some regularities that characterize industry evolution.

15.2. An evolutionary view of corporate development: Gist-brocades.

This section analyzes aspects of corporate strategy and corporate development from the perspective of the EVT formalism. We perform a broad-brush analysis of the development of the Dutch company Gist-brocades. This presents a stepping-stone for the formulation of some general features of corporate strategy and corporate development. For convenience sake, we study the period of evolution of Gist-brocades unto the point where it merges with DSM in 1998.
Gist-brocades emerges as "Nederlandse Gist en Spiritus Fabriek" (NGSF) in 1870 as a producer of baker's yeast and the associated product alcohol. Its founding rests on a clear market need as the community of bakers asks for more reliable alternatives for the yeast sources available in those days. The manufacturing of the products of the new company rests on fermentation technology using agricultural resources, such as grains. Fermentation is in those days a quite elusive and new technology. The scientific understanding that microorganisms are the source of useful products such as ethanol creates the new technology. Today, we call this the competence to use a microorganism, in the case of baker's yeast the eukaryotic organism *Saccharomyces cerevisiae,* to produce useful products. In this case, these products are baker's yeast for the leavening of bread and ethanol for e.g. beverages and perfumes. In the 1990s, Gist-brocades is a co-leader in the European market for baker's yeast and it discontinued the production of ethanol earlier.

The first decades of the 20th century, witness the addition of industrial chemicals such as acetone and butanol, geared at totally different markets, to the portfolio of products of the company. The basis of this expansion is that fermentation, in this case fermentation of bacteria, allows the production of these products from agricultural resources, such as grains and sugar containing crops. The company achieves this step into a new market by expansion into a slightly different, but much related, competence than the one the company is familiar with through its baker's yeast activities. Fermentations based on bacteria expand the competence base and the information set of the company by adding a different, but closely related, type of information. Today, we consider this an example of competence-based expansion. It allows the company to extract economic value from a market need new to the company. The First World War fuels this opportunity as it shifts the pattern of demand for industrial chemicals.

We expand a little bit on the notion of a competence. A competence (Prahalad and Hamel (1990)) is a complex integration of disciplines, technologies and other types of knowledge and routines. In general, it allows the production of useful products for a number of markets. In addition, it is difficult to copy, because it is a complex integration of disciplines honed by competing for sources of economic value over a number of years. Finally, it has a distinctive influence on the ability of the company to make products that allow it to couple to sources of economic value, fueled by market needs. It has a distinctive impact on competiveness. A competence is, in line with the general formalism of EVT, of an informational nature, it is part of the "DNA" of the company. The recognition of the importance of competence development results the founding of an R&D department in NGSF at the end of the 19th century. In this respect, the company follows a general feature of industry in those days.

Today, the production of industrial chemicals is the realm of the fossil resources based industries and the products disappeared from the portfolio of the company. Clearly, the company is unable to follow the shift in the resource base for the production of the products, as this requires a very different information set. With the disappearance of these products, this disappearance of part of the phenotype, the fermentation competence remains part of the genotype of the company.

At the end of the Second World War, a new opportunity arises to expand the business scope of NGSF. During the War, the Allies develop Penicillin production based on the fermentation of the mould *Penicillium chrysogenum.* This rests on the discovery of Penicillin as an antibiotic substance in 1928 by Fleming (1929). The Allies use penicillin at the end of the Second World War to treat battlefield infections, an important cause of impairment of soldiers. Again, the fermentation capabilities, competence based expansion, allow NGSF to enter a new market, the pharmaceutical market, shortly after 1945. At the date of the merger with DSM in 1998, Gist-brocades is the largest producer of Penicillins and derived substances worldwide.

Penicillins naturally occur in two varieties, called Penicillin G and Penicillin V. These products share a chemical structure called the β-Lactam moiety, but differ in a side chain attached to that moiety. Penicillins prove to be quite successful antibiotics and variations on the chemical structure of the natural penicillins develop. These are the so-called semi-synthetic penicillins. Ampicillin and amoxicillin are notable examples. Today, amoxicillin is still widely used in antibiotic therapy. To enter the market for semi-synthetic penicillins based on the fermented product Penicillin G, the company needs to acquire a new skill, a new information set. Expertise on the chemical modification of the fermentation product needs development. The company develops it in a largely organic way, based on sizeable R&D-efforts.

At the end of the 1950s, the kernel develops for an important new major step in the beginning of the 1960s. A new competence based expansion in fermentation comes in reach. The introduction of protein splitting enzymes in household detergents leads to a large new market for products based on fermentation. A range of other enzyme-based products adds to the potential. The market for industrial enzymes develops. The company evolves to a large player in industrial enzymes, but it sold its activities in some of the larger commodity enzymes in the 1990s. The expansion into the industrial enzymes market needs new additions to the competence base. Application expertise and modern genetics, e.g. recombinant DNA technology and later on genomics, for the production of enzymes, augment the company's genome. This again largely derives from R&D driven organic development, although the company acquires some critical technologies from third parties. A distinctive new competence complements the skill bases in fermentation technology and chemistry, resulting in broadening of the company's information set.

At the end of the 1960s, the earlier entry of the company in the pharmaceutical market through penicillins prompts a major move. The company merges with Brocades, a company operating in the end user market for pharmaceutical products, to form Gist-brocades. This is an example of a customer-based expansion of the company, completely different from the earlier competence-based expansions.

With the merger, the company acquires new capabilities in marketing pharmaceutical specialties and the development of new pharmaceutical products. That merger takes place in an era when the pharmaceutical industry shows strong growth. In the eighties and nineties of the 20th century, the pharmaceutical industry shows signs of consolidation and the players in the industry rapidly grow in size and become dependent on sharply increasing efforts and investments in the development and marketing of new medicines. The activities of Gist-brocades do not keep pace with these developments and the company decides to sell its pharmaceutical activities to the Japanese pharmaceutical company Yamanouchi at the end of the eighties. However, it retains its position in fermentation and chemistry based ingredients and active products for β-Lactam antibiotics.

The beginning of the 1980s, welcomes a new customer-based expansion of Gist-brocades, again largely by acquisition. It expands its position in the bakery market by moving into products for the production of bread and pastry beyond baker's yeast. The company enters the markets for bread improvers and pastry ingredients. This leads again to a sizeable expansion of its competence base, the information set on which it operates again diversifies. The move also leads to new outlets for its existing technology base in enzymes, as bread improver enzymes complement its product portfolio. Such market based expansions, witnessed by Gist-brocades' expansion in pharmaceuticals and bakery products are customer-specialist types of expansion.

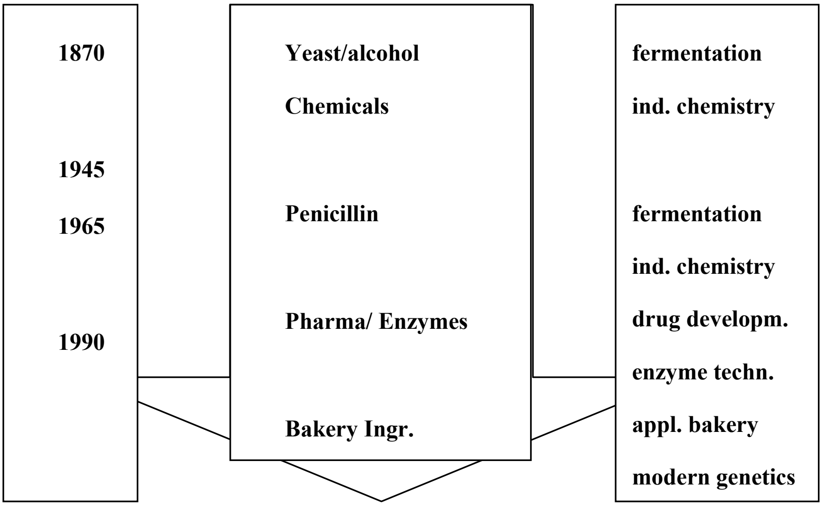

Fig. 15.1. The corporate history of Gist-brocades.

Fig. 15.1 summarizes this bird's eye view of the development of Gist-brocades unto the merger with DSM.

After this discussion of Gist-brocades and its historical development and evolution, we generalize the picture. A company often starts out as a narrow market-competence combination. It enters a "niche" market representing a need to which it couples to generate economic value. To supply the product to couple to the economic value source, it needs an adequate information set, a competence base. It can acquire it from the outside or it can develop it organically. In addition, the reverse can apply, the company may have a captive competence base, e.g. based on new scientific developments, for which it sees opportunity to create and couple to a source of economic value in the market. In this way, an initial customer base and matching competence base result. In the beginning, this competence base is far from perfect and it further develops, and the customer base expands by co-evolution with the gradually perfecting competence base or information set.

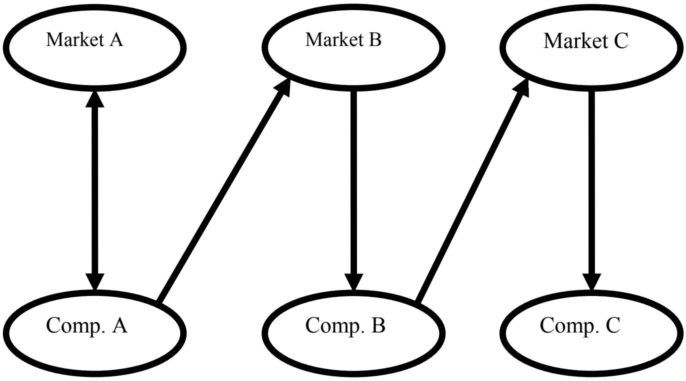

Fig. 15.2. Co-evolution of market exposure and competence base.

After a while, the company sees a new outlet for its activities. As an example, this exists in new markets that become available to the company based on its existing competence base. This is a critical strategic moment in the history of the company. In almost all cases, such expansion involves larger statistical entropy, uncertainty, than furthering the existing customer base/competence combination. It involves a substantial risk and exposure, and the future success is far from certain. The company management needs to make a bounded rationality decision to grasp the opportunity or not. Both decisions have an important historical significance in terms of the unknown future development paths that become open to the company. In terms of the theory of evolution of dissipative structures, the company reaches a bifurcation point.

Number of markets	Number of competences	
	one	many
many	competence driven	one stop shop
one	niche player	customer driven

Fig. 15.3. Corporate development strategies.

After entering the new market, the company no doubt finds out that its competence base does not optimally suit the requirements of the new market and it sees the need to evolve its information set to serve the new market in an optimal way. This results in gradual organic adaptation of its competence base to the new customer base, or, alternatively, acquisition of an already existing competence in the market. Such acquisition is again a major critical strategic move that may further shape the range of futures that are open to the company. Fig. 15.2 shows this development. After a while, the process repeats itself and the competence base needs to change accordingly. The figure illustrates the co-evolution of the company and its competence base.

Fig. 15.3 illustrates the strategies deployed in the development of a corporation. An emerging company starts out as a niche player based on one market and a product. From there it can move in two directions. It can take its competence base as a starting point and move to new markets. Its competence base stays rather homogeneous, although it augments its basic competence with market specific aspects of the competence. Market specific knowledge and applications expertise integrate in the competence base. In many companies, the growth resulting from this strategy leads to a partial splitting of the company in market specific business units. The business units develop their own market specific competence base, and this relaxes the span of control of management. In addition, a greater rate of experimentation is possible without getting in problems with the copying fidelity limit. The problem associated with this move, is that the focus on the basic competence diminishes and this challenges harvesting the synergy fruits of the combined skill bases. This mitigates if the business unit structure does not follow the entire internal value chain, e.g. if the basic competences remain concentrated in a central R&D unit.

The second approach with reference to Fig. 15.3 is that the company sticks to its market and develops a broader range of competences to support a growing range of needs in the specific market. This generally results in a less focused competence base and often requires acquisition of new competences. Examples of both of the developments highlighted above, appear in the historic development of Gist-brocades.

Of course, a mixture of these strategies is also possible, and the ultimate picture is that of the one-stop shop: Many markets and many competences. History in industry shows that this latter strategy is difficult and often proves unsuccessful in the end.

The bifurcation diagram depicted in Fig. 15.4 closes this discussion. The start of the diagram shows the start up of the company and it enters a phase of gradual development unto point A in the diagram.

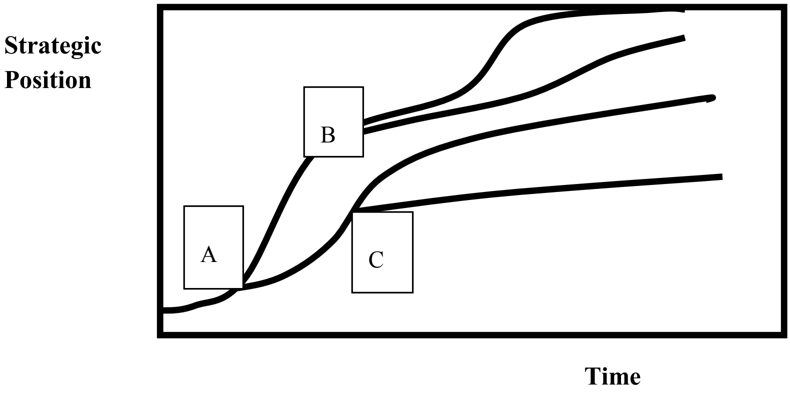

Fig. 15.4. A bifurcation diagram of the development of a corporation.

At point A, we need a critical decision regarding corporate development. It reflects a critical choice of strategic direction. An example of such a point is the entrance of Gist-brocades in the penicillin market in the last stages of World War II. After this decision, the company locks into either the branch AB or the branch AC. Both branches lead to different critical strategic opportunities at point C or B. By repetitive strategic decision points, very different historical paths are open to the development of the same primary initiative. The difficulty in managing such a process is that, at the time of the strategic decisions, the future opportunities that may result are unknown, whilst the possible paths of further development critically depend on the option chosen. A reconstruction of the historical path of the evolution of the company is only possible ex-post. Only looking back allows identifying the logic behind the evolution of the company. In fact, the evolutionary picture is importantly shaped by bounded rationality decision-making and hence contains elements of chance.

Many examples of such developments appear in the history of companies. The history of Gist-brocades discussed above, features the entrance in penicillin production at the end of World War II. Based on this move, it attempts an entry in the by then growing market for ethical pharmaceuticals when it merges with Brocades in the late 1960s to form Gist-brocades. This entry did not last as Gist-brocades sold it ethical pharmaceuticals business in the late eighties. The company became a producer of food ingredients, enzymes and intermediates and active ingredients for penicillins and cephalosporins. The nowadays leading ethical pharmaceuticals company Pfizer made the same decision in the said world war. It entered production of penicillin based on fermentation expertise that emerged from its activities in the production of citric acid by fermentation. It spins of the production of citric acid, discontinues the production of penicillins towards the end of the 20th century and becomes one of the leading pharmaceutical houses with a successful track record in the development of new ethical drug ever since the 1960s. This example shows that companies with a similar competence base in the past can show historical developments that result in very different companies with a presence in very different market and drastically different competences and information sets later on.

15.3. The evolution of some leading multinationals.

In this section, we analyze the historical development of some of today’s large industrial corporations, just as we performed such an analysis for Gist-brocades in Section 15.2. This provides an insight in the general features of industry evolution in selected cases. We select five companies. Two pharmaceutical companies Pfizer and Roche, a company in the energy and chemicals sector, Royal Dutch Shell, a producer of consumer products in the personal care and food sector, Unilever, and a producer of life science products and advanced materials, DSM. We acknowledge the information on which we base the analyses. We derive it from the accounts of the historical development of the respective companies on their official websites.

Royal Dutch Shell

The Royal Dutch Shell Group is one of the world’s largest businesses. Its complex corporate organization consists of more than 2,000 companies worldwide. Collectively, the group operates in oil and gas exploration, production, refining, transportation, and marketing. It has large interests in chemicals, being the world’s ninth largest chemicals business in the late 1980s. In addition, it is active in e.g. coal mining, forestry and solar energy.

The Royal Dutch Shell Group results in 1907 from a merger of Royal Dutch and Shell Transport. Royal Dutch originates in 1890 after receiving a concession to drill for oil in Sumatra, in the Dutch East Indies. The promoters of this venture find oil in 1885 and in 1892, the venture exports its first oil. The predominant use for oil in the late 19th century is as paraffin or kerosene, used for lighting and heating. However, Sumatra's oil is particularly rich in gasoline, the product used by the internal combustion engine. This is a favorable property due to the growth in demand for fuel for the motor car. We see that a shift in the resource base of the industry and a new need in the market provide the economic value force that allows the new venture to grow.

"Shell" Transport and Trading Company's origins lay in the activities of a London merchant. He starts his career in the 1830s, selling goods made from shells from the East. The business gradually expands the number of products in which it trades, including oil. Shell Transport grows rapidly. By 1900, the company possesses oil fields and a refinery in Borneo and a fleet of oil tankers. By the early 1900s, Shell Transport makes a series of costly mistakes. This finally results in the merger with Royal Dutch that creates the Group.

The Group expands rapidly. The total assets of Royal Dutch and Shell Transport grow more than two and a half times between 1907 and 1914. It acquires major production interests in Russia in 1910 and in Venezuela in 1913. The group also moves into the US, the homeland of Standard Oil. By 1915, the group produces nearly six million barrels of crude oil a year in the US.

World War I brings mixed fortunes for the group but the 1920s are a decade of growth. In 1919, Shell purchases the large Mexican oil fields controlled by the U.K. oil company Mexican Eagle. Venezuelan oil production expands rapidly, much of it controlled by Shell. In 1922, the Shell Union Oil Corporation forms in the United States to consolidate Shell interests there with those of the Union Oil Company of Delaware. The American business grows rapidly. This decade also sees the first steps in diversification. In 1929, a new company, N.V. Mekog, forms in the Netherlands to produce fertilizers from coke-oven gases. This is the group's first venture into chemicals. In the same year, the Shell Chemical Company in the United States arises to produce fertilizers from natural gas. The Depression years bring problems. From the late 1920s, there is a chronic problem of overcapacity in the oil industry.

During World War II and the invasion of the Netherlands, the head office of the Dutch companies moves to the Dutch West Indies. Shell plays a major role in the Allied war effort. The refineries in the United States produce large quantities of high-octane aviation fuel, while the Shell Chemical Company manufactures butadiene for synthetic rubber

The 1950s and 1960s are golden years of growth for oil companies, as demand for petroleum products expands. The Shell group and the other "seven sisters" of leading international oil companies, Shell, British Petroleum, Exxon, Texaco, Chevron, Mobil, and Gulf, retain a strong hold over petroleum production and marketing. The Shell group supplies nearly one-seventh of the world's oil products in these decades.

During the 1950s and 1960s, Shell diversifies into natural gas and offshore oil production and further expands its chemicals operations. This includes moving forward in the value chain in agrochemicals. This latter move proves a not lasting addition. The company divests its agrochemicals activities in the early nineties. In 1959, a joint Shell/Esso venture discovers natural gas in the Netherlands in Groningen. This turns out to be one of the world's largest natural gas fields, and by the early 1970s, it provides about half of the natural gas consumed in Europe. Shell discovers oil in the North Sea and it finds major reserves of offshore gas on the Australian northwest shelf. By the end of the 1960s, the Shell group also manufactures hundreds of chemicals in locations all over the world.

The structure of the world oil industry alters during the 1973 oil crisis when the Organization of Petroleum Exporting Countries (OPEC) unilaterally raises crude oil prices. The oil companies find themselves forced to allocate scarce oil supplies during the crisis, causing problems with several governments. The Shell group, like all the Western oil companies, suffers nationalization of much of its crude oil production in developing countries. The search for oil in non-OPEC areas increases and is successful. In the late 1980s, it is responsible for producing 5% of the world's oil and 7% of its gas.

In response to the problems of the oil industry, the Shell group diversifies its business in the 1970s, acquiring coal and metal interests. In 1970, the company acquires the Billiton mining and metals business in the United Kingdom, an activity that it divests again in the late nineties. In addition, the company creates its Renewables activities in 1977 to explore alternative sources of energy and chemicals such as biomass, wind, solar and hydrogen based technologies. Chemical manufacturing expands strongly. This expansion proves unfortunate, since world economic growth after 1973 slumps, with the major recession of the late 1970s and early 1980s causing acute problems. As a result, severe overcapacity develops in the chemical industry. The U.K. company Shell Chemicals experiences problems and reduces capacity. Similar overcapacity occurs in the oil refinery business throughout most of the 1980s. In the 1980s, the group rationalizes its exposure to chemicals and other noncore businesses. However, the group remains the world's largest producer of petrochemicals and a leading supplier of agrochemicals, in particular insecticides, herbicides, and animal health products. As said, the company divests its agrochemicals business in the nineties.

In 1990, Shell is an enormous business enterprise. One of the immediate challenges for the group concerns increasingly stringent environmental requirements that affect the chemicals and the petroleum industries.

The story of the evolution of Shell highlights a change of the resource base of the industry reflecting the importance of finding sources of free energy or economic value to fuel economic growth. Oil and motorcar fuel are an important factor in the Second industrial revolution, just as coal and the heat engine in the first industrial revolution. Oil fuels transport when the popularity of the motor car, powered by the combustion engine, grows. In addition, the increasing use of electricity is a factor. Furthermore, chemical know how facilitates takeover of the chemical industry by oil and natural gas at the expense of coal tar based technology. During its life time the group diversifies its resource base and moves outside energy into metals. However, to date largely only its energy and basic chemicals activities remain. The company moves downstream in chemicals unto agrochemicals and fine chemicals. These activities largely disappear by divestment. The presence of Shell in chemicals largely restricts itself to base chemical and the most important share of its revenue derives from being a supplier in the energy resources market. Its presence in the value chain of the chemical industry limits itself to the early stages.

In the consumer market, its most significant presence is in fuels.

Shell's strategy comprises both an upstream and a downstream part. Upstream involves the search for and the development of new oil and gas reservoirs. The downstream focuses on refinery, chemistry and distribution activities. Investments in e.g. new gasification technologies aim at the discovery of new methods to use resources efficiently, including the renewable ones.

Pfizer Inc.

Pfizer Inc. is one of the leading healthcare companies in the world. Pfizer exists of four groups: Pfizer Pharmaceuticals Group, Warner-Lambert Consumer Group, Pfizer Animal Health Group, and Pfizer Global Research and Development. Among Pfizer's products with annual revenues exceeding $1 billion, are drugs for treatment of hypertension and angina; a cholesterol reducer; an antidepressant; an oral antibiotic; an antifungal product and a drug for the treatment for erectile dysfunction. Warner-Lambert markets a number of leading consumer brands, including over-the-counter healthcare mainstays. Pfizer Animal Health is a world leader in medicines for pets and livestock. On the development side, Pfizer Global R&D spends almost $ 5 billion a year to develop new products. Its product pipeline contains more than 130 possible new products. The multitude of alliances that Pfizer fosters with academia and industries, significantly contributes to the efficacy of its R&D efforts.

In 1849 Charles Pfizer, a chemist, and Charles Erhart, a confectioner, start a partnership to manufacture bulk chemicals. Products are iodine and boric and tartaric acids. The company starts operating in the citric acid market for the growing beverage applications based on sourcing the product from agricultural resources but changes over to more reliable sourcing from fermentation-based production in 1919. Fermentation becomes a competence of the company, a part of its genome. As we will see, this later proves important to the company's further evolution. The company enters the production of vitamins, mainly based on chemical synthesis. By 1940, Pfizer produces a wide range of industrial and pharmacological products. Those latter products show particularly relevant in its further evolution.

While Pfizer technicians practice excellent fermentation technology, Alexander Fleming makes his historic discovery of penicillin in 1928 (see also Section 15.2). Recognizing penicillin's potential to revolutionize treatment of infections, major causes of early death, scientists research the production of large quantities at high quality for years. During World War II, commercial production becomes possible due to the efforts of the Allies. Because of its expertise in fermentation, the government approaches Pfizer. Soon afterwards, Pfizer starts research in both surface cultures in flasks and later on in submerged fermentation in large tanks. Still later, the company announces its entrance into large-scale production of penicillin in the by then largest production facility for the antibiotic. In Section 15.2, we identified such move as a competence driven expansion of the company, in this case based on competences in large-scale fermentation. The company adapts it information set to apply to production of a product beyond citric acid, the product leading to its competence in large-scale fermentation. By 1942, Pfizer produces the first penicillin vials for the medical departments of the Army and Navy.

After the world War, Pfizer enters mass production of penicillin to make it available for the consumer market. This marks Pfizer's first large entrance into the manufacturing of pharmaceuticals. A few years later, Pfizer produces enough penicillin to supply 85 percent of the US market and 50 percent of the world market. In 1946, sales reach $43 million. Competition of many others, e.g. Gist-brocades that also enters in that period (Section 15.2), results in price reduction and in a strong growth of the market and mass use. In fact, the market and the production technology co-evolve in the usual evolution pattern. Pfizer's bulk chemicals business decreases as former customers enter and the industry structure changes accordingly. Pfizer's instrumental role in developing antibiotics proves beneficial to society but a poor business venture in its own right, although it lays the foundation for the growth of the antibiotics and the pharmaceutical market. The dwindling prices of penicillin create a significant crisis and herald a bifurcation point in the company's history.

In 1950, the company introduces a new antibiotic, oxytetracyclin. In the postwar years, pharmaceutical companies search for new broad-spectrum antibiotics useful in the treatment of a wide number of infections. Penicillin and streptomycin help to expand the frontier of medical knowledge but not always offer refuge before the advent of semi-synthetic penicillins. Pfizer's breakthrough discovery of oxytetracyclin, a broad-spectrum antibiotic that soon proves effective against a wide range of infections, comes to the rescue.

With oxytetracyclin, the company changes from solely producing the active ingredient for the drug and moves ahead in the industry value chain to become a full-fledged pharmaceutical company supplying hospitals and retailers. This significantly changes the company's competences, its information set starts evolving in a new direction. It acquires marketing and distribution capabilities with a strongly growing sales force. This is a critical bifurcation point that shapes the company's future. After 12 months on the market, oxytetracyclin sales account for one-fourth of Pfizer's total $60 million in sales. Aside from modernizing marketing campaigns, Pfizer successfully develops a diversified line of pharmaceuticals. Thus, the company reduces its dependence on sales of antibiotics by introducing a variety of other pharmaceuticals.

At the same time as Pfizer's domestic sales increase strongly, the company improves its presence in the foreign market. Pfizer's foreign marketing expands into 100 countries and accounts for $175 million in sales by 1965. By 1980, Pfizer is one of the only two U.S. companies among the top ten pharmaceutical companies in Europe and the largest foreign healthcare and animal health product manufacturer in Asia.

Pfizer also diversifies by acquisition. Between 1961 and 1965, the company spends $130 million to acquire more than ten companies, including manufacturers of vitamins, antibiotics for animals, chemicals, and cosmetics.

The company increases its emphasis on research and development. With increased funds allocated for research in the laboratories, Pfizer joins the ranks of other pharmaceutical companies searching for the innovative and therefore profit making, drugs.

Although the company sales are $2 billion by 1977, Pfizer's overall growth slows down through the period of the late 1970s and early 1980s.This leads to a change in strategy. Firstly, funds for research and development increase to $190 million by 1981, an increase of 100% from 1977. Secondly, Pfizer begins a comprehensive licensing program to market drugs developed elsewhere. The result is an expansion of sales and profits both from licensed drugs and from the products that result from its increased R&D spending. By 1983, sales reach $3.5 billion.

By 1989, Pfizer operates businesses in more than 140 countries at net sales that are $5.7 billion, but net income declines. Research and development expenditure quadruples during the 1980s, and Pfizer plans to continue investing heavily in research and development.

Pfizer heads into the 1990s with numerous drugs in development, including preparations in the areas of anti-infective, cardiovascular, anti-inflammatory and central nervous system medications. Net sales in 1990 reach $6.4 billion. Research and development costs rise 20 percent, in line with Pfizer's determination to invest heavily in the development of new drugs.

Pfizer International launches many new products worldwide in 1990. The company's antifungal drug becomes the world's leading drug of its kind during this time. Sales of Pfizer's newest products account for 30 percent of all pharmaceutical sales, up from 13 percent in 1989.

Net sales in 1992 reach $7.2 billion and research and development expenses hit $863 million.

In the early 1990s, a wave of mergers and acquisitions characterizes the global pharmaceuticals industry. Pfizer instead builds its product pipeline organically rather than through acquisition. By 1995, the R&D budget is $1.3 billion. The one major acquisition that the company completes during this period is not in the area of human pharmaceuticals but in the animal health realm. In 1995, Pfizer spends $1.45 billion in buying the animal health unit of SmithKline Beecham, the largest acquisition in Pfizer's history. The purchase transforms the company's Animal Health Group into one of the largest providers of medicines for both livestock and pets.

Not content with its own rich product pipeline, Pfizer enters into a series of partnerships whereby it co-markets drugs developed by smaller pharmaceutical firms based on its powerful sales force. In 1997, Pfizer helps Warner-Lambert bring a cholesterol-lowering pill to the market. It soars to the top of the anticholesterol niche in the United States, achieving sales of $865 million in 1997 and $2.2 billion in 1998. Another 1997 introduction of a co-marketed drug developed by Japan's Eisai quickly becomes the leading drug prescribed to treat the symptoms of Alzheimer's disease.

Focusing further on pharmaceuticals, Pfizer sells its Medical Technology Group in 1998. In the same year, Pfizer introduces a leading drug for erectile dysfunction. It is by far the most famous of the new "lifestyle" drugs. It is an instant blockbuster: Sales for 1998 total $788 million and exceed $1 billion in 1999. Having participated in three record-setting launches in the late 1990s, Pfizer becomes one of the fastest growing drug companies in the world in revenue terms. Sales increase from $11.31 billion in 1996 to $16.2 billion in 1999. The overall growth for the 1990s is even more impressive as revenues increase with almost 300 % from the 1989 total of $4.2 billion. During this same period, Pfizer increases its R&D budget six fold, reaching nearly $2.8 billion by 1999, or more than 17 percent of sales.

In early February 2000, Pfizer and Warner-Lambert reach agreement on merging. The merger completes in June 2000. The combined company retains the Pfizer Inc. name, with the consumer product lines of the two firms combined within a unit called Warner-Lambert Consumer Group. The new Pfizer has 1999 revenues of $27.3 billion, making it the world's number two drug company. The company has the industry's largest R&D budget, totaling $4.7 billion in 2000. In terms of prescription drugs, Pfizer now markets eight products with annual sales in excess of $1 billion. The Warner-Lambert Consumer Group is the leading marketer of OTC (Over The Counter) healthcare, confectionery and shaving brands, with 1999 revenues of $5.5 billion. The Pfizer Animal Health Group is the world leader in medicines for pets and livestock, with 1999 sales of $1.3 billion.

In the period since its founding the company succeeds in evolving from a company producing active ingredients for food and pharmaceuticals into a full-fledged creative pharmaceutical company. Its information set changes drastically and very little of the old heritage remains as far as its "genome" is concerned. Over the years, the company divests fermentative bulk production of penicillin and citric acid. Those products and production technologies are a vital element in the development of the company until the last decades of the 20th century. This part of its information set disappears in its history whilst being crucial in its early development. We also recognize this feature in what we learn from the general theory of evolution.

Roche.

In 1894, banker Fritz Hoffmann-La Roche develops the idea of internationally marketing medicinal products with standardized dosages and effects under uniform brand names. The manufacturing basis for the idea is the extraction of the active ingredients from natural raw materials. As early as 1897, Hoffmann establishes subsidiaries in Italy and Germany in pursuit of his aims. The first success appears in 1898 in the form a cough syrup based on the company's antitubercular agent. It rapidly becomes a success, also due to the marketing strategy of the company. By 1912, Roche has branches in nine countries on three continents. In 1904, Roche introduces a heart tonic, and in 1909 a painkiller. Both are very important medical innovations at the time. The company's growth slumps during World War I as both France and Germany blacklist Roche for its international connections. The Russian Revolution in 1917 proves an even greater blow, as Russia developed into Roche's biggest market in the preceding years.

Developments in chemistry bring Roche its first synthetic product, a sleeping drug. This marks a change in strategy for the development of pharmaceuticals involving the search for synthetic rather than natural products. This happens earlier in the market for dyes. We recognize a change in availability of resources and technologies that heralds a bifurcation point in the evolution of the market for human pharmaceuticals. In the 1980s, the majority of the pharmaceutical products is of a synthetic or semi-synthetic rather than natural origin.

In 1933, Roche acquires the Reichstein process for synthesizing vitamin C and begins to scale it up for mass production. The process is also an example of biotechnology "avant la lettre", just as the process of Pfizer for the production of citric acid. This openness to biotechnology rather than completely relying on synthetic chemistry reflects part of the company's genome. Some emerging pharmaceutical companies evolve out of the markets for synthetic dyes and rely completely on what we now term classical chemical synthesis. Roche has a different history with an emphasis on the medicinal part of the development of a drug rather than the way in which its production takes place.

World War II is a test of strength for an international enterprise like Roche. Although sharply rising demand, particularly for vitamins, provides a significant boost in sales, war taxes in various countries and difficulties in conducting international transport and completing financial transactions, leave the company with few opportunities to earn a profit and develop.

In 1945, soon after the end of the war, a new application emerges for vitamin B derivatives in the field of hair care. Roche establishes a new subsidiary to develop this new line of business. After the War, Roche introduces a number of successful new products.

In the sixties, a tranquilizer and Roche's first anticancer chemotherapy drug appear on the market. During this exceptional growth period for the relatively young pharmaceutical industry, Roche's introduction of its tranquilizer in 1963 is a milestone whose impact on medicinal practice is still manifest. In addition to infectious diseases, Roche's research department targets mental illnesses and central nervous system drugs. The same year, Roche acquires a fragrances and flavors producer and in 1964 a luxury perfume producer to complement its own flavors and fragrances operations.

The rapid technological progress of the post-war period leads to the conviction that expansion of medicine into areas beyond the simple treatment of disease with chemically synthesized substances is a real possibility. Moreover, the illnesses of civilization, such as overweight, become urgent medical problems.

Roche launches a broad based diversification program. It acquires an electronics research unit. In 1968, Roche purchases a diagnostics unit and founds the Basel Institute for Immunology. The company enters the manufacturing of devices in 1969. The newly acquired or established companies form the core of Roche's future Diagnostics Division.

In research, the group starts its biotechnology activities in the early 1970s. This results in isolating interferon, leading to Roche's first genetic engineering based cancer drug in 1986. This is an early example of success based on the integration of a new competence in the company's information set.

In the 1970s, the company develops a process for producing monoclonal antibodies, a key area for Roche today. This is again an example of integration of a new field of technology.

Roche launches the first significantly successful drug for Parkinson's disease in 1974. In the same year, a successful mild tranquillizer follows. In 1978, the company starts collaborating with the newly founded biotechnology company Genentech Inc. Later on, it acquires a majority stake in said company. This is an example of complementing Roche's own genome with that of a creative emerging company and shows that often collaboration beats captive development if a company needs a totally new information set, i.e. in the newly growing area of biotechnology.

To improve a difficult market situation, Roche initiates a broad-based consolidation program within the company. It divests numerous activities less related to the company's core. This is a feature of industry evolution that we repeatedly observe since the late 1970s.

In 1982, the company introduces a series of significant new products, including an anesthetic, an antibiotic, acne medication and the first effective psoriasis treatment. The new products bring a significant improvement in Roche's business outlook. The company introduces the first product based on monoclonal antibodies, a cancer test, the same year.

In 1994, Roche acquires the American Syntex Group, a major company with highly regarded research. In 1995, Roche introduces an inhibitor against AIDS, followed by a monoclonal antibodies based cancer drug.

Further acquisitions, including such companies as Boehringer Mannheim, make Roche the global market leader in diagnostics and provides an entrance in the diabetes business.

Roche introduces an antiviral drug against the common cold in 1999. This drug shows to have broader applications in fighting viral infections. In addition, it introduces treatments for obesity and breast cancer.

As part of Roche's ongoing strategy of focusing on the core, it divests the fragrances and flavors division, acquired earlier, in 2000. In addition, the sale of the Vitamins and Fine Chemicals Division to DSM in 2002 is an example of this strategy of focus.

The year 2002 is one of significant expansion by both introductions of newly developed products and an acquisition to strengthen the position in Japan.

Finally, Roche's sells its founder product line of over-the-counter drugs to Bayer AG in 2004.

Consolidated sales for 2007 are CHF 46 billion, generated by approximately 78,000 employees.

The history Roche shows interesting features. Roche emerges as a company developing and selling what we nowadays call Over The Counter, OTC, drugs. Interestingly, the remains of this old company genome disappear in recent divestitures. Other than some of its present competitors in the pharmaceutical industry, it does not originate from the advent of industrial chemistry in the synthesis of synthetic dyes from coal tar. It has a medical focus ever since its emergence. The history of the company shows the impact of creative marketing in its early evolution. In addition, it adopts new competences, such as biotechnology, in an early stage of their development, organically and by cooperations and acquisitions. The company shows a strong dedication to the pharmaceutical market although it diversified by acquisition and organic growth in the second half of the 20th century. In later years, it spins of the non-core activities including one of its founder businesses in the vitamins and fine chemicals area. It nowadays wholly focuses on the health sector.

Unilever.

In the late 19th Century, at Oss in Brabant, the Netherlands, Jurgens and Van den Bergh, two family businesses of butter merchants, have thriving export trades to the UK. In the early 1870s, they become interested in a new product made from beef fat and milk, margarine, that allows mass-production as an affordable substitute for butter. Later on, this margarine changes its raw materials base and becomes of a mainly vegetable origin.

Over in the north of England in the mid 1880s, a successful wholesale family grocery business run by William Lever, starts producing a new type of household soap. The product contains vegetable oil. It lathers more easily than traditional soaps made of animal fats. Again an example of a change in resource base. Unusual for the time, Lever gives the soap a brand name and sells it wrapped in distinctive packs. This is an early example of the marketing approach so characteristic for companies, like present-day Unilever, that operate in the branded consumer products area.

Jurgens and Van den Bergh open their first factories to produce margarine in the Netherlands in 1872 and in 1888, they expand into Germany. By 1898, Van den Bergh already has a 750 people sales force and launches a new branded margarine.

Lever & Co starts production of soap and by the end of 1887 produces 450 tons a week. Mid 1890s, it sells nearly 40,000 tons of soap a year in the UK and starts expanding into Europe, America and the British colonies with factories, export businesses and plantations. At the end of the nineteenth century, the company introduces a new type of product, soap flakes. It makes housework easier than using the traditional hard soap bars. In 1904 in the UK, Lever Brothers launch another convenience product for housework, one of the first scouring powders. In 1906, Lever Brothers has a thriving export trade and factories in three European countries as well as in Canada, Australia and the US. It also starts enterprises in the Pacific. In 1909, Lever Brothers develops a palm plantation in the Solomon Islands and at the same time Jurgens and Van den Bergh set up a joint palm plantation venture in German Africa. Both companies pursue this as a strategy to secure access to raw materials in a period of increasing demand for these resources, after agricultural oils become the resource of choice for their disparate products in oils and spreads and cleaning and personal care products, respectively.

Lever Brothers follows a characteristic of the Second Industrial revolution and founds its first research laboratory in 1911.

In the 1920s, Lever Brothers buys a popular sausage company that produces ice cream to sell in the summer when demand for sausages falls. This marks a broader entry in the food sector. Jurgens and Van den Bergh create the Margarine Unie, Dutch for Margarine Union. The union quickly gains new members, creating a large group of European businesses involved in the production of almost all foods created from oils and fats. A Lever Brothers company launches the first vitamin-enriched margarine. The Margarine Unie acquires the French-Dutch Calvé group with factories in the Netherlands, France, Belgium and Czechoslovakia.

At the end of the 1920s, in 1929, Lever Brothers and Margarine Unie sign an agreement to create Unilever. This ultimately results in an amalgamation of all activities. In 1930, Unilever is officially established. Apparently, the similar resource base and the associated processing know how, form the incentive to combine the companies' genomes. In addition, they both have competences in consumer marketing.

Mid 1930s, soap production moves further from hard soaps to flakes and powders designed to make household cleaning easier. This leads to further expansion in the soap market. In the same period, a margarine follows that mimics the properties of butter further by addition of Vitamins A & D. A campaign to improve public perceptions of margarine results in strongly increasing sales.

In the early 1940s, Unilever started its research in food and nutrition. In addition, Unilever becomes the majority shareholder in a company owning the UK rights to a method of food preservation new to mass markets, deep-freezing. Years later, deep-freezing enjoys a surge of popularity when it shows the best way of preserving food. Around the same time, Unilever acquires activities in freeze-dried vegetables and canned goods. In 1946, Unilever launches the first frozen peas in the UK.

Unilever introduces a branded shampoo in the UK in 1954. This becomes its leading shampoo brand. By 1959, it is available in 18 countries worldwide. It introduces a branded soap in US.

In 1956, Unilever Research establishes its Biology Department, which in the 1980s becomes the BioScience, Nutrition and Safety unit.

In the late 50s and the early 60s, Unilever launches its first margarine in a tub, replacing the traditional block wrapped in greaseproof paper. In addition, the company launches a pioneering "health" margarine after the medical community asks Unilever to develop a cholesterol lowering food product. Initially, it is only available from pharmacies, later on through the food retail chain. In 1963, the first packaged and branded ice cream cone launches in Europe.

The company acquires Lipton International in the early seventies and Unilever's tea business becomes one of the largest in the world. In addition, Unilever further develops its businesses by a number of new product introductions and acquisitions.

It increases its presence in the US by acquiring National Starch at the end of the 1970s. It is an ingredients company rather than a consumer food company. It produces adhesives, starch derivatives and specialty organic chemicals. This acquisition is the largest by a European company in the US at this time.

In the eighties, Unilever announces its Core Business Strategy and large acquisitions and disposals follow over next decade. The acquisition of Naarden doubles Unilever's business in fragrances and food flavors.

In the nineties, Unilever grows further and expands its product range organically and through acquisitions. In addition, regional expansion into Eastern Europe and Asia takes place. Unilever also realizes a new landmark in nutritional products. The company decides to eliminate practically all trans-fats from food production in a rapid response to new research suggesting that their effect on blood cholesterol is at least as adverse as that of saturated fats.

Towards the end of the nineties, the company sells the ingredient businesses it acquired earlier and further focuses to become a company operating in consumer products rather than ingredients for use by the food industry.

In the beginning of the 21st century, Unilever acquires Bestfoods in the second-largest cash acquisition in its history. It further retreats to core activities by reducing its brands from 1,600 to 900. Further divestments contribute to this process and this further continues in later years.

Unilever emerges from two companies operating in branded consumer products in the food and the detergent/personal care sectors respectively. These two types of activities still reflect the profile of Unilever, albeit that its position in the value chain fluctuates from being totally resource and intermediates integrated, i.e. spanning the whole value chain, to a focus on the consumer end only. The merger of the two companies probably derives from synergies in raw material and conversion technologies as well as a shared perspective on the importance of branded products and consumer marketing. The company expands organically based on its sizeable consumer marketing and research activities and by acquisition. It expands its coverage of the growing needs of the consumer and the associated diversification of the number of product categories the consumer requires. The increasing per capita GDP due to the industrial revolutions is one of the factors fueling this growing demand.

Over the years, the company increases its emphasis on the consumer end of the value chain and marketing of branded products. It first divests it presence in primary agricultural resources in plantations and later its activities in ingredients.

Many overriding principles of industry evolution in the 20th century appear.

DSM

In 1902, the Dutch State Mines emerges as a stated owned mining company. It produces its first coal in 1906. A head office in Heerlen opens in the same year. Between 1910 and 1920, the number of mines increases and a new coke plant marks the start of the chemistry activities of DSM. The mnemonic DSM still reflects its origin as Dutch State Mines. In 1927, the chemistry activities become more prominent by the construction of a fertilizer plant in Geleen. Recently, the company divests its remaining fertilizer activities in the process of restructuring the company. In 1920, the coke plant delivers the fist coke gas to the local community. In 1928, the company creates its own gas distribution firm.

In 1939, DSM follows the general pattern on Industrial Research emerging during the second Industrial Revolution by creating the Central Laboratory in Geleen. The innovative technologies the company develops based on its research, lead to patents and a licensing subsidiary at the end of the 1940s.

In the 1940s, the closure of one of its coke plants marks the retrenching from coke production. DSM's diversification into chemistry becomes manifest in the 50s when its starts production of caprolactam, urea and polyethylene. To secure its base chemicals supply, the company invests in a naphtha cracker going into operation in 1963. The diversification into chemicals continues with the construction of a melamine plant.

The seventies herald completion of the abandonment of the company's founder business when it closes one of the last mines. The last mine closes in 1973.

In the eighties, the chemistry activities move from base products into fine chemical and specialty chemicals. It moves downwards in the value chain by the foundation of activities in resins, acquiring a Venlo, the Netherlands, based fine-chemicals company and an acquisition in synthetic rubber. In addition, it ceases to be State owned and becomes a publicly listed company.

In the nineties, the expansion in performance materials enhances and further acquisitions in fine chemicals follow. This firms up its profile as an industrial chemicals, performance materials and fine chemicals company.

In 1998, the acquisition of the Life Sciences company Gist-brocades (Section 15.2) materializes and the Life Science Products activities of DSM result from the integration of the activities of Gist-brocades and the fine chemicals activities of DSM. This sets the scene for reshaping the company in the first decade of the 21st century.

In the 21th century, DSM spins of its activities in base chemicals and commodity polymers and becomes a company focused on advanced materials and life science products. A further landmark is the acquisition of the Vitamins and Fine Chemicals division of Roche that profiles the nutrition activities in Life Science Products. The divestiture of the founder business of Gist-brocades in baker's yeast materializes and the process of divesting industrial chemical nears completion.

The company drastically changes its business profile in the last odd twenty years from mainly active in industrial chemical to a specialties profile in advanced materials and life science products. This also involves a change in competences base from process technology oriented to a range of competences, such as biotechnology, nutrition sciences and materials sciences. An important driver of its change process is organic growth based on its new competence profile. In addition, growth results from a large number of sizeable acquisitions. Acquisitions are also an important factor in realizing the diligent change in its competence profile.

15.4. The paradox of strategic planning.

In his book, "The Blind Watchmaker" Dawkins (1986) analyzes two approaches towards the explanation of the existence of complex systems such as personal computers, airplanes, motor cars and organisms. Here we include organizations like firms and universities in this array of complex systems. The first approach is the hypothesis of planned design and we are probably inclined to consider the engineering efforts underlying personal computers, airplanes and motor cars as examples of complex systems that are a product of planned design and this is in my view an, although limitedly, defendable position. Also the design of such systems involves elements of learning by doing beyond rational design. Most people that accept the theory of evolution of Darwin (and Wallace) regarding the evolution of biological species understand that biological complexity arose by a process that does not involve planned design but is a product of random changes and directed evolution by environmental selection. The apparent goal orientation arises by selection of the most efficacious information sets, the DNA, out of the variation that random mutations provide. This also applies to our species *Homo sapiens* that is a product of this evolutionary process rather than a creation based on intelligent design. We have argued that this also to an extent applies to organizations like firms and universities. Of course, their development involves elements of design, however, the inherent complexity of the information sets that characterize these organizations, leads to bounded rationality. Hence, planned design only limitedly applies to such organizations. Variation of their information sets by limitedly deterministic change shapes these organizations by environmental selection of the most efficacious organizations. This has become an excepted feature of contemporary approaches to organizational economics where the evolutionary approach (Section 8.13) introduces this notion.

The theory we develop in this work analyzes the root causes of our inability to rely on planned design in complex systems such as economic institutions like firms and other parts of society such as universities, as caused by information that is lacking in our "model" of these organizations. It particularly highlights critical bifurcation events in which these informational limitations do not allow reliable predictions of the outcome of our (management) actions.

In analyzing the history of the evolution of life, the element of chance in the path evolution takes is very clear. Random mutations at the DNA level provide an impetus for the evolution of increasingly complex and adapted organism if combined with competition and environmental selection. It is also clear that exogenous and endogenous changes in the environment are crucial in the bifurcation events that shape the path of evolution. The evolution of life thus has an important historical dimension in which crucial events shape the paths that are available for evolution to proceed. At every instant in history, the paths available for evolution to proceed, do not follow from the state of the organisms and their environment at that moment in time.

In this chapter, we analyze a limited sample of firms and their evolution over time. Also for firms, their evolution only limitedly results from "intelligent design", critical bifurcation events, involving limitedly rational decision-making, seem to shape the evolution of industrial corporations. Different firms that at critical bifurcation points have similar opportunities to shape their subsequent development, may end up in very different markets using very different information sets (competences) later in their evolution. The rationale of the decision-making at the critical strategic events in the company's development can mostly only be derived ex post.

This author has been working in strategic planning in industry for some twenty years. Strategic planning involves designing the future of the company starting from an analysis of the present position of the company and its environment and the formulation of actions to shape its future evolution in a desired direction, i.e. to realize attractive strategic goals. We discussed an approach to strategic planning in Section 8.15. The discussion presented above may cast some doubts on the rationale behind strategic planning in view of the inherent limitations to the prediction of the future in complex systems. Are companies accessible to strategic design or is a paradox inherent in exercises in strategic planning? Of course, many if not all companies do engage in strategic planning of some sort and devote scarce resources to this activity. From an evolutionary perspective, this could be construed as evidence that strategic planning adds to the survival value of companies and hence must result in competitive advantage. I personally believe that this is the case but that the considerations above have important consequence for the practice of strategic planning.

The first aspect that needs explicit consideration in the process of planning is the element of uncertainty and unpredictability. Both our internal analysis, defining the companies present genotype and our ability to shape it in a future perspective, and the external analysis, the analysis of the competition and the other aspects of the environment, e.g. as defined by the Porter analysis (Section 8.14), should reflect the inherent limitations of our information. Both the internal and external analysis should reflect the main areas of uncertainty and risk. Furthermore, we need to specify the assumptions on which our assessment of the likely trends in the future evolution of the environment and of the company rests. These provide a checklist if unexpected developments in the environment or internally occur. This may lead to the necessity to reconsider the strategic objectives of the company.

Another consideration in the design of the planning process, concerns the way in which we incorporate the co-evolution aspects of the development of complex systems and their environment as we highlight these in this work. This means that we cannot analyze the internal perspective in isolation from the external perspective, we should explicitly take into account the impact of our actions internally on the changes in the environment that result from these actions. Taking a static perspective that considers our firm and the environment as separately evolving, is inadequate.

Furthermore, we need to be careful if our analysis reveals that we need drastic changes in the information set characterizing our capabilities and competences. Here we risk reaching the limits of the changes allowed by the copying fidelity limit characteristic for complex evolving systems (Section 7.3). It may very well be that organic development does not allow the planned changes and that we need to consider cooperations, joint ventures or mergers and acquisitions to achieve our goals. Alternatively, our strategic goals may be unrealistic and a review of these goals may be necessary.

A no doubt important part of a proper strategic planning process concerns challenging existing beliefs and practices that underlie a company's information set. If we do not challenge in this respect, the creative function of mutation of information sets that is one of the vital drivers of evolution, erodes. The balance between conservation and mutation should be carefully considered. If we completely avoid the risk and uncertainty inherent in changing the company's genome, we risk being no longer able to react to changes in the competitive landscape or other aspects of the environment.

Another pitfall rests in assuming change in the competence base and the market position of our company in the period that defines the planning horizon, whilst neglecting the fact that also competitors and the other aspects of the environment change in that period. Strategy is a play on an evolving stage.

A further remark concerns the method used in the strategic planning exercise. Of course, it helps if we use a sound and well thought of method. However, we insist that the creativity of the stakeholders engaged in the exercise and the quality of their interaction in the planning process is far more important than mechanically working through the planning agenda the methodology prescribes.

15.5. Conclusion.

Obtaining generic conclusions from the observations on this limited sample of the "ecosystem" of present-day industry proves a difficult task. A number of factors contribute to this difficulty.

Firstly, there is the nature of the material on which the analysis rests. As said, we use the companies' perspective on their respective historical development as obtained from their own websites. Although the events that feature in the historic development are objective facts, the very selection the company makes and the present position of the company may affect interpretation of the importance of certain steps.

Secondly, the selection of companies reflects activities that are successful participants in the present-day landscape of the industry. Many of the emerging industrial activities of the past that did not grow into prosperity or have not survived, are not present in the sample. No doubt, their histories provide important clues to the evolutionary features of the industry as a whole.

Thirdly, the selection features companies with which the author is to a varying extent familiar.

Notwithstanding these difficulties and limitations, we attempt an analysis of some overriding features of the evolutionary tree.

We clearly recognize the importance of a change in the resource base of society in the appearance and the growth of new industrial activities. An example is the shift to fossil resources, initially coal in the first industrial revolution and later on to oil and gas in the second industrial revolution. The multinational Shell emerges from such a development. The precursors of Unilever initially emerge from the introduction of new resources for a butter substitute and for cleaning products. In these two cases, we detect signs of a cyclical interaction as both the availability of the resources and the activities of the companies to make these resources available and useful in terms of products in demand in society, drive the evolution of the changing resource base. Industries and their resource base co-evolve.

A second evolutionary force results from the evolution of the science and technology base of industry. The appearance of modern chemistry revolutionizes the availability of new products to serve existing needs and makes it possible to serve new needs. The synthetic dyes are a case in point. Also the birth of the pharmaceutical industry traces back to new approaches to drug development (synthetic drugs). Also in this arena, we see signs of co-evolution. All the companies we studied engage in industrial research and interact with public research. In this way, the companies and the science and technology base of industry co-evolve.

What we also note is the importance of the interaction between public R&D in academia and institutes and captive industrial R&D. A case in point in this respect are the developments in biotechnology in the last three decades of the 20th century. A bifurcation point in the development of biotechnology is the advent of genetic engineering in the early 1970s. Academia pioneers the early development of the enabling technology but it becomes quickly integrated into an industry infrastructure through, in addition to captive in house organic entry, small start-up companies, and cooperations between established industries and these start-ups and/or the academic community. In addition, acquisition of startup companies by established players plays an important role. We see such efforts in almost all the companies in the sample we analyzed (e.g. Roche, DSM).

In addition, changes on the demand side result in forces that change the evolutionary pattern of the industry as a whole and in addition affect the fate of individual activities. Again, there is a cyclic interaction in which changes that originate on the demand side lead to developing new ways of satisfying that demand by industry. Alternatively, new technologies or new resources, made available by the initiatives in industry, lead to a growth of the demand in the market. We already stressed that the cyclic nature of the interaction between supply and demand defies a clear separation between cause and effect in such cases.

In the evolution of Shell, we see an example of such development. The combustion engine triggers demand for automotive fuels and Shell happens to avail of resources and technology to supply fuels for the combustion engine. This was instrumental in shaping the future of Shell and for that matter the combustion engine.

The emergence of the pharmaceutical industry, as witnessed by the companies Roche and Pfizer, provides another example of such a new demand. The marketing and technology efforts of emerging companies engaged in the development and the supply of such products satisfy this demand and trigger the growth of the pharmaceutical market.

An example of such a development is also clear in the history of the market for penicillin and the anti-invectives market in a broader sense. The success of penicillin derives from the scientific discovery of Penicillin by Fleming (1929). In society, there exists a clear need for anti-infectives, as infections are a major cause of death. Still Fleming discovery does not immediately lead to a booming market success; society does not avail of the technology to make large quantities available at affordable prices. At the end of the Second World War, this changes when the nature of the demand (military use) leads to industrial activities that change the supply side in such a way that larger quantities become available. This results from largely preexisting technology positions of established industrial players (e.g. Pfizer and Gist-brocades) and strongly soaring technology development efforts that lead to quickly decreasing prices and growing availability and demand in the years after the World War. This made anti-infective therapy an established part of medicinal practice. In addition, it heralded a period of strong growth of the pharmaceutical industry by the introduction of both a broader range of anti-infective products and new successful products in other therapeutic area. We again see a case of co-evolution of technology, industry infrastructure and market demand.

We also see such a development in the Unilever history, where the environment and the company were instrumental in the creation of the present food and personal care market and the presence in the market of present-day Unilever.

A factor that also needs highlighting is the strongly rising per capita GDP in the wake of the industrial revolution. This is both the result and the cause of the changing supply and demand pattern in the years after 1850. This is again an example of a co-evolution process. It is perhaps one of the most important co-evolution events in the history of humankind to date. It is no accident that all the companies in our sample find some of their roots in the second industrial revolution.

We also recognize another feature of the evolution of complex non-equilibrium systems in the companies in our sample. We refer in this respect to the importance of critical bifurcation points where strategic decisions shape the fate of the company and determine unknown future options for further development. These can drastically change the future appearance of companies even if they start out in markets or with competences that make the various options for further development possible at the moment of the strategic choice at the bifurcation point. This is exemplified by the cases of Gist-brocades and Pfizer, where a critical event at the end of the Second World War, leads to very different future evolutions of the companies under consideration. The above remarks also point to a limitation of the practice of strategic planning. Future opportunities depend on decisions made at a point in time when these future opportunities are still unknown. Still the choices made determine whether the new opportunities will materialize.

In the sample of companies we analyze, we also recognize another feature that derives from the intrinsic limitations to the prediction of the future evolution in complex systems. In some companies, we still recognize the nature of the original initiative from which the company emerges in the resource base, the competence base and/or the market presence. This applies to Shell and to a perhaps lesser extent Roche. In the case of Gist-brocades, only the competence base in fermentation survives. For DSM its original position in the coal mining area completely disappears. This also applies to Pfizer whose mainstay today is in the development and marketing of ethical pharmaceuticals.

Finally, our analysis highlights that judicious use of cooperations, mergers, acquisitions and divestitures shapes, in addition to organic growth, the evolutionary fate of the companies in our sample.

CHAPTER 16. TRACING THE SOURCE AND THE DIRECTION OF EVOLUTION: A SUMMARY.

16.1. Introduction.

In the foregoing, we analyze a wide range of evolutionary phenomena ranging from the very large stellar objects, to the very small, the molecular machinery underlying life phenomena. In this chapter, we present a round up and arrive at general features of evolution, both in material systems, e.g. organisms, and in immaterial systems, such as science and technology.
The first concept we reintroduce is the learning by doing cycle (Fig. 16.1) that is very basic to any type of evolution. We discuss this archetype of an evolutionary cycle in Chapter 6 in the context of the systems theory of evolution.

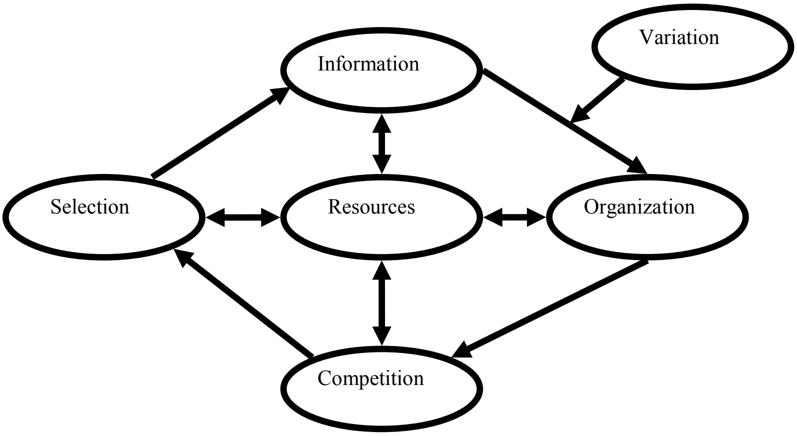

Fig. 16.1. Evolution through change by chance: The learning cycle.

Fig. 16.1 highlights the pivotal role of information. All the systems we studied develop, progress and communicate information, be it in a tangible form (e.g. an organism's DNA) or in an abstract or intangible form, such as in a scientific theory. The information translates in some kind of organization, e.g. the organism's phenotype, or the products and services and other aspects of the manifestation of an industrial enterprise. In this process of translation of that information set changes appear, either by more or less directed change, e.g. by R&D or the gravitational force in the universe, or by error. This variation is vital, as it is the source of progress. The phenotype competes for scarce resources such as free energy or, in our broader interpretation in terms of the EVT formalism, economic value. We recognize many examples of such resources and we reiterate these in a general way in the next section. This competition may also take place in the intangible world, e.g. where rivalling scientific theories compete for explanatory power in the light of experimental evidence. This competition for scarce resources leads to selection of the most competitive phenotype and hence its related, changed, information set. The process leads to the selection of information sets that are coding for structures that are more efficacious with respect to the selection the environment provides. This closes a positive feedback cycle that is the basis of sustained evolution. We extensively highlight that closing the cycle leads to a blurring of cause and effect.

In addition, a cause and effect cycle between the structures and their environment closes. This in principle holds for all the elements of the cycle. The resources, the variation, the information set, the phenotype, the competitive environment and the selection process relate causally and are both cause and effect of the evolution taking place.

The second aspect that we reintroduce is the concept of sustained evolution (Fig. 16.2 and Chapter 6).

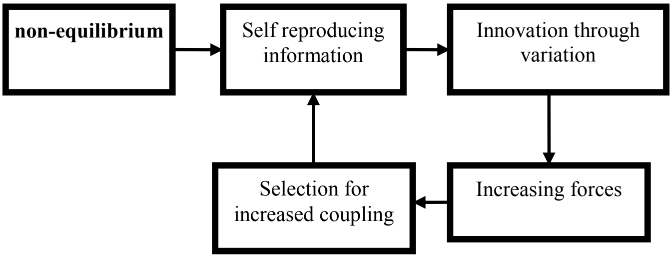

Fig. 16.2. Sustained evolution.

Firstly, there needs to be a source of non-equilibrium, i.e. a source of potential value. In the case of the evolution of the systems around us in the universe and on earth, the Big-Bang creates this source and this potential value remains the same in the estimated 13.5 billion years of evolution to date. In the process of evolution of the universe, its statistical entropy increases.

As the cosmic background radiation witnesses and more particularly the variations in the background radiation, at a crucial time in the early universe, differences exist in local energy densities, be it in the form of mass or radiation. This variation in density forms the initial self-organizing information. These differences in energy density amplify by gravitation and lead to increasing inhomogeneities in the local energy density in the evolving universe. It triggers the subsequent ignition of the stars, the initial source of free energy being the fusion of hydrogen to helium. One of the systems that form in this way is our solar system that also contains all of the higher elements, made by fusion reactions in the stars. This also induces the formation of the solar system and the planet earth. Life emerges on earth as a self-replicating information set. Early life forms, the early replicators, feed on the scarcely available forces derived from inorganic fuels and organic materials that form under the influence of solar radiation. DNA based self-replication appears and this later leads, amongst others, to the first photosynthetic bacteria that harness the free energy in solar radiation. This drastically increases, both in size and in number, the forces and fuels the development of heterotrophic organisms that feed on the organic material produced by the photosynthetic autotrophs. The food webs, as we know them today, appear. In the process of evolution on earth, the brain emerges. It increases in sophistication when the genus *Homo* starts to evolve and our species *Homo sapiens* appears. With the appearance of the brain, new replicators beyond DNA evolve and exogenic evolution emerges. Exogenic evolution results in the scientific revolution and sophisticated science and technology, the first and second industrial revolutions and our present-day socioeconomic system with its markets and firms.

The cycle of sustained evolution depicted in fig. 16.2 emerges clearly in this broad-brush description of the processes that finally lead to our present-day socioeconomic system.

In Chapter 6 where we introduce the systems theory of evolution, we highlight the conditions for the evolution of dissipative structures, such as organisms, firms and economies:

- The existence of non-equilibrium resulting in sources of potential value.
- The appearance of self-replicating structures that develop and communicate the information needed to transform these sources of potential value into economic value that fuels their maintenance and further development.
- A mechanism allowing the exploration of information space by the introduction of variation in the information sets of the self-replicating structures. This introduces the elements of both change and chance in evolutionary processes.
- The development of an increasing variety of information sets that both create and couple to an increasing variety of sources of economic value and compete for these ultimately scarce resources. This leads to the selection process that closes the cycle of evolution.
- The sources of economic value start increasing in diversity and in fact, the structures that develop become sources of economic value for other structures. In biological evolution, this leads to primary producers such as plants that directly use solar radiation, herbivores that feed on plants, predators that feed on herbivores and omnivores such as humans that feed on both. Our socioeconomic system is omnivoric in its character as our economies build on all resources that derive from solar radiation, including fossil resources.
- In addition, the variety of information systems increases. Systems that evolve based on new ways of storing progressing and communicating information complement the immortal coils of the DNA macromolecules. Exogenic software and hardware start complementing and dominating the DNA hardware.

The conditions we identified above typically leads to a pattern of evolution that shows a succession of structures that evolve following a succession of life cycles. This results in a pattern of evolution as shown in Fig. 16.3.

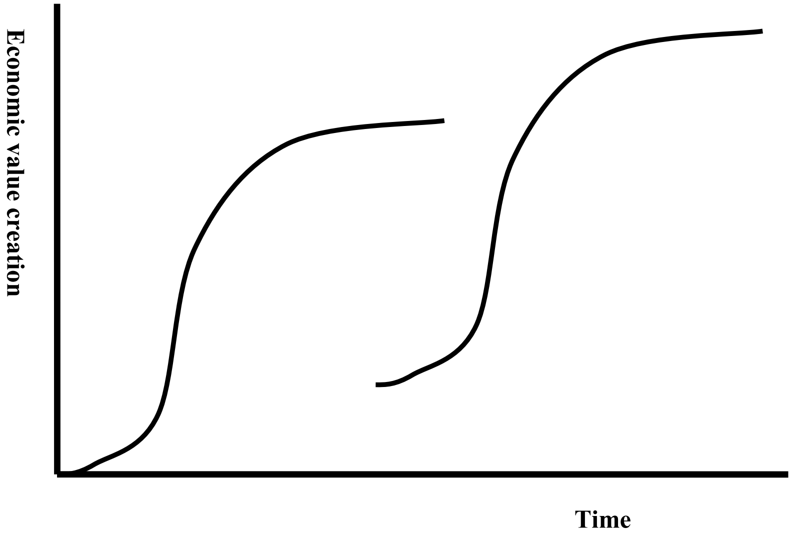

Fig. 16.3. Economic value creation in sustained evolution.

As Fig 16.3 shows, the evolution in response to the availability of a source of potential value results in a primary development of structures that transforms potential value in harvestable economic value. The primary information set that underlies these structures evolves through the familiar life cycle that characterizes dissipative structures (Section 6.9, Section7.5). After its emergence, it goes through a phase of strong growth, subsequently growth decreases in pace and the structure reaches maturity and decays. The decay phase is not shown in our figure. In the maturity stage, the harvesting of economic value reaches the maximum that the characteristics of the information set of the structure allows. In the growth phase there is ample room for perfection of the information set within the limits allowed by the its characteristics and there is only limited evolutionary pressure that allows information sets with drastically changed characteristics to challenge the evolving primary set. After a while when the information set reaches the limit of its potential to harvest economic value, such evolutionary pressure develops. This allows the emergence of a new initiative that starts at a lower level of harvesting economic value but has an information set that allows harvesting more economic value when it evolves into maturity. This may, as an example, be a larger information set, a phenomenon we observe in biological evolution. The phase in which a new initiative appears is often a phase of crisis in which predictability of the evolution of the system is lost. It is a critical bifurcation point in the terminology of the systems theory of evolution.

In the following, we analyze the elements of the eternal golden braid of evolution and try to indicate regularities that appear in the variety of evolving systems on earth and in the universe.

16.2. Non-equilibrium, free energy and economic value.

Non-equilibrium allows the development and creation of forces that become sources of economic value. The essence of these forces exists in gradients in economic value and/or differences in cost of information. In the systems we discuss, we identify a wide variety of sources of forces. In the early stages of the evolution of the universe, the electromagnetic force is responsible for the formation of neutral atoms. This changes the intensity of the interaction of matter and radiation to such an extent that radiation decouples from matter. The interaction between matter and radiation that remains is vital to the further evolution of life on earth as it allows photosynthesis. After the formation of the nucleons by the strong force, the electromagnetic force creates neutral atoms, particularly hydrogen and helium, the life supply of sources of matter for the universe. By the gradual decoupling of matter and radiation we progressively enter the matter dominated era and after the formation of neutral atoms gravitation increasingly becomes the force that leads to the formation of large concentration of mass that ignite to form galaxies and solar systems. This is initially the result of the transformation of hydrogen into helium and later on the synthesis of heavier elements. In a sequence of processes, this leads to the formation of our solar system and the earth some five billion years ago.

A vital element of the evolution of life and the socioeconomic system, with e.g. its firms and markets, is the development of self-replicating information. What exactly fuels the emergence of life is still subject to debate. Several competing theories exist for the explanation of this crucial step in evolution. The first life-like structures show some or all of the general features that drive evolution. The initial structures were autotrophic, i.e. did not depend on resources that were a product of already existing life forms. We do not delve further into these speculations and assume that RNA and ultimately DNA based organisms emerge on earth quite early in its evolution.

A very significant development needs to take place. Today, the overwhelmingly dominant source of non-equilibrium on earth derives from solar radiation. This radiation provides an abundant source of free energy and potential value that derives from nuclear fusion processes in the sun. In order to transform this potential value to useful work life has to develop an

effective coupling mechanism to benefit from this potential. This depends on inventing ways of interaction between radiation and matter in a way not destructive to matter itself. The interaction exists in the absorption of a photon by an electron causing it to jump between possible energy state in an atom, i.e. by the very process responsible for the adsorption lines in the spectre of the radiation we receive from the sun. Most probably photosynthetic bacteria of the cyanobacter species invent this coupling mechanism. The photosynthetic bacteria are almost as old as life itself, they originate 3-4 billion years ago when the common ancestors of all bacteria and eukaryotes appear. The first photosynthetic bacteria did not evolve oxygen. Later on, probably some 2.5-3 billion years ago, organisms appear that generate oxygen. This is a landmark event as it leads increase of oxygen in the atmosphere. The first photosynthetic organisms were prokaryotes. The eukaryotic oxygen generating organisms arise in the oceans by symbiosis between a host organism of a non-photosynthetic nature and a photosynthetic prokaryote. This event takes place less than two billion years ago. When life moves from the seas, some 500 million years ago, terrestrial photosynthesis starts to develop. This results in a quick increase of oxygen to its present level when the sinks for oxygen on earth (mineral and the oceans) saturate. In conjunction, the carbon dioxide level in the atmosphere becomes very low. Through the process of photosynthesis, a massive new source of economic value, stored in plants, becomes available and evolution enters a new stage. Herbivores start feeding on the plant material and in their turn become a resource that invites the development of new ways of coupling when the carnivores appear. As we explain, the human omnivore fits in this pattern of development of sources of economic value. An increasing diversity of sources of value and their exploitation by development of ways of coupling results in what we call the food webs that characterize the biosphere on earth.
Fig. 16.4 shows a typical food web.

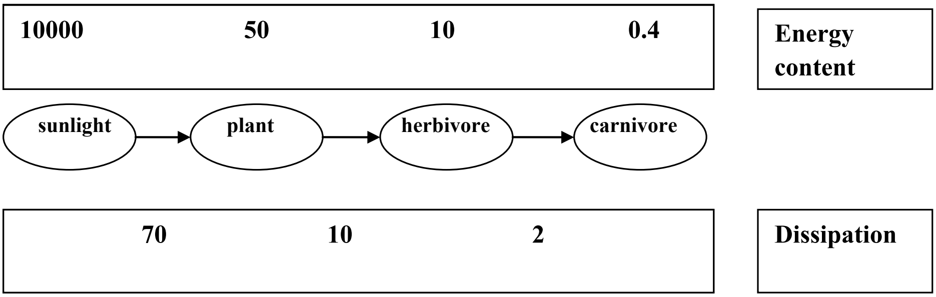

Fig. 16.4. Typical flows through a food web. Energy flows in arbitrary units.

Primary producers, mainly plants transform sunlight into their own biomass. If measured in equivalent energy units, the conversion of sunlight to biomass to plants proceeds at an overall efficiency of less than 1%. The second law requires that in the process of conversion of photons into biomass dissipation takes place. For primary producers the dissipation is of the same order of magnitude as the amount of energy converted to biomass. The next step in the food web involves herbivores that feed on plants, the so-called secondary producers. These convert plant biomass with an efficiency of about 20%. The dissipation is again about equal to the amount stored in biomass. Tertiary producers, carnivores, consume the herbivores at an efficiency of about 5% and the dissipation is 4-5 times higher than the amount stored in biomass. The food web shows that in the overall conversion of solar radiation to the top of the food web specialization takes place. We compare this with the kind of specialization that evolves in other chains geared at the exploitation of economic value, e.g. the industry value chain that we discuss later.

There is another way in which photosynthetic life on early earth contributes to the resources available to the socioeconomic system. Life in the past that decays fossilizes to resources such as peat, coal, crude oil and natural gas. Today this is the most important source of energy and resources for our industry. Around 80% of the resources we use, mainly as source of free energy, are of a fossil nature. In 2008, the energy consumption of the world is about 15 Terawatt (TW). The total influx of solar radiation, estimated at $1.2.10^5$ TW, is higher by almost a factor 10^4. The estimated formation of new fossils of about 0. 5 TW, is less than 5 % of the present use. It is clear that somewhere in the future, we must find other ways of coupling our economy to the free energy available in solar radiation. This applies even if we ignore the greenhouse effect that results from the use of fossils.

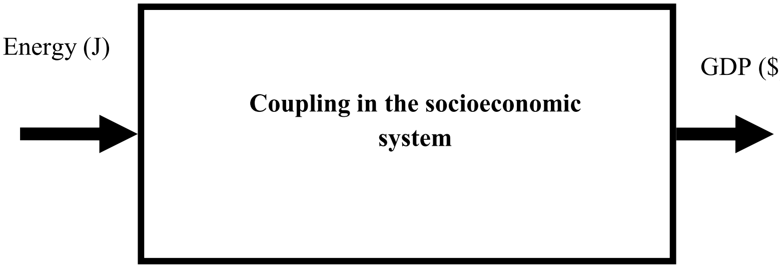

Fig. 16.5. Value transduction by coupling of energy consumption to GDP.

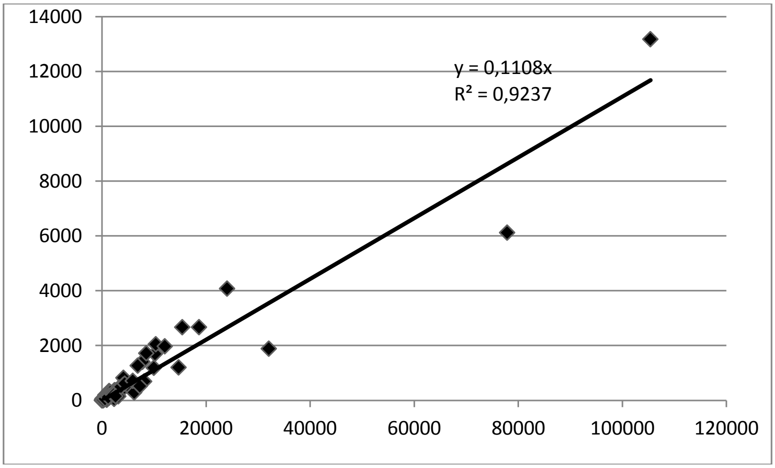

Fig. 16.6. GDP in billion international $ vs. energy consumption in millions of GJ.

The energy consumption of society is a good indicator for economic activity expressed as Gross Domestic Product (GDP). A strong positive correlation exists between GDP and energy consumption. We can tentatively describe the relation between energy consumption and GDP by the concept of the economic value transducer (Fig.16.5).

The correlation between GDP and energy consumption is strong indeed as we shown in Fig. 16.6. The figure shows the correlation (expressed as R^2, the square of the coefficient of correlation) for a wide selection of world countries. The coefficient of correlation for a linear relation is high indeed.

This reasoning also applies to the relation between GDP and the production of the greenhouse gas carbon dioxide (Fig. 16.7).

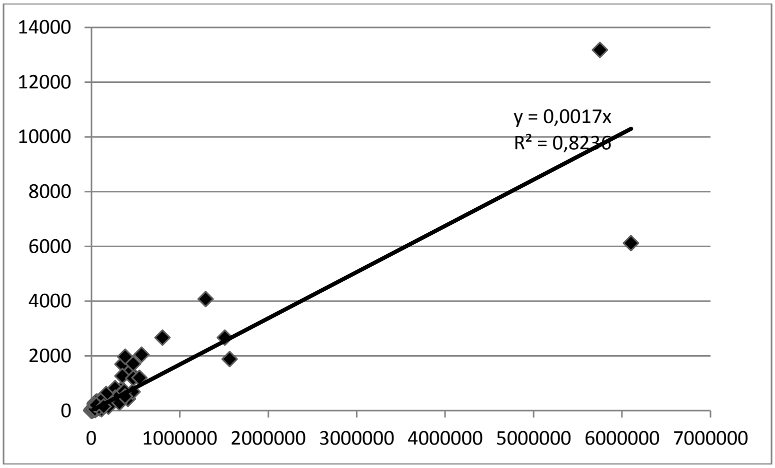

Fig 16.7. GDP in billion international $ vs. carbon dioxide emission in kilotons.

From the slopes of the linear correlations between GDP and energy consumption and carbon dioxide production respectively, we can estimate the amount of Carbon emitted per unit energy used. It calculates at 0.0177 tons of Carbon per GJ. This compares favorably with the number we calculate based on the energy resource mix the world used in 2007: 0.0191.

In doing the calculation, one comes to realize that the ratio of carbon emission to energy consumption depends on the nature of the energy resource used. On using natural gas the emission is about half of that on use of biomass, one of the so-called renewable resources. This also shows that when we use of biomass instead of fossil resources, carbon emissions will go up if we fail to produce the biomass used in the same amount in photosynthesis. Using existing biomass would lead to an increase in carbon emission. Table 16.1 present a summary of the relative amount of carbon emission per unit energy consumed for a number of common energy sources.

Table 16.1. Relative carbon emission for various energy sources.

Energy Source	**Relative Carbon emission per unit energy.**
Natural gas	1
Oil	1.33
Coal	1.99
Biomass (cellulosics)	1.90

We extend our line of reasoning by looking at the costs of energy. Oil at $75 per barrel leads to a cost of energy of $ 12.5/GJ, we can compare this with a GDP of on the average 111 $/GJ, i.e. a multiplier of about 9 compared to the basic costs of energy.

We can apply the same line of reasoning to industries that use economic value in the broader sense. In Chapter 8, we introduce the concept that industries and economic activity in general, derive economic value from bridging the gap between resources and information sourced from the environment and needs in society. Industry uses captive information to produce products and services that satisfy those needs in a competitive way. This leads to an economic value force to which the industry's phenotype, its products and services, couples in an increasingly effective way. These forces that result from information asymmetries drive our economy. Over the years, just as in the ecosystem on earth, the number of sources of economic value identified by industry drastically increases and this results in a strongly increasing GDP. This is in agreement with what we observe in the co-evolution of the earth and its biosphere and for that matter the tendency in the evolution of the whole universe. Also in industry, we see a strong evolutionary radiation.

The positive correlation between the consumption of energy and the GDP that characterizes our present-day economies and their evolution in the last 250 years, leads to another conclusion that we introduced in Chapter 5. We refer to the conclusion that coupling can exist between physical forces that derive from a gradient in free energy and economic forces that derive from a gradient in economic value. This is a prerequisite for the evolution of the socioeconomic system from the purely physical energy that the Big-Bang emergence of the universe creates. It follows that indeed the potential capacity to do work in the physical sense can lead to the creation of economic value by diligent coupling of the downhill force that results from the free energy forces resulting from the use of energy resources and the uphill process of the creation of GDP in the socioeconomic system.

Another striking observation rests in the longevity of information sets, DNA survives in the editing process that takes place in evolution whilst most of the actual biological species, the phenotypes, that appear in evolution disappear when time proceeds. The same holds for a company, most companies develop strongly different products and services over the years but the companies' information sets survive, albeit edited.

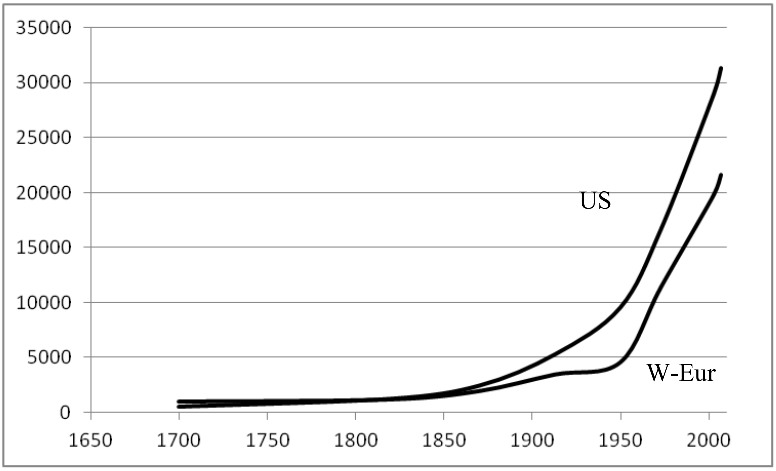

Fig. 16.8. Per capita GDP (1990 US Int. $) for the US and Western Europe as a function of time.

The process of industry evolution (Chapters 14 and 15) is apparent if we consider the historical evolution of per capita GDP for Western Europe and the US (Fig. 16.8). The data are from Maddison (2007).

The data of Maddison show that per capita GDP hardly increases unto 1700 AD. After the first industrial revolution an increase by a factor of up to 2 results. The second industrial revolution results in an increase of per capita GDP by a factor of roughly 15. This applies to both the economies of the US and Western Europe. It is also clear that the US overtakes Western Europe at the end of the 19th century and the effects of the world wars seem to be visible as a relative loss of position of Western Europe.

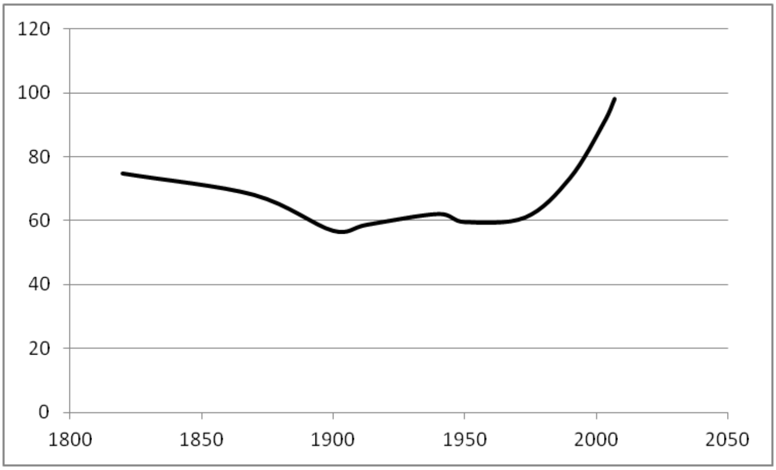

Fig. 16.9. Energy efficiency of the economy. GDP (1990 Int. $/GJ) versus time (Data from Maddison (2007)).

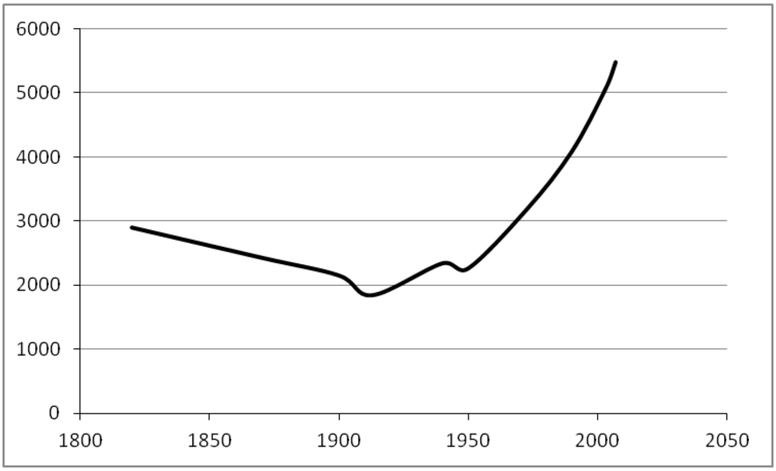

Fig. 16.10. Carbon emission efficiency of the economy. GDP (1990 Int, $/ton C). (Data from Maddison (2007)).

In Fig. 16.9, we show the energy efficiency of the generation of GDP (US $/GJ) as a function of time. Initially, the energy efficiency of GDP decreases with time up to about 1900. Subsequently, it stays more or less constant onto the early 1970s. Presumably, due to the oil crisis in 1973, the energy efficiency of the economy starts to increase in the more recent years.

We can also analyze the energy efficiency of the evolving economic system in terms of the ratio of GDP to carbon emission. Fig. 16.10 shows the time evolution of the carbon emission efficiency of GDP ($/ton C).

The pattern of the carbon emission efficiency of GDP mirrors that of the energy efficiency unto roughly 1900. After 1900, the energy efficiency stays constant until the early 1970s whilst the carbon efficiency increases, This is most probably due to the shift to more carbon efficient energy sources such as oil and gas rather than the less efficient coal.

Fig. 16.11. Time evolution of the world population in millions.

From the foregoing discussion, we conclude that the first and second industrial revolutions lead to an important increase in economic prosperity. From the data of Maddison (2007) we conclude that unto the first industrial revolution in the mid eighteenth century world GDP hardly increases. A strong increase starts after the first and particularly the second industrial revolution. This also reflects itself in the size of the world population as is evident from Fig. 16.11, again based on the data of Maddison (2007). It is clear that the industrial revolutions trigger a strong increase in the rate of growth of the world population in addition to a strong increase in per capita wealth.

It is interesting to note a further aspect of the development of industries over time. In the old days of the artisan, the whole process of delivering a product or service was in one hand. The artisans sourced materials from the environment and transformed these directly into a product for the consumer. Modern industry evolves away from this simple concept; typically, a value chain or a business column develops in industry. Fig. 16.12 shows a stylized example of an industry value chain, referring to the chemical industry, e.g. the production of pharmaceuticals.

In most instances, resources in the environment and needs in the market do not match directly. Industry bridges the gap that exists between the resources that are available and products that satisfy a need in society. These two positions are the alpha and omega, the beginning and the end, of the industry value chain depicted in Fig.16.12. In principle, one firm could perform the whole bridging operation. However, practical experience shows that this is generally not the case. A chain develops in which we first transform resources in general purpose type of base products, followed by steps in which, through intermediates and specialties, the finished product results that satisfies the need at the consumer end. The development of such a specialization critically depends on the differences in the competences, differences in information sets, needed to excel in the various stages of the industry value chain. At the consumer end, knowledge of consumer needs is critical. The big consumer products companies in the food industry, such as Unilever (Section 15.3), Danone and Nestlé, are examples of companies that excel in such capabilities.

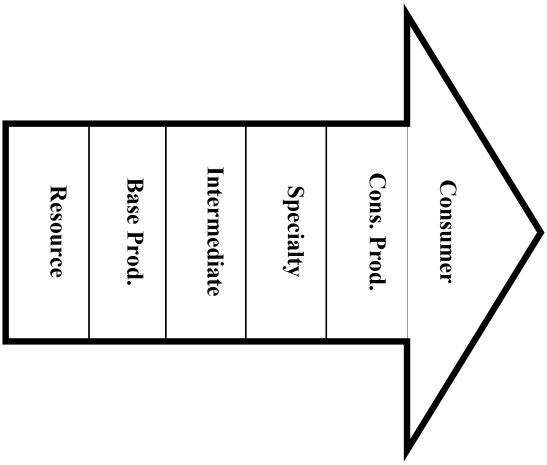

Fig.16.12. The value chain of an industry.

These companies spend large funds to perfect their information regarding consumer behavior and consumer needs. At the resources end, knowledge about effective exploration and sourcing of basic raw materials is critical. An example can be found in the fossil resources (such as oil) industries. Companies like Royal Dutch Shell (Section 15.3) spend vast resources in exploration technologies to locate and access new oil and gas reserves. This involves a complete different skill or information set than that of companies that operate at the consumer end. Apparently, those information sets are so different that it pays to specialize. This development of specialization increases since the second industrial revolution, although periods of a reversal of the trend also occur. The overall tendency, however, seems to be in the direction of increasing specialization (Sections 15.3 and 15.4).

The development of this kind of specialization involves at least the following aspects:

- The size of the industry, i.e. the amount of economic value available for the players in the value chain. If it increases, the value chain starts to differentiate. It concerns the total economic value resulting from transforming the resources into the consumer product. This is the concept of the wealth of the environment in biology or the extent of the market in economic theory (Nicolis and Prigogine (1977a)). The growth of economic value creation by industry in the period after the second industrial revolution certainly contributes to the increase in specialization during that period.
- The fact that sourcing an upstream product from a third party leads to economic benefits compared to in house production. This reflects the superiority of the information set of the player producing the upstream product because of specialization. In view of the uncertainties involved in sourcing the product from a supplier outside the firm, and hence the associated decrease in economic value, this advantage must be larger than the transaction costs of settling and policing the contractual relation. This brings us back to transaction cost theories (Section 8.11). These define various types of uncertainty, i.e. information asymmetries in transactions with outside parties, dedication of

assets and know-how, and unpredictable, opportunistic, behavior of the supplier and buyer. This becomes particularly important if the economic value gained by external sourcing is much smaller than the economic value at stake at the position of the buyer. In this respect, the market for branded pharmaceuticals is a case in point. Here the active ingredients price at only a fraction of the finished product.

- The existence of economies of scale and scope. Often products upstream in the value chain serve more than one market need, i.e. are part of many industry value chains. Also this aspect is of an informational nature.

Of the above-mentioned points, the first one is straightforward. The second point is more involved. The separation of the industry value chain in activities in different firms is stable only if it results in an increased economic value for all the firms in the chain. This is likely to apply if significant differences exist in the nature of the information sets necessary to operate optimally in the different stages of the value chain. Specialization allows the players to concentrate on their part of the total game and hence to develop the most suitable information set. Economies of scale and scope assist this, i.e. if different value chains depend on the same upstream product. This allows the upstream player to couple to a larger flow of economic value and hence to invest more in the perfection of its specific information set in learning by doing. If such economies of scale and scope do not exist, it is difficult to see why separation of the industry value chain would result in a situation stable against the forces of evolution. The specialization also allows the use of a smaller information set, and hence the rate of evolution increases as the copying fidelity limit can go down. In fact, in the chemical industry such economies of scale and scope automatically develop by the very workings of the industry and the tendency to use a given technology or competence for as broad a range of products as possible.

It is also possible to organize sourcing of upstream products through the creation of a separate entity within the downstream player. This becomes likely if the upstream product serves only one single business column. In this respect, a Hypercycle type of cooperation could result within the downstream firm.

The concept of the value chain shows a strong resemblance to the food web in biological systems. A differentiated food web evolves in the advanced stages of evolution, if we reach large enough overall forces driving evolution. This in both cases also depends on the observation that it pays to specialize. Another example we briefly highlight, concerns the force that drives the evolution of science. Here the driving force relates to the need to understand and harness the world around us. Science in its purest form derives from scientific curiosity, but in the authors opinion even the purists of the scientific profession, partly derive their curiosity from problems existing in society. Here the demand is clearly the curiosity driven or problem driven need to get the scientific understanding that allows us to harness the potential in solar radiation in an increasingly efficacious way. We return to this subject when we discuss cycles in the following section. Also in science, we see an increasing specialization and there definitely exist a value chain type of approach to science. Often people that formulate hypothesis and models are others than those specializing in the testing of these models and hypothesis by discriminating experiments.

16.3. The ubiquity of cyclic interaction.

We started this chapter introducing the learning cycle that drives sustained evolution. Such cycles driven by competition for scarce resources frequently appear in systems that evolve. These cycles provide feedback to information sets through the processes of competition for ultimately scarce sources of economic value. In this section, we analyze some of the cyclic interaction processes that drive sustained evolution in almost all sectors of the biosphere including the socioeconomic system.

The learning cycle.

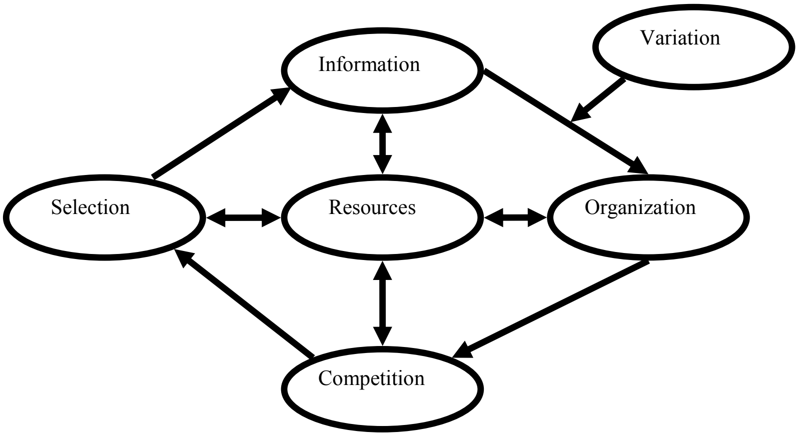

Fig.16.13. The learning cycle.

First, we revisit the learning cycle we discussed in Section 16.1. (Fig.16.13). Fig 16.13 represents the archetype of all the cycles that we discuss in this section. The starting point is an information set that in most cases initially derives from a fluctuation in a system. In some cases, the information set is tangible, such as the distribution of energy and mass in the universe. The initial information freezes when radiation and matter decouple some 300,000 years after the Big-Bang and it is still visible in the inhomogeneities in the cosmic background radiation. This information translates into some kind of organization, i.e. the stellar systems and the stars in the universe.

The forces of gravity cause these stellar systems to compete for the available energy or mass in the universe and this leads to the selection of a new matter and energy density distribution. The closed nature of the cycle provides the feedback loop necessary for the system to sustain its evolution. Important is that in principle in all the stages of the cycle interaction with scarce resources takes place. A double pointed arrow represents the interaction with the resources in the environment. The environment and its resources develops in the process of evolution in addition to driving evolution. We made this explicit discussing the concept of sustained evolution in Section 16.1 and Section 6.8. Fig. 16.2 shows that dissipative structures and their underlying information sets are both the product and the source of an increasing variety of forces.

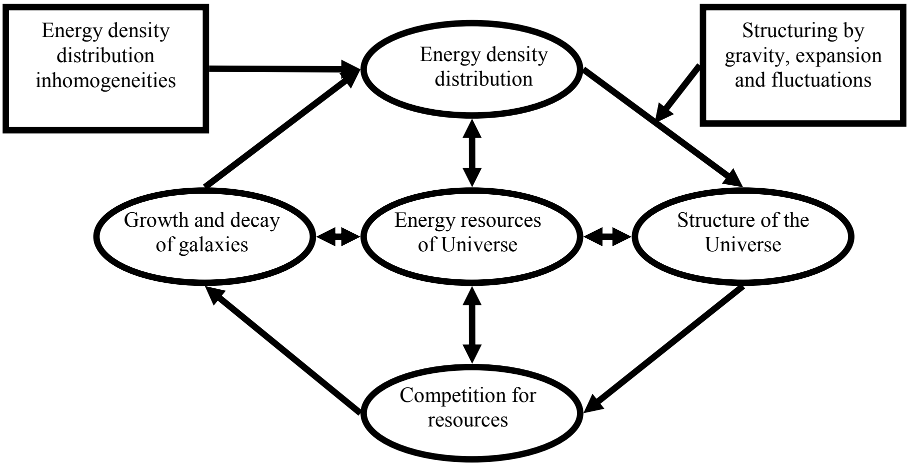

Fig.16.14. The evolutionary cycle of the universe.

Cyclic interactions in the evolution of the universe.

In Fig.16.14, we show our view of the cycle driving the evolution of the universe. As said, there exists a cosmic blueprint of the universe. Somehow, initial spatial differences in the energy density appear in the early universe. These stabilize and evolve after the decoupling of matter and radiation. The forces of gravity lead to a translation of this initial pattern into a structure, denser areas tend to grow autocatalytically and expansion and fluctuations influence that process of translation of the inhomogeneities. The initial inhomogeneities have the same function as the initial replicator molecules in biological evolution. Structures appear that compete for the resources created in the Big-Bang. The structure in the universe both shapes the forces of gravity and is a product of the forces of gravity. This is the very element of autocatalysis governing the evolution of the universe. In this process, galaxies grow and masses of hydrogen and helium collapse under the forces of gravity. This causes fusion processes to ignite the stars. The stars go through the familiar life cycle of emergence, growth, maturity and finally decay in novae or supernovae and add material to the resource fund of the universe. We recognize many of the features of the structures and processes that are a characteristic of the developments described, as typical from the perspective of the systems theory of evolution.

The carbon cycle on earth.

A vital element of life on earth is a photosynthesis driven carbon cycle (Fig.16.15). The early atmosphere on earth contains large amounts of carbon dioxide. This leads to a relatively high temperature also due to other greenhouse gasses such as methane. The early earth atmosphere contains hardly any oxygen and the initial life on earth is anaerobic. It uses no electron acceptors or electron acceptors other than oxygen. The life forms based on the information that evolves through the DNA replicators, quickly lead to an abundance of life forms using the source of potential forces in the solar radiation.

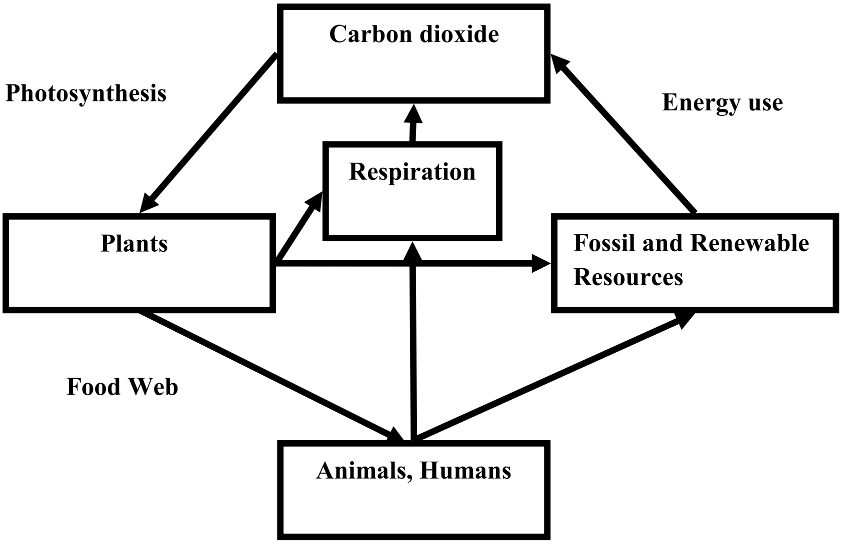

Fig.16.15. The carbon cycle.

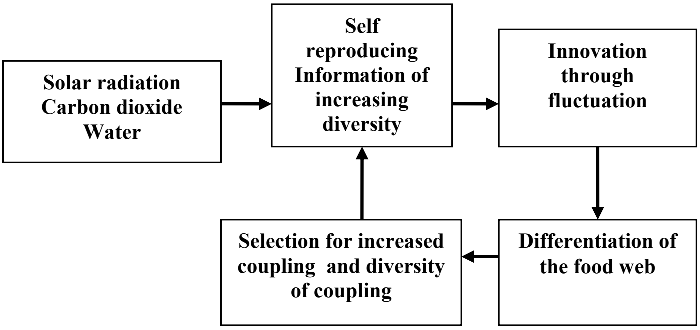

Fig.16.16. The learning cycle of evolution on earth.

This triggered the sustained evolution cycle depicted in Fig.16.16. One of the early innovations in the DNA based information sets was photosynthesis. The first ways emerge allowing direct exploitation of the resource the solar radiation provides. These first processes do not involve splitting water to obtain reduction equivalents and no oxygen evolved. Later on photosynthesis involving the production of oxygen develops. This starts the production of oxygen on our planet. Initially this does not result in appreciable amounts of free oxygen in the atmosphere as it dissolves in the seas and disappears by oxidation of minerals.

An endosymbiotic development leads to the appearance of eukaryotes containing chloroplasts. The organelle responsible for photosynthesis in algae develops. About half a billion years ago, land plants develop and photosynthesis starts to produce massive amounts of oxygen in the atmosphere. This opens new avenues for heterotrophes that develop oxidative phosphorylation as a mechanism to produce metabolic energy from plants; the herbivores appear. The brain starts developing and exogenic evolution starts challenging the monopoly of

DNA as replicator in biology. The appearance of the herbivores triggers the evolution of carnivores and finally the human species appears in which exogenic evolution results in the socioeconomic system of today. This introduces new sources of economic value in markets for goods and services. The cyclic interaction results in increasing exploitation of the potential value in solar radiation and results in an increasing variety of forces, an increasing variety of information sets and an increasing variety of organism and organizations in the differentiating food web and the socioeconomic system. We witness a sequence of life cycles in which organism and organizations appear, grow into maturity and decay if challenged by innovations in the exploitation and creation of sources of economic value. The present dynamic socioeconomic system with its markets, firms and other institutions, results. This includes the widespread use of fossil resources that provides the additional economic value to allow society to develop by coupling to the forces.

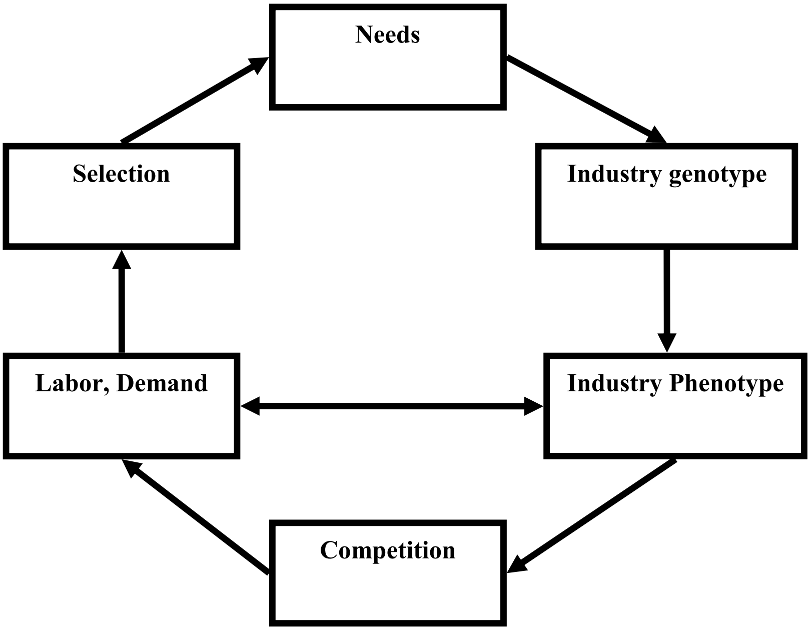

Fig 16.17. The supply and demand cycle.

The supply and demand cycle of industry.

Also in industries and economies, we see cyclic phenomena. Fig. 16.17 shows a representation of that cycle. A need in the market results in an opportunity for industry to couple to that need by developing an information set that allows it to produce products and services that satisfy the need. The industry's genotype starts to evolve the ability to couple to the potential value in the market need. The information set gives rise to the industry phenotype. It exists among others of its tangible and intangible assets and its labor force allowing the realization of the firm's products and services. The labor force leads to the development of purchasing power in the economy. In addition to the need, this is a prerequisite for demand. The classical theory of production distinguishes two factors that determine supply: Labor and capital. We maintain this distinction here, albeit that we want to stress that the factor capital does not only lead to tangible assets but also to intellectual assets such as know-how developed on the basis of information work, e.g. R&D, or licensing agreements, joint ventures and collaborations. We feel that in the years since the second industrial revolution, the value of intellectual assets shows a relative increase compared to that of tangible assets. This process enhances in the last 150 years of industry evolution. Other intangible assets, such as brands, show the same tendency. In many instances, the development of intellectual assets is a process that is much slower and less predictable than deploying tangible assets. This also results in the conjecture that information work increases in importance compared to the physical sources of work.

Generally, more than one firm supplies products and services geared at the need under consideration and as demand is a scarce quantity, competition results in selection of the products most appreciated by the existing need. This competition and selection process feeds back into the need and the industry genotype. This closes the cycle.

The co-evolution of humankind and society.

With the invention and perfection of the brain, a new phase in the evolution of life on earth emerges. New types of information sets emerge. In addition to the hardwired ones in the DNA molecules, a diversity of other ways of storing and communication of information appear that are not physically bound to DNA. E.g. through primitive forms of communication, through the development of language, in written forms, in science and in electronic storage and processing. The question is whether these different information sets evolve in their own right, i.e. whether they develop in distinct, separated, evolution processes. We have already seen that that is not the case. Science, technology, culture and society definitely co-evolve. To obtain a meaningful picture we have to consider the holistic information set. It includes the information in DNA and all the other more or less intangible information sets. This leads to an important conclusion about the evolutionary fate of humankind as a biological species. The exogenous aspects of the evolution of the socioeconomic system will no doubt feed back to evolution at the DNA level, i.e. will change the direction of the evolution of the biological species *Homo sapiens*, including the capacities of the brain. In fact, this is not a phenomenon new to the course of evolution. It is likely that the challenges of the environment resulting in the need to develop the exogenous component of human evolution towards the use of tools and mental models of reality, triggers the expansion of the brain in human evolution and closes the cycle between purely biological and exogenous evolution. In this way, humankind and society co-evolve in a complex pattern of closed cause-effect cycles. In fact, this holds for many of the cycles we discuss. There are definite feedback loops between the cycles we describe in this work. This complicates the analysis of the processes in the biosphere on earth and the socioeconomic system. It may lead to behaviour that is “unexpected” if we limit our analysis to the evolutionary cycles in isolation.

The evolution of science, the hypothesis experiment cycle.

Also science and its evolution follows a cyclic pattern (Fig. 16.18). The central resource in Fig. 16.18 is human intelligence, the source of abstract reasoning and the kernel of the process of observation. The science cycle involves observations on phenomena in reality leading to the need for understanding nature to access sources of economic value, i.e. to harness parts of nature. Science often considers the process of observation objective, but today we understand that assumptions appear already in this stage of the science cycle. Such problems emerge in quantum theory that challenges the existence of an objective reality and considers observations to interact with reality. We extend this notion in a broader sense. We perform observations on systems and as we argue (e.g. Chapter 4, Chapter 5) that information is not a free commodity but comes at a cost, observations come at a cost. To obtain information we have to perform information work on the system. Furthermore, we cannot divide the amount of information indefinitely, the smallest amount of information being one bit. This leads to the conclusion that in each observation process we have to perform a finite amount of information work, this involves a force that disturbs the system we observe. We conclude that we do not observe reality but reality as influenced by our observation.

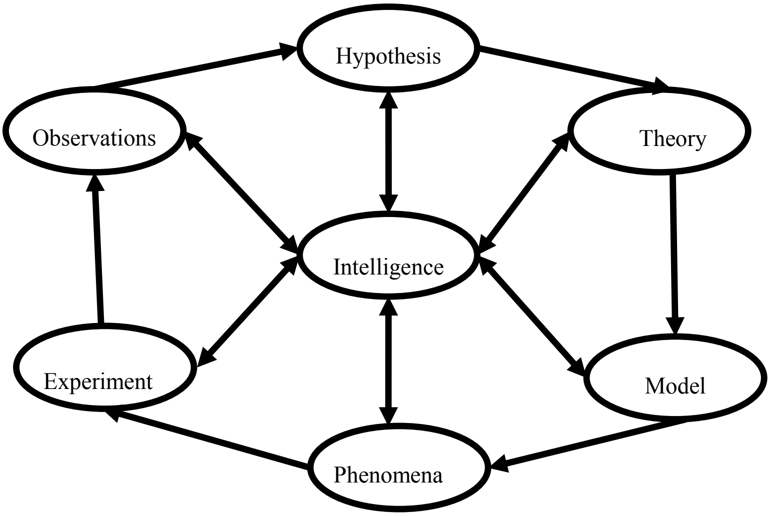

Fig. 16.18. The cycle of the development of science.

From observations, a process of induction leads to the formulation of a hypothesis about the nature of reality. The hypothesis forms the basis of a theory such as gravitation, relativity or the theories of quantum mechanics. A model concerns the application of theories to a restricted part of reality with the objective of arriving at testable predictions of the behavior of a system. Scrutinizing these predictions by experiments and additional observations results in acceptance of the hypothesis or its rejection and revision (Popper (1963)). The experiments we choose to perform and the conclusions we reach from the observations, are not independent. The choice of experiments and the interpretation of the results, critically depend on the model we use. These scrutinizing experiments are the selection mechanism governing the evolution of science. Theories acquire the status of scientific laws when withstanding the challenge of verification for a sufficiently long period. A paradigm of science is a collection of laws or theories accepted by the scientific community at a given moment in time.

There are several interesting questions regarding the science cycle. One we want to highlight here is that theories and models often lead to predictions that defy experimental validation at the time of their formulation. An example is the effect of gravitation on the path radiation follows in a gravity field. The bending of the rays of light that the general theory of relativity predicts is subject to testing during the famous eclipse expedition of Eddington in 1919. The existence of the neutrino is another example of a theoretical prediction verified much later. In addition, the positron is an example of a particle theoretically predicted by Dirac.

As we indicated earlier, an interesting question concerns the relation between science and technology. Both science and technology are information sets that can drive the transformation of potential value into economic value. The difference between science and technology is that technology is the set of tools or competences used by society to develop desirable products and services, whilst science primarily aims at understanding reality. Technologies find their value in increasing the competitiveness of products and services in the satisfaction of society's needs. Science can lie at the roots of new technologies. This is, however, not a necessary condition for technologies to exist. Often technologies develop in a stage where their full scientific understanding is beyond the scientific state of the art in the period under consideration. An example is the development of the production of penicillin at the end of the Second World War where the scientific understanding of the nature of microorganisms and the detail of the biosynthesis of penicillin were only limitedly available. It is true that new scientific theories may lead to new technologies for application in industry. Conversely, established technologies or new emerging technologies may lead to new scientific questions that defy scientific understanding at the time of the first application of the technology in products and services. Technology and science are part of a closed cycle and the question, which holds the driver's seat has no unique answer. This is a consequence of the cyclic nature of the interactions in our evolving world.

There is one further element that we want to include. It concerns the interaction between (university) teaching and the progression of science and technology. Certainly, teaching established theories and facts in science and technology, plays an important role in the furthering of science and technology. If we were not communicating the information sets underlying science and technology, there would not be an evolution in the direction of further perfection of these instruments vital to the competitiveness of our species. Every generation would have to start from scratch and evolution would not proceed, as the copying fidelity would be far below the limit allowing directed evolution to take place. However, there should be room to challenge existing beliefs. In absence of such challenge, evolution comes to a halt. We discuss this in Chapter 12 where we analyze the evolution of science and the causes of the scientific revolution. Therefore, the art of judicious challenge of existing beliefs needs to be as vital a part of the academic curriculum as the teaching of the scientific state of the art.

To analyze the interaction between science and technology in more detail, we refer to the innovation cycle that is a characteristic of industry (Fig. 16.19).

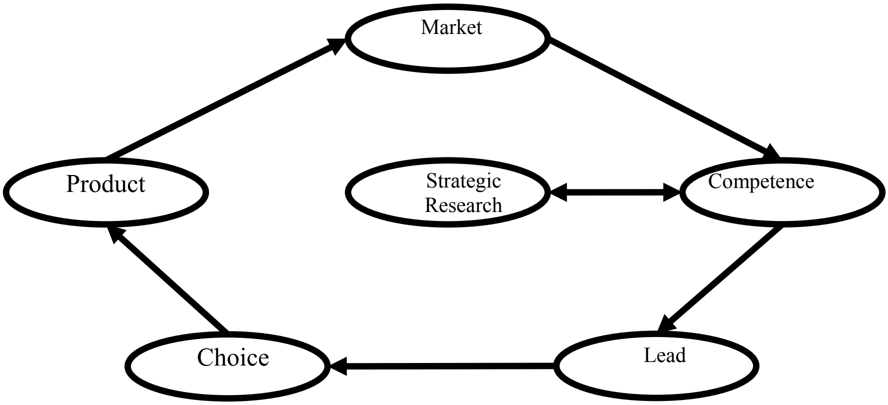

Fig. 16.19. The innovation cycle.

The cycle illustrates that developments in the market if confronted with the competences of a company, may result in leads for the development of new products and services or new processes. A process of limitedly rational choice and the subsequent development processes, create a new product. What follows is confronting the product with the market. In case the product proves to be less successful than expected, this may lead to the need to change the competence base to allow adaptation of the product to better suit the observed needs. In fact, the cycle does not always start at the market need. New product ideas not always derive from developments in the market, i.e. do not always depend on market pull. In fact, the more breakthrough new products often derive from developments that start at the competence level where new technological possibilities provide the inspiration for radically new products.

An interesting aspect of the cycle is the interaction between the competence base and so-called strategic research. Although for some time less in vogue, many corporations have strategic long-term research programs that do not directly aim at developing new products. These derive from the strategic need to fill gaps in the competences based on a vision concerning expected developments in the markets in the future. These strategic research programs often provide a fruitful area for interaction between science in the academic sense and industrial research. Scientific curiosity is no doubt an important inspiration for scientific progress but it helps when this curiosity follows needs in society including the competence needs of its industrial component. In fact, given the clear impact of technological innovation on economic growth, government policy with respect to innovation tends to stimulate a focus of academic research on needs in society. In the Netherlands, this resulted in cooperation between industry and academic institutions partly funded by government: This is the so-called leading technological institutes approach.

The author does not want to imply that all academic research should relate to existing or expected needs in society. There should be sufficient room for pure curiosity driven research. This is an important catalyst for the optimization of scientific progress. Its necessity, also in an evolutionary perspective, derives from the inherent unpredictability of the evolution of science and the uncertainty about the evolution of the needs in society. The author thus wants to emphasize that there needs to be a proper balance between need driven and curiosity driven research. The funding policies of governments and other institutions should be fine-tuned to reflect the need for such a balanced approach.

16.4. Evolution and sustainability.

If we consider the evolution of humankind, the relation between evolution and energy consumption clearly emerges (Section 16.2). An analysis of the present-day Gross Domestic Product (GDP) and energy consumption for a wide selection of countries, leads to the following correlation:

$$\text{GDP (billion US \$)} = 0{,}11\text{x (energy consumption in millions of Gigajoules)}$$

This strong correlation between the evolution of energy use and the development of economic wealth is also apparent from an analysis of the evolution of the GDP in the history of humankind (fig. 16.8).

Per capita GDP hardly increases unto 1700 AD. After the first industrial revolution (1750), it increases by a factor of up to 2. The second industrial revolution (1850) triggers a roughly 15-fold increase. This relates to the causes of the first and second industrial revolutions. The first one rests on the heat engine and introduces fossil resources, coal, as energy source. The second revolution involves a shift to oil and gas.

The whole history of evolution, including that of humankind and the economic system, depends, in a direct or an indirect way, on energy resources available from solar radiation. Solar radiation is the overwhelmingly major source of free energy and economic value that drives evolution on earth. The first and second industrial revolutions result from using of fossilized energy resources derived from photosynthesis in the past. In the end, this resource is not sustainable. The present-day use of fossils vastly surpasses their creation. If we take the correlation between GDP and the use of energy resources as given, this observation leads to concerns about the sustainability of the present economic wealth. In addition, other complications exist. Using fossil energy resources results in an increase of the concentration of carbon dioxide in the atmosphere and the greenhouse effect of this gas may result in global warming. This becomes a political issue regarding the increase of the usage of energy, society's energy efficiency, and the use of renewable resources. Renewable resources derive from the direct use of solar radiation to produce biomass or electric power through solar cells, or its indirect use through e.g. wind and water. These alternatives make sense only, if the net effect is indeed decreased dependence on fossil resources, i.e. if the use of fossil energy in these direct applications of solar energy does not nullify the gain in renewable energy. In addition, these measures must make sense economically, i.e. should not result in an unacceptable decrease of the GDP.

Both the exhaustion of fossil resources and the changes in the environment on earth by an increase in the concentration of greenhouse gases, are examples of co-evolution of biological species and their environment. As we have seen (Section 10.6) this is not a new phenomenon, life and the conditions of earth co-evolved in the past. The question is whether the socioeconomic system is at stake if we leave developments to the blind force of the co-evolution process. This causes increasing challenges for policy makers and results in several approaches to mitigate undesirable aspects of these processes. The analysis of available options from these two perspectives on the use of fossil resources is not straightforward. In this section, we attempt a contribution from the perspectives of the material presented in this book.

We have seen that thermodynamics and its extension to EVT provides a powerful tool for the analysis of systems that exchange resources with the environment and in which transformation of such resources takes place. In Section 3.4 and Chapter 5, we applied our formalism to the formulation of constraints that apply to any system in which transformations and transactions take place and exchange of resources with the environment occurs. We show that the total effect of the processes in the system must result in a decrease in economic value; at least if the system is not in equilibrium and macroscopic processes do take place. If a system is in a steady state, i.e. of its macroscopic state does not change, a positive flow of economic value from the environment to the system must compensate the destruction of economic value in the processes in the system. This immediately implies that sustainability in terms of economic value cannot apply to the system and its environment together. If a relevant part of the environment shows change due to its interaction with our system, its effect needs consideration in the analysis and hence it becomes part of the system we have to analyze. In fact, the industrial revolution, triggered by a strong increase of the use of in principle non-renewable, i.e. not sustainable, sources of economic value, causes changes in the environment on earth that become relevant in recent years and we cannot ignore this in policy analysis.

We argue that we cannot define energy in an absolute way. We need a reference state to which we relate changes in energy. For our analysis here, we conveniently choose that reference state as the free energy or the economic value of the resources at the ambient conditions on earth (of course we take into account that this reference state may change if the co-evolution process causes the ambient conditions to change).

We state that, if we want to keep the relevant system in a steady state, we need a supply of economic value from resources that are inexhaustible or are outside the relevant system. This principle allows us to do economic value accounting on the flows that leave and enter the system. Accounting on the flows leaving and entering the system greatly reduces the complexity of the analysis in complex systems in which many value-bearing assets appear and many transformations and transactions take place. In almost all cases, the meaningful exchange flows with the environment are far smaller in number than the assets in the system and the processes that take place in the system. Examples of such a reduction of complexity exist in microorganisms and for that matter in the socioeconomic system. In fact, microorganisms are champions of sustainable operation. Many compounds in a microorganism are part of a closed cycle and do not cross the boundary between the organism and its environment in a meaningful amount. The same applies to a well-designed factory where many compound appearing in processes never leave the factory in significant amounts.

Present-day economies couple to forces resulting from the energy contained in mainly fossil resources. We use these resources at a rate exceeding their creation today. These resources are not part of a closed cycle and the sustainability of our economies cannot rely on these resources in a long-term perspective. We need the information (the technology) to more effectively couple to the energy available in solar radiation. This is a prerequisite for sustainability of both the environment and economic activity.

We describe the processes on earth from the perspective of a heat engine by determining the equivalent temperature of the energy in the source and the sink applying to earth. Present-day estimates put the so-called black body "temperature" of the sun at 5780 K and that of the earth at 288 K. This allows a heat engine with a maximum efficiency of 95%. Considering the energy contained in the solar radiation, this allows a power generation of more than 150.10^{15} Watt. We can compare this to the total energy consumption of humankind of 15 TW at 0.01 % of the potential. What we conclude is that the free energy potential in solar radiation is more than adequate to allow sustaining our economies and the growth of our economies. What we need are more efficient ways to couple to the potential of solar radiation using technologies not leading to undesirable side effects, such as exhaustion of non-renewable resources, or undesirable changes in the conditions on earth (e.g. a significant greenhouse effect).

The beginning of this section identifies the relation between GDP, an indicator of economic wealth, and energy consumption. As we explained in Section 16.2 this is an example of coupling. In our socioeconomic system, we create forces by consumption of energy resources and couple these to economic work, resulting in the GDP. These resources contain a potential value and economic work serves to generate economic value. Information work results in transforming part of the potential value into the capacity to perform economic work. It relates to decreasing statistical entropy by information work. In order to sustain our economy we need to couple to a source of economic value.

One way to sustain economic activity at a reduced level of energy dissipation, or to grow it at the same level, is to increase the energy efficiency of the economic system. This is the analogue of finding better ways of coupling the output of a heat engine to the input. It is clearly desirable as more GDP results from a given quantity of free energy. In Section 16.2 we show that this indeed takes place in the post Oil Crisis years.

If we want to grow our economies in absence of increased efficiency options, we need to find new energy resources as our present resources are not part of a closed cycle and are hence subject to exhaustion. Potential availability of such energy resources is not the problem as we highlight in this section. We use only a fraction of the economic value potential of the solar radiation and our present energy usage is only 15% of the yearly rate of biomass formation in photosynthesis. The problem is lack of a technology that allows use of alternatives for fossils on a large enough scale. Society's genome did not yet evolve the information to turn the potential energy in solar radiation to a free energy resource large enough to fuel our economy if we cannot rely on availability of fossils.

As said, there is may be an additional problem involving another cycle that is not closed. A problem associated with combustion of energy resources is that it involves carbon dioxide production. The cycle does not close for carbon dioxide as the present rate of consumption far exceeds the rate of fixation of carbon dioxide in fossilization. In fact, this also applies when using renewables, such as wood, if we do not close the cycle by a concurrent increase of the carbon dioxide fixation in photosynthesis. If we calculate the carbon dioxide production per unit of energy consumption the best option is use of methane, using oil results in an increase of 30-40%, coal use scores 2 times higher than methane. Biomass use comes close to coal; here carbon dioxide production exceeds that for methane by 90 %. How can we reconcile this last observation with the cellulosics debate regarding a sustainable energy supply? The answer lies again in the need to close the cycle by increasing carbon dioxide fixation in photosynthesis to compensate for the biomass used as an energy resource.

We summarize our conclusions so far as follows. The widespread use of fossils leads to two types of problems. Firstly, fossils do not fit in a sustainable energy supply, as the cycle does not close. This leads to exhaustion of the available resources and potential greenhouse problems. The extent to which this greenhouse effect will cause problems is still subject to debate. There is some room for increasing the efficiency of the creation of economic wealth from free energy. The same problem applies to most other resources, also for these we need to close the cycle to assure sustainability. Only if we tap directly into solar radiation without disturbing any cycle, true sustainability results.

An inherent problem of all C-based technologies is that we may face global warming if we do not close the carbon dioxide cycle by increasing the photosynthetic production of biomass or again drastically increase the efficacy of coupling energy consumption to the generation of economic wealth.

We close this section with a summary on sustainability issues in evolution of the socioeconomic system and introduce some hints to avoid pitfalls in sustainability accounting.

The first remark regards the definition of the boundaries of the system. We repeatedly encounter this problem. The boundaries allow identification of the resources we import and the products we export to the environment. All other items that appear in the system must either not participate in processes or be part of a closed cycle. For each of the resources and waste products, we need to ascertain whether these are part of a closed cycle, if this is not the case scarcity and exhaustion may result. In addition, wastes not cycled may change the environment in an undesirable or unacceptable way (e.g. the greenhouse effect).

Furthermore, we inevitably need to dissipate free energy to sustain economic wealth and we need to ascertain that this ultimately directly couples to the use of solar radiation.

Another important cycle we need to consider refers to the supply and demand cycle in the economy. Disturbing this cycle may induce price changes of resources thus jeopardizing the economic viability of the system. Profits in the operations in our economy may turn to losses and this may preclude the economic viability of the operations. In addition, government policies such as taxes or subsidies may induce such economic effects.

We now analyze two processes from the perspective of the sustainability characteristics introduced above. The first generation renewables based energy supply processes, emerges in the late 70s and early 80s. These processes rely on the use of agricultural resources such as corn. First, we discuss the macroeconomic perspective, i.e. from the perspective whether the processes contribute to solving society’s energy supply problems. In this analysis, the production of the agricultural resources used and for the matter all significant other resources used, needs to be included within the system’s boundary. The only free lunches are solar radiation and ambient temperature heat. Most of the analysis to date show it questionable whether the first generation renewables technology makes sense from the macro perspective.

This analysis proves to be less straightforward as it seems. An important problem relates to the accounting of the dissipation of non-renewable energy. How much of this energy can be attributed to the co-products, in the case of most bioethanol programs these products are used in the feed sector. The correct approach would be to attribute these by-products the dissipation of non-renewable energy in the production of these by-products in other processes, i.e. in the production processes of the products these by-products replace. Even then, there is the question of relative sizes. Are our by-product flows of such magnitude that these can be accommodated by the end-use from which these derive their value and what are alternative end-uses for these products? The question of relative sizes is also important from a microeconomic perspective if these by-products disturb markets supported by products from other sources. This may cause a shift in the economies of the process as the by-product credits are often an important factor in the economic viability of bioethanol.

Similar considerations apply to the question whether these approaches contribute to solving the greenhouse gasses dilemma.

The example also illustrates another problem we highlighted. In bioethanol, the introduction of a new application for existing resources leads to competition for e.g. corn in its existing end uses in food and feed. The introduction of a new outlet for these resources makes their scarcity felt and the prices increase also in the food and feed related market. A fair approach involves growing additional corn for use in the bioethanol sector. This development was foreseen (e.g. Kooreman and Roels (1989)). Still, political forces led to ignoring the problem that was sure to emerge and indeed materializes. From the socioeconomic perspective, energy for dissipation is a lower quality application than its direct or indirect use as human food. Subsidy policies to stimulate the use of bioethanol as an automotive fuel increase the disturbing effect.

Another question refers to the sustainability of bioethanol from the microeconomic perspective, i.e. from the perspective of a firm that operates in the market for bioethanol. This analysis rests on the price of the main inputs that are not captive to the firm, e.g. corn and energy resources, and the market price for its main products being ethanol and the by-products, such as animal feed and corn oil. Tax incentives and subsidies disturbed this picture and where main drivers behind the introduction of bioethanol.

There is important progress in recent years in the so-called second-generation renewable technologies. These rely on using cellulosic resources such as wood or agricultural wastes. Enzymatic and microbial conversions of these resources to energy carriers show great improvements. However, the pre-treatment technologies and the recovery of the ethanol from the rather dilute solutions, continue to introduce a high energy burden. The author lacks the information to judge whether these processes contribute to net generation of energy resources even if we assume the agricultural waste to have zero free energy in the accounting exercise. If we again consider the macroeconomic perspective, the accounting should at least include: use of free energy in the production of the enzyme used, in the production of the yeast, in the pre-treatment (e.g. for heat or the production of the acids used), in the enzymatic treatment and the fermentation to produce ethanol, in the recovery of ethanol and in the production of the cellulosic resource or waste. We should also compare this with alternatives such as gasification and liquefaction.

There is again a separate consideration regarding the micro-economic perspective. If the wastes become a viable energy resource, they become valuable and their prices increase. Furthermore, we have to consider alternative uses, such as in animal feed.

Availability of cellulosic wastes may preclude a large contribution to a sustainable energy supply if fossils dwindle. From my perspective, a real solution is a closed cycle with regard to the cellulosic resources and dedicated production through forestry needs to be included in the process. Again, alternatives like liquefaction and gasification are a yardstick.

Our take away message is that the only free source of economic value is that contained in the solar radiation. To be on the safe side all other resources should be part of a closed cycle or truly unlimitedly available in the time perspective we choose to adopt. Even then, we should always be aware of the limitations of our ability to foresee future developments in this respect.

In this discussion, we ignore solutions based on nuclear technologies, such as nuclear fission or fusion. Of course, in a complete picture their economic and social feasibility needs careful consideration.

16.5. Reflections on predictability issues in evolution.

In this section, we present some preliminary reflections on predictability issues regarding the evolution of the global system on earth. This is still work in progress. Many aspects have not reached a stage of development that allows definite conclusions.

As we indicated, evolution of the system on earth involves many evolutionary feedback cycles. We also indicated that these cycles cannot be considered in isolation as co-evolution phenomena lead to a coupling of the developments in the various cycles. In this respect, we highlight the co-evolution of the DNA hardware and the technological and culture aspects of human society, in Section 16.3. In Section 16.4, we discover that the supply and demand cycles of the market for automotive fuel and that for animal and human feed, start to co-evolve by the introduction of the use of bioethanol, derived from agricultural resources that are also used in food and feed applications. As we indicate, such coupling may lead to drastic changes in the supply and demand cycle in the food and feed market. In fact, drastic changes, bifurcations, in the development of one evolutionary cycle, generally lead to bifurcations, resulting in unexpected developments, in one or more other cycles that are coupled and co-evolve. This may drastically change the global evolution in a complex system such as the socioeconomic system on earth. Microeconomic evolutionary bifurcations may lead to unexpected and unacceptable changes at the macroeconomic level.

In fact, such phenomena frequently appear in biological systems. The invention of photosynthesis, resulting in oxygenation of the atmosphere on earth, causes a massive extinction event regarding anaerobic life on earth. Also exogenous effects may cause such a crisis. In this respect, we refer to the extinction of the dinosaurs when a meteorite hits the earth and a temporary strong decrease in solar radiation reaching the earth results. The consequences this has for the primary production of sources of free energy in the form of plants, leads to a strong decrease of the food supply for the dinosaurs. In fact, this coupling and its consequences, applies to the overall evolution of many of the food webs on earth.

We further note that the development of the GDP importantly relies on the availability of sources of free energy as the use of energy couples to the generation of economic value in the overall economic system. Apart from efficiency improvements that are the result of more efficient coupling of energy consumption to GDP, new sources of economic value for the economy can only result if we succeed in making more free energy available from the energy potential provided by solar radiation. This strongly resembles the sustainability issues we discussed in Section 16.4. It also means that economic gain in evolutionary cycles that are not directly driven by the primary source of free energy and economic value in solar radiation, are either not sustainable, if these depend on e.g. fossil energy sources, or are dependent on the free energy provided by the process of primary coupling to solar energy. The economic gain in these derived evolutionary cycles, e.g. those supporting markets at the microeconomic level, derives from asymmetries in information between, e.g. buyers and sellers. Such information asymmetries may cease to be relevant if drastic changes in the overall global evolution occur by the disturbance of one or more of the separate evolutionary cycles that support the global evolution.

We conclude that drastic changes in one evolutionary cycle may challenge the stability of the global evolution and may induce instabilities at e.g. the macroeconomic level. This results in a loss of predictability of the future evolution of the system.

These considerations may be relevant to recent developments where turbulence in the housing market and the introduction of new financial instruments in the banking system, induce a massive disturbance of the global financial system. This results from the fact that the information asymmetries involved in the profitability of markets for financial products, cease to be relevant due to unexpected developments in the housing market and the associated mortgage market.

We can illustrate this further if we take refuge to the Capital Asset Pricing Model that we discuss in Section 5.3. That model derives uncertainty of the future value of financial assets (and hence, in our model, economic value) from the standard deviation of the returns on assets in the past. Of course, this measure of uncertainty and economic value may cease to be relevant if drastic changes in the overall financial system lead to an increase in the standard deviation of the relevant returns in the future.

In the following, we further illustrate the concept of co-evolution of separate evolutionary cycles, using a simplified example. We analyze an industry value chain (Section 16.2) from the economic perspective i.e. in terms of our EVT formalism. For simplicities sake, we consider an industry value chain in which only two organizations appear. The first organization is the primary producer, the upstream organization. It converts a resource that is available in the environment into an intermediate product. We assume that the resource is renewable as it derives directly from the solar radiation that is the ultimate driver of the processes on earth. The organization downstream produces the final product of the value chain; its resource is the product of the upstream producer. This is a highly simplified picture of an actual value chain in industry but it serves to analyze some general features of such a chain. Both organizations evolve their own information set that allows it to extract economic value from its resource. In both cases, the resource presents a given potential value. The economic value the organizations can realize from their activities is lower than the potential value of their resource. Firstly, they avail of only part of the information needed to fully harvest the potential value in their resource. Their lack of information determines the potential value that is unavailable due to the statistical entropy of their state of information. This reduces the economic value they can extract from the potential value with an amount equal to the product of their cost of information and their statistical entropy. Secondly, in the process of conversion of the resource to their product, dissipation needs to occur due to the restrictions posed by the second law. Thirdly, part of the potential value conserves in the product. The potential value corrected for these three quantities represent the economic value that the organization creates. In this way, the upstream and the downstream producer contribute to the total creation of economic value in the value chain that connects the initial resource to the product of the chain. The operation of both companies are made possible by the nature of their (captive) information sets that are, of course, different. The upstream producer does not avail of the information to convert its product to the final product of the chain. The downstream producer does not avail of the information to produce the product of the upstream company from the primary resource of the chain. We see a clear example of specialization and we discuss the reason why such specialization evolves in Section 16.2. The total economic value produced by the participants in the value chain is by definition higher than the economic value created by each of the separate players. In a practical situation, the price mechanism that governs the market for the intermediate and the final product can be an additional contribution to the creation of economic value in the chain but we assume that these transactions do not add additional economic value to the output of the total chain. The two organizations in the chain evolve their information set in supply and demand cycles that are coupled and engage in co-evolution. If the environmental conditions do not change (and the companies behave responsibly), a situation will evolve that is characterized by a smooth development of the economic value creation by both companies. Of course, the balance of power may shift and the share of each company in the total creation of economic value may change. However, the two companies are mutually dependent as we assume that the producer of the first product depends on the demand created by the downstream producer. If there is no longer a demand from the downstream side, value creation in the first step stops. The same holds for the downstream step that depends on the product of the upstream producer for its resource. If the upstream producer stops supplying, value creation stops. As said, in absence of drastic changes, the co-evolution process will lead to a selection process that results in a smooth evolution of the overall chain.

However, if drastic changes occur, e.g. if for some reason there is no longer a demand for the final product of the chain, or if the primary resource is no longer available, the chain will no longer function and the information sets of the company lose their significance as an instrument to create economic value. In addition, from the macro perspective of the total chain, both microeconomic contributions of the separate activities to the total creation of economic value are lost to the economy's "GDP".

In an actual economy, many of such dependent cycles appear and this will surface as a complication in case of drastic changes in one or more of the microeconomic evolution cycles in the economy. This complicates analyzing the economic system in cases of drastically changing conditions.

An additional complication results from the workings of the financial system that supports our economic system. The need for a financial system results from the fact that the activities of players may result in the creation of economic value that exceeds the amount needed to satisfy their needs. In that case, they generate an excess of liquid capital. The reverse situation may also exist, in that case, the economic value creation is less than the amount necessary to satisfy the player's needs. This leads to a supply and demand cycle on its own, a capital market evolves. In general, this capital market does only create additional economic value beyond the value created in the other supply and demand cycles in the so-called "real economy", if it leads to a more effective allocation of capital to the most promising investment proposals. No doubt, the smooth functioning of capital markets can be essential to harvest the fruits of the evolution that supports the socioeconomic system, in an optimal way. In organizing the markets for capital, the banking system plays an important role in orchestrating supply and demand in financial markets. The lending and borrowing policies of the banking system depend on the judgment of the risk associated with the lending to players that are in need of capital. This results in a separate supply and demand cycle in which the banking firms operate. A crucial factor is the judgment of the risk associated with lending; this influences the economic value creation or destruction in the bank's business model. In normal times when no drastic changes take place in one or more of the other evolution cycles in society, this is rather straightforward. In phases of turmoil, the banking system is subject to the problems we indicated for our model value chain earlier in this section. From the macroeconomic perspective, this may lead to loss of economic value and this may results in crises.

16.6. Conclusion.

In this Chapter, we present an overview of some important features of the evolution of complex non-equilibrium systems. We discuss the increasing variety of information sets, both tangible and intangible, that participate in the evolution of the socioeconomic system on earth.

In addition, we highlight the increasing variety of forces that evolves from the potential value in the non-equilibrium situation that results when the universe emerges. We show that the forces, being the sources of free energy and/or economic value, induce competition between the information sets that characterize the structures and organizations in the system. In fact, the structures and organizations on the one hand and the forces on the other hand become part of interrelated cause and effect cycles that are a characteristic of complex evolving systems. This leads to problems in the analysis of such systems in its own right, but it also leads to the conclusion that the environment and the system show increasing interaction and significant parts of the environment may start to participate in the co-evolution in the system.

We also highlight that, following the first and second industrial revolutions the sustainability of our growing economic wealth in terms of GDP, strongly relies on the availability of sources of free energy. Today, the majority of these free energy resources are of a non-renewable nature and this leads to sustainability issues with respect to both the environment and the socioeconomic system.

We conclude that the potential value that reaches earth in the form of solar radiation is more than adequate to sustain the present GDP and allows for strong further growth. The challenge is to develop the information (or the technology) to couple to this resource in such a way that we realize both sustainability and avoid undesirable side effects.

Such undesirable side effects may result from the disturbance of some of the cycles that drive the evolution of the socioeconomic system. An example is the carbon cycle, if we influence it to the extent that increasing concentrations of carbon dioxide lead to an unacceptable greenhouse effect. In addition, we need to be careful that we do not introduce unacceptable disturbances in other supply and demand cycles that exist in our economy. We encounter this problem if we engage in the use of agricultural resources for the production of fuels and thus disturb the supply and demand pattern in other economic sectors, such as the markets for food and feed products.

This also applies to the use of cellulosics if we do not close the cycle of use and production through dedicated forestry.

Finally, we discuss the impact of complex interacting co-evolution cycles on predictability issues regarding the evolution of our economic system.

CHAPTER 17. EPILOGUE

17.1. Introduction.

This book analyzes the principles of an Economic Value Theory (EVT) and a systems theory of evolution and shows their application to tracing the origin of the firm. In this chapter, we broadly review the perspective that we develop and illustrate some prospects and limitations of the analysis we perform.

We extend thermodynamics by the introduction of economic work to complement the traditional sources of work and introduce an Economic Value Theory. We also show in a number of ways that economic work can couple to the traditional physical sources of work. This results in the observation that a continuous line of evolution connects the Big-Bang that created the universe with our present-day society with its markets, firms and other institutions. Our analysis largely rests on two important bodies of theory: Macroscopic thermodynamics and its statistical foundations and the Neo-Darwinian theory of evolution.

Developments in thermodynamics in the second half of the 20th century reconcile evolution of complex structures, so-called dissipative structures with the basics features of thermodynamics, such as the first and second laws of thermodynamics and the concept of temperature. This is one of the important scientific developments on which this work rests.

17.2. Summary of the highlights of the basic theory.

The theory this work develops, rests on the concept of information and its value. This shows to be the basis of macroscopic thermodynamics and it extension to EVT, as well as the root mechanism underlying all the evolutionary processes we observe in our universe, including the biosphere on earth. In fact, information enters the theoretical developments in two different, rivaling ways. Firstly, we discover the role of information in macroscopic thermodynamics en its extension into a full-fledged theory of value transformations and transactions. Secondly, there is the role of information in all evolution phenomena, both in the purely physical world, e.g. in the evolution of the universe and in biological evolution, and in the case of exogenic evolution. In exogenic evolution, also “non-physical” information evolves. In fact, these two aspects of our overall theory drive information in opposite directions. The second law of EVT implies that processes and transactions result in the destruction of information. In the approach to evolution, we introduce a learning cycle that creates the information that allows so-called dissipative structures to evolve in the direction of increasing information. This is the very essence of the functioning of such organized structures.

The EVT formalism is, just as classical thermodynamics, a macroscopic theory. Our EVT formalism in its extended form includes thermodynamics but it is not limited to the purely “physical” forms of work and information. The extension of the theory rest on recognizing that economic work exists and that information includes information that does not relate to physical objects. Exogenic information is an example of such information. This type of information proves to be an important driver behind the evolution of humankind and society.

EVT presents a reduced information picture of a complex reality and quantifies the information that is lacking in such an approach in terms of the statistical entropy of the description. It also defines the cost of information and states that information is not a free commodity. To obtain information we need to do information work, this involves a minimum expenditure of value in the extended sense we use it in our theory, i.e. considering both energy in the thermodynamic sense and value in the economic sense. In absence of information work on a macroscopic system, our information about the state of the system can only decrease if the system under consideration is not at equilibrium in the thermodynamic or in the extended EVT sense. This is the essence of the second law of EVT that states that processes in a system can only result in the destruction of information that has an economic value. In the terminology of this book, we define a process in which information decreases as a process in the "natural" direction. We also use the term downhill for such a process by analogy to the natural direction in which rivers flow.

The Neo-Darwinian theory of evolution states that the evolution of living systems involves the limited fidelity communication of valuable information, i.e. information that increases the competitiveness of a species. Limited fidelity refers to the fact that the genetic code, the DNA, is communicated subject to error. This sometimes results in a code that is more valuable, i.e. leads to an organism with an improved competitiveness. In this theory, competitiveness is defined through a selection process that results from the interaction of the organism with its environment. The evolution process leads to a learning by doing cycle in which information is generated. The cycle results in an evolution against the natural direction the second law dictates. In our theory, we generalize this to communication of information beyond the DNA or RNA based system in biology. We generalize it to include other systems of coding and communicating information. We indicate these collective alternative systems with the term exogenic information. In this way we generalize the Neo-Darwinian approach to evolution, to include systems in which also other communication systems than the DNA based system exist. This approach strongly resembles the one leading to an extension of thermodynamics to EVT. The mechanism of learning by doing we just described results in an evolution process that is uphill in the terminology of EVT, i.e. proceeds in a direction against the natural direction the second law dictates. In such a process, valuable information is created rather than destroyed.

These combined extended theories lead to a systems theory of evolution that includes the evolution of organizations that appear in the socioeconomic system. The firm is an example of such an organization.

The mechanism of the evolution of complex structures rests on applying the concepts of the two extended theories to systems that are not in equilibrium in the thermodynamic and/or economic sense. This applies to the universe and the systems that appear in the universe, such as the earth and its biosphere and the socioeconomic system on earth. The source of this non-equilibrium is the energy that was created when the universe emerged in the Big-Bang. It is the source from which all aspect of the biosphere and the socioeconomic system derive.

To understand the mechanism by which our system evolved and continues to evolve we need coupling of the two directional processes, i.e. the processes in the natural direction that involve a loss of valuable information and the creation of valuable information in the process of evolution. This requires defining the forces that drive these processes. The concept of a force develops in a branch of thermodynamics, non-equilibrium thermodynamics or irreversible thermodynamics that emerges in the beginning of the 20th century. In this theory, the force that drives processes is defined as a difference in the ratio of free energy and the temperature. In our extended EVT formalism we define such forces as the ratio of the change of economic value (economic value includes but is not limited to free energy) to the cost of information. A downhill process, in the natural direction, involves a positive force, i.e. a process in which economic value is lost. Evolution, a process in which economic value or valuable information is created, involves a negative force, it is an uphill process.

To complete the theory we introduce the notion of coupling of uphill processes to downhill processes. The downhill forces result from the non-equilibrium situation that arises when the universe emerges. The uphill forces result from the coupling to those downhill forces that the information that is created in evolving structures allows. A prerequisite for an evolution to proceed is that in the combined uphill and downhill processes valuable information is destroyed. In this way, the overall development confers with the second law. The structures that evolve in such an evolutionary process are termed dissipative structures as these maintain their structure and develop further by dissipation of some of the economic value in the downhill processes that the overall non-equilibrium situation allows.

A further observation is that dissipative structures and the forces that allow their creation co-evolve, i.e. the structures derive from the forces and also create the forces. This creates a cyclic interaction in which the forces and the structures and the underlying information, co-evolve. It leads to a blurring of cause and effect as the forces result from the activities of the structures and the structures develop as a result of the forces.

A further complication is that in the course of an evolution in a non-equilibrium situation as in the universe, many forces and many structures employing a wide variety of ways of storing and communicating valuable information evolve. To all these combinations of forces and information sets, the mechanism of co-evolution applies. Furthermore, these individual co-evolution cycles become coupled in what we term “Hypercycles”. This leads to severe difficulties in prediction the behavior of complex evolving systems and to behavior that is counterintuitive from the perspective of the reduced information characteristics of our picture of reality.

The theory shows that where asymmetries in economic value exist or develop, it is possible for systems to move away from equilibrium by the mechanism of coupling to the resulting forces. In that case, organization phenomena result in the appearance of dissipative structures as discussed above. In the near equilibrium region, i.e. where linear phenomenological equations and the reciprocity relations apply, near equilibrium dissipative structures evolve that are stable. Once such a structure appears, it does not evolve further. The interaction between individual actors determines the stable steady state and the creation of more complex organized forms of matter, or broader organizations, does not occur. This leads to a conflict with the phenomena that we daily observe in the biosphere on earth, with its rich variety of highly organized forms of matter and in human society with its economies, institutions and industries. This conflict disappears on moving beyond the strictly linear region. If the forces increase, we reach a critical limit where the structures, stable in the linear region, become unstable. This does not necessarily imply that the structures characterizing the linear region immediately disappear. However, the structures become conditionally stable. Here the limitations of the macroscopic description, based on the averaging the microstates of the system to arrive at a reduced information macroscopic picture of the microscopic reality, enter the picture. In fact, constant fluctuations of the macroscopic state variables take place due to the microscopic reality of the system. The macroscopic picture averages fluctuations away, but in reality, these take place and serve to test the conditionally stable steady states. The nature of the fluctuations determines the further evolution of the system at the critical points of stability. At these so-called bifurcation points, several evolutions become open that determine the future developments in the system. As the fluctuations do not appear in the macroscopic description the “choice” the system makes at the bifurcation points cannot be predicted from the perspective of the macroscopic observer. This introduces uncertainty in the evolutionary path of the system. In addition, this introduces a historical dimension in the evolution of the system as the future evolutions that remain open to the system depend on the fluctuation that determine the fate of the system at the critical bifurcation points. This leads to the notions of inevitable “change” induced by the unpredictable “chance” of the fluctuations taking place. At the bifurcation points, steady states become unstable. This requires the existence of, or the creation of, sufficiently large forces. Beyond the linear range, more complex dissipative structures may and generally will appear.

An example of such non-linearities beyond the strictly linear region is autocatalysis, where a structure increases its own rate of growth. We frequently encounter autocatalysis in biology and in many phenomena in our socioeconomic system. We conclude that organism and organizations in the socioeconomic system are examples of dissipative structure that operate beyond equilibrium and beyond the strictly linear region. In this analysis, we follow and extend on, earlier work of Schrödinger (1945) and Prigogine and co-workers (Glansdorff and Prigogine (1971), Nicolis and Prigogine (1977), Prigogine (1980) and Prigogine and Stengers (1984)). These organized structures develop because of coupling to sources of free energy or economic value that derive from an initial non-equilibrium situation. The structures become increasingly ordered and develop in the direction of decreasing statistical entropy. This development is a direct consequence of the fact that, beyond the strictly linear region, the macroscopic low order branch becomes unstable with respect to some fluctuations that occur in the system. We enter a regime of sustained evolution.

An important conclusion is that the organizations are both the result and the source of the forces that exist and develop in the environment. This introduces the notion of co-evolution by cyclic interaction as an essential feature of the evolution phenomena we identify in this work. Therefore, the forces drive evolution and evolution drives the development of forces. In fact, organizations and organized matter are both the product and the source of gradients in economic value.

If we reach a threshold of complexity of the information sets, the diversity of the possible structures becomes very large and the domain explored in the course of the evolution of the systems and their environment is much smaller than the range that is accessible. We enter an arena of sustained evolution in which steady states are only locally stable and may become unstable if fluctuations in the system and its environment become larger. This leads to a succession of life cycles in which structures appear, grow, become mature and decay by displacement by other emerging structures. This is a typical consequence of sustained evolution. It culminates in complex systems existing of entities based on information and its corresponding functionality. A closed cycle of information and its functionality emerges. As said, dissipative structures are models of biological systems and in this work and in earlier work, we extend this to include industries, economies and markets. In fact, any structure that stores, processes and communicates information can become subject to evolution and is a source of evolution.

This notion includes a broad range of evolving information sets, including science, technology, language, and art. We also conclude that both the variety of the information sets and the variety of the forces increase in the process of evolution. This leads to an increasing number of interacting cycles between information sets, corresponding structures, institutions and organizations. Furthermore, this leads to a co-evolution of the variety of information sets, i.e. different information sets start to interact and change the direction of their mutual evolution.

A few other general features derive from the theoretical development in this work.

The first one derives from the distinction between the system and its environment, a common notion in contemporary macroscopic modeling. We show that evolution proceeds in the direction of an increasing interaction between the system and its environment. An assumption in modeling is that change processes only take place in the system and that the environment does not change. This assumption breaks down if the level of interaction between an evolving system and its environment increases to such an extent that the environment starts changing. We show that this indeed frequently is an inevitable consequence of the increasing interaction between the system and the environment. This leads to the necessity to include an increasing part of the environment in the system we want to analyze.

A further complication rests in the complexity of the systems we have to analyze in studying systems such as the socioeconomic system. It contains a large number of interacting entities that are in itself already complex and for a full microscopic description, we would need an amount of information that is not accessible in an economic way or even is fundamentally not accessible. That is why we resort to macroscopic models in which many microstates are lumped together in one macroscopic state based on a process of averaging, i.e. each macroscopic state represents the average of the probability distribution of the underlying microstates. This leads to problems if the selection processes that direct the evolution of the system, favor extremes of the distribution function rather than the average. This is the case if non-linearities characterize the processes in the system. In that case, the macroscopic method contains insufficient information to predict the future evolution of a system and hence fails to predict the direction of the evolution. This leads to situation in which developments in the system come as a surprise to an observer that has only a macroscopic reduced information picture of the system.

A further problem rests in an important modeling principle. We have to assume that the model and the system we model are independent, i.e. we have to assume that the existence of the model does not influence the processes that take place in the system. We indicate that this is a questionable assumption if the actors of the system know the conclusions of the model and use these to change their behavior.

17.3. The Origin of Markets and Firms.

In this section, we summarize the developments that connect the Big-Bang with our ‘present-day socioeconomic system. We discuss many of the elements involved in the pertinent sections of the book. Still it seems worthwhile to provide the reader with a concise overview.

The story starts when the non-equilibrium situation arose that marks the birth of the universe in the Big-Bang some 13.5 billion years ago. This event created the lifetime supply of energy of the universe or, taking a broader perspective, all the potential value available to drive the universe’s evolution.

A few minutes after this event, the nuclei of the light elements, mainly hydrogen and helium, emerge in a mass ratio of 25% Helium and 75% hydrogen. These nuclei are charged particles that have not bound to electrons to become electrically neutral. The neutral elements become firmly established some 300,000 years after the emergence of the universe.

As witnessed by the background radiation that still exists in the present-day universe, there are, when radiation decouples from mass as a consequence of the particles becoming charge neutral, small inhomogeneities in the energy and mass density in the evolving universe. Under the force of gravity, these inhomogeneities become the blueprint for concentrations of mass in the form of hydrogen and Helium. These are the precursors of the galaxies and stars and these constitute the “DNA” of the early universe. When the concentrations of mass become large enough, they start to collapse under the force of gravity. This results in an increasing temperature and finally the stars ignite and start to emit radiation fuelled by fusion reactions. First hydrogen fuses to helium, later on the other elements form in the fusion reactions. In this way, elements like Carbon, Oxygen, Nitrogen and Phosphorus form, i.e. vital constituents of life on earth emerge. The final fate of the stars is that these collapse and explode in novae and supernovae when their fusible elements become exhausted. This leads to expelling the elements formed in the stars into the universe where these form the source of new stars and planetary systems. One of these galaxies is the Milky Way in which our sun resides. This galaxy emerges a few million hundred years after the birth of the universe. The sun and the solar system, including earth, develop some 5 billion years ago.

A landmark event is the appearance of life on earth. This introduces the first self-replicating information sets that we highlight as an important prerequisite for sustained evolution of the type we see on earth. Very soon, the information sets that characterize life, take the form of DNA and RNA based macromolecules. In the beginning, life fails to avail of an information set that allows direct coupling to the potential energy available in the solar radiation. It relies on inorganic resources and indirect products of the solar radiation. We also identify emerging co-evolution cycles, e.g. the co-evolution of the phenotype and the genotype, and the co-evolution of life and the forces that life creates based on the energy potential on earth. The first sign of co-evolution of life and the environment on earth start to be visible.

The process of perfection of life's information set in the competition for scarce resources and in the creation of new forces from the potential on earth, results in the invention of photosynthesis. This introduces another major bifurcation point in the evolutionary process. It sets the scene for an important co-evolution of the environment on earth and the life forms that evolve on earth. Photosynthesis results in the generation of oxygen. Photosynthesis originally develops in the seas, but later photosynthesizing life colonizes the land. This also marks the point of a strong increase of the oxygen concentration in the atmosphere. It quickly increases to its present-day level. The appearance of the land plants and the availability of oxygen, induce at least two important bifurcations. With the land plants new potential source of free energy become available, this induces the evolution of the herbivores that feed on that newly available resource. In addition, life forms start to use oxygen in the generation of metabolic energy from plant material. This process gives access to far more metabolic energy than the anaerobic processes that reigns before the advent of oxygen. The herbivores induce the appearance of carnivores and the land animals in a broader sense.

The primates, ancestors of humans start emerging in the period in which the dinosaurs are the reigning species on earth. The further development of the primates most probably results from another critical bifurcation event. In this case, an event of an exogenous nature. Apparently, a large meteorite hit the earth some 65 million years ago. This leads to a shielding of the solar radiation by material expelled into earth's atmosphere. This strongly decreases the supply of plant material from photosynthesis. It leads to a mass extinction of the dinosaurs and creates the evolutionary space in which the primates further increase in number and diversity.

The appearance of the animals marks another important bifurcation event. In animals, the brain appears and the first, be it primitive, steps in the evolution of new evolving information sets result. Communicating and perfecting information is no longer the exclusive domain of DNA and RNA based macromolecules. New replicators appear that open new routes to the further evolution of life.

Some 5-7 million years ago, the early ancestors of the humans, the Hominins, appear. These have already relatively large brains at one third of the brain of present-day humans. Some 2.5 million years ago, the genus *Homo* appears and the first seeds of technology appear when tool making and the perfection of hunting emerge. The pace of exogenic evolution starts to increase. In fact, this early development of technology based on exogenic transfer of information, induces a co-evolution cycle between the hardwired DNA based information and the exogenic process of the communication and development of technology. This leads to an expansion of the brain and a concurrent increase in the pace of technological progress.

The origin of *Homo sapiens,* our species, traces back to 200,000 years ago or less. With the appearance of our species, the pace of exogenic evolution increases. Language develops as a new means of communication. It reaches a level of sophistication less than 100,000 years ago. These developments also rely on abstract and symbolic thinking as other landmarks in the evolution of new ways of understanding reality and transferring information. Roughly 12,000 years ago, the technology starts including animal domestication and agriculture as a way of harnessing the environment. This is the result of the increasing pace of exogenic evolution resulting in a further sophistication of technology and access to new ways of transferring potential value into economic value.

With the advent of modern writing some 6000 years ago, another decisive step follows. This also marks the advent of abstract thinking that later on will prove vital to the scientific method as a way of harnessing reality through exogenic evolution of new information sets.

The other landmark event that enhances the pace of exogenic evolution is the scientific revolution in the 17th century. It introduces the information sets that evolve as the basic precursors of scientific progress as a way of harnessing the potential value available in the environment.

The scientific revolution is no doubt one of the important triggers of the industrial revolutions that start in Britain in the 18th century. These importantly rely on the technology to get access to a new abundant source of potential value in fossil resources, products of photosynthesis in the past. Firstly, this relies on coal as an energy source. Later on, in the 19th century, the use of oil triggers an increase in the pace of the industrial revolution, if we enter the second industrial revolution. This process of industrialization rests on a strong further evolution of exogenic information sets, such as science and technology, that show an increasing pace of co-evolution. Increasing investments in public and private R&D further fuel this. The phenomenon of industrial R&D strongly increases as a driver of socioeconomic evolution ever since the advent of the second industrial revolution.

In studying the developments since the advent of the industrial revolutions, we make an important observation. Economic wealth, expressed as per capita GDP hardly increases until the onset of the industrial revolution. It shows an exponential growth, particularly when the second industrial revolution takes off. In addition, we note a strong correlation between a growing GDP and the level of consumption of energy resources by society. There is a clear relation between the creation of economic value in the socioeconomic system and society's access to sources of free energy. We see this as a strong evidence for the fact that the capacity to do useful work that is available in energy resources, couples to performing economic work resulting in the creation of economic value. This process rests on the availability and the evolution of largely exogenic information sets as embodied in e.g. science, technology, firms and other aspects of human culture.

The forces we use to create economic value do directly derive from the non-equilibrium situation and the potential to do work that emerged in the Big-Bang on the creation of the universe. I think it is fair to say that our line of reasoning closes the gap between the creation of the universe and our present-day socioeconomic system with its industries and markets. We effectively trace the origin of the firm as dependent on the co-evolution of a variety of information sets of both a tangible, e.g. DNA, and an intangible nature, e.g. science, technology, and the institutions in the socioeconomic system.

17.4. Prospects and challenges.

We hope to have convinced the reader of the validity and the merits of the approach this work develops. It clearly connects the sequence of events that results in today's socioeconomic system. In addition, the theory sets a direction to the further evolution of our world. It certainly allows us to explain the sustained path of evolution in a qualitative sense.

However, we also show that the actual path evolution takes and will take, is beyond the predictive powers of the theory. We trace this to a fundamental problem. The complexity of reality precludes the development of a full model of reality in all its microscopic complexity. This forces us to rely on so-called macroscopic models of a complex microscopic reality. These models are characterized by uncertainty that quantifies in a statistical entropy. The statistical entropy results from the circumstance that the reduced information picture of reality is based on lumping many microstates together in one macroscopic state based on averaging the properties that characterize the microstates. In fact, the macroscopic state fluctuates between the many microstates that are compatible with our macroscopic picture of the system. The probability of the system being in a given well-defined microstate is given by the probability density function of the microstates that defines the statistical entropy. By definition, the macroscopic model does not contain information about the nature of the fluctuations and the likelihood that these will materialize.

At some critical points, called bifurcation points, the macroscopic state of a system may become conditionally stable. This may be the case and generally will be the case in systems that are characterized by forces beyond a certain critical level. In that case, different future evolutions in terms of the macroscopic features of the system may become open to the system. The actual path the evolution will take then depends on the fluctuations that appear in the system. The nature of the fluctuations determines the "choice" the system makes between the various evolutions that may occur. By definition that information is not present in the macroscopic description and is hence beyond the predictive powers of any model, including the EVT based approach.

We also show that this is a fundamental problem. In principle, we can think of a model that contains far more information than the information present in the macroscopic description. In practice that is impossible because information is not a free commodity. It comes at a cost in terms of value or energy. This leads to the concept of temperature in thermodynamics and to the concept of the cost of information in the full EVT formalism. It also implies that in order to obtain that information, we have to interact with the system by performing information work on the system. This in principle changes the probability distribution of the microstates and hence the future evolution. We conclude that the information to describe the future evolution of the system is not contained in the present state of the system at the macroscopic level. The theory thus fails to provide full guidance in the question what the future evolution of a complex system will be.

What can we do, based on the theories developed in this work? To illustrate a possibility, we discuss an example in industry: We return to the process of fermentative production of a basic raw material for semi-synthetic penicillins, the production of Penicillin G. As discussed earlier, the production of so-called β-Lactam antibiotics traces back to the discovery of Fleming in 1928 (Fleming (1929)). The Second World War induces its industrial production in the early 1940s. The varieties of the production organism used in the early days produce only minute quantities of the antibiotic, and the first quantities came at a unit cost prohibitive to its wide spread application. Over the years, the production costs of Penicillin G decrease by a factor of at least 10,000. This was due to a learning by doing approach towards the development of the competence involved in production of penicillin. The literature (Van der Beek and Roels (1984)) shows that the efficiency of the penicillin production process improved dramatically over the years. Penicillin derives from a so-called fed-batch fermentation process in large stirred vessels, called fermenters, in a submerged fermentation of the mould *Penicillium chrysogenum.* Over the years 1962-1982, the productivity, expressed in units per unit fermenter volume and unit time, increases almost fivefold. This was, among others, the result of developing superior varieties of the production organism, varieties devoting more of their metabolic energy to the production of the desired product than the original production organisms. This results from a kind of artificial evolution, in which we introduce random changes in the genetic information of the production organism, followed by selection of better producing mutants. This artificial evolution, based on very little detailed knowledge about the changes introduced in the genetic information, proves to be a very effective strategy. Later on, following the market introduction of penicillin as an antibiotic, the scientific understanding of the biosynthesis increases and this results in new possibilities for optimization of the production process. However, it becomes clear that the process of random selection of superior strains optimized many enzymes in the biosynthetic route towards penicillin before obtaining the knowledge of details of the biosynthetic pathway. This shows the power of random mutation combined with directed selection, i.e. the process on which Darwinian evolution relies. In recent years, DSM elucidated the genome, i.e. the complete information set of the organism at the genetic level, and a more directed approach to the optimization of the organism is in principle possible (Van den Berg et al. (2008)). This is, however, not a straightforward task because even if we know the total genome, prediction of the results of the full interplay of the enzymatic machinery is beyond the power of the mathematical methods used in the prediction of the production characteristics of genetically modified strains.

The example of the development of the penicillin production process over the years illustrates an important point. Certainly, the metabolism of the mould responsible for the industrial production of penicillin is far too complex to undertake a directed manipulation of its genome and to predict the impact of this on the production of penicillin in large-scale fermenters. The advent of genomics and genome sequencing, does allow the identification of all the genes in the organism and in principle allows us, although this is already more tricky, to identify the function of these genes. Prediction of the structure of the cell and its metabolism based on the interplay of the identified gene activities is, however, still far beyond the reach of the present state of the art. The author is of the opinion that this will still involve a lengthy and costly research agenda and is in doubt whether this will ever be realistically possible. Still we were able to improve the productivity of the production organism and decrease the cost of producing penicillin by orders of magnitude. This progress relied, as stated above, certainly initially, on wholly random mutation of the genome and clever selection methods, resulting in an artificial evolution of the strains in the desired direction of more economic production of penicillin. Apparently researcher were able to arrange this game in such a way that the course of evolution proceeds in an in the eyes of the industrial players desirable direction. The researchers successfully created an environment in which the evolution took a desired course. In this process, as in evolution on earth, most of the mutant strains did not survive the harsh rules of that mutation and selection process. In more recent years, the knowledge in biochemistry and genomics makes a more rational and directed approach possible, at least in principle, and some successes emerge in the present industrial research practice. However, many of the rational design approaches lead to improvements already realized in the decades of random mutation and directed selection. In modern approaches to the development of industrial strains, clever combinations of rational and random approaches remain the preferred and productive approach and this will not change in the near and even the remote future. The organism, in the past a true black box with no means to look inside the box, maybe changes into a grayish box, but the phase of complete transparency and rational manipulation is still far away. We still find us in a situation that in the design of industrial strains we are, as is the case in the managing of a firm, in a stage where bounded rationality prevails. A significant statistical entropy is still present.

The development that we highlighted shows that, in absence of detailed knowledge of the information set, and this is of course always the case in socioeconomic systems, where the analogue of genome sequencing does not exist, useful strategies are available to introduce desired characteristics. This could be a direction in which an evolutionary strategy regarding the socioeconomic system could be useful, although the possibilities for experiment are much more limited than in microbiology, where the introduction of variation only occasionally leads to improved strains. This is hardly an option for the management of a firm as we are not likely to be in a position to repeat an experiment if it is unsuccessful. The question is whether we can devise methods of control and intervention, from both the company and the macroeconomic perspective, that result in an evolution considered beneficial. This involves an uncompleted and challenging research agenda that is beyond the scope of this work.

Furthermore, we also highlight the problem regarding the application of models to system where actors are aware of the existence of the model and its predictions. This leads to loss of independency of model and system, a basic assumption in modeling.

A complicating feature adds to the complexity of analyzing aspects of the socioeconomic system. We indicated that phenomena of cyclic co-evolution apply to many processes in these systems. These separate cyclic co-evolution loops may start to interact in the course of the evolution taking place in the system. This often leads to unexpected behavior. As an example, we refer to the cycle that drives scientific progress and its interaction with the technology cycle in industry.

We discussed another more involved example in Section 16.4 where we analyzed the bioethanol programs in the context of the renewables approach to automotive fuels. These programs results in a co-evolution of the supply and demand cycle for automotive fuels with that for animal feed and human food. In addition, government policies, such as tax incentives that favor bioethanol, further disturb the evolution of the cycles that are in place. This leads to very complex interactions and the danger of creating new co-evolution cycles.

In fact, elements of government policy will co-evolve with private objectives of the actors in the system. This leads to a blurring of cause and effect as far as the government interventions are concerned.

Recent experience clearly shows the limitation of our picture of the world's economic reality and the statistical entropy of that picture. We witnessed, or maybe we are still witnessing, an economic crisis that some claim to have predicted, but certainly not the majority of the mainstream experts. Certainly, many players in the stock market did not foresee this development as is clear from the development of the S&P 500 index as depicted in Fig.17.1.

Fig.17.1. The evolution of the S&P 500 until 2010.

Given the collapse of the index, the economic developments were a surprise to many investors. Of course, we know that governments and regulators reacted to this crisis by taking measures geared at restoring our economy. In the jargon of Darwinian evolution, they attempted measures to direct the evolution in a more desired direction. At the time of the writing of this book, this seems successful (at least it has to date not been clearly detrimental). It is also instructive to note how the experts differed and continue to differ of opinion about the measures that are most likely needed. Of course, the problem is not that no models exist that aim at directing and getting grip on the economy. Models widely exist but are apparently subject to significant statistical entropy.

LITERATURE CITED

Aghion P., P. Howitt (1998), Endogenous Growth Theory, MIT Press, Cambridge (MA)

Alchian A.A. (1950), Uncertainty, Evolution, and Economic Theory, J. Political Economy, 58, 211-221

Alchian A.A., H. Demsetz (1972), Production, Information Costs, and Economic Organizations, American Economic Review, 62, 777-795

Andrews F.C. (1975), Equilibrium Statistical Mechanics, Second Edition, John Wiley&Sons, New York

Atkins P. (2007), Four Laws that drive the Universe, Oxford University Press, New York.

Beinhocker E.D. (2007), The Origin of Wealth, Random House Business Books, London

Brillouin L. (1962), Science and Information Theory, Second Edition, Academic Press, New York

Bunch B., A. Hellerman (2004), The History of Science and Technology, Houghton Mifflin Co., New York

Carnot S. (1824), Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance, Bachelier, Paris (French)

Coase R.H. (1937), The Nature of the Firm, Economica New Series 4(16), 385-405

Cockell C., Carfield R. (2008), An Introduction to the Earth-Life System, Cambridge University Press, Cambridge

Dam B. (2009), Duurzame Energie, een materiële kwestie, Intreerede TUD (Dutch)

Darwin C.R. (1859), On the Origin of Species, John Murray, London

Darwin C.R. (1871), The Descent of Man and Selection in Relation to Sex, John Murray, Londen

Dawkins R. (1976), The selfish Gene, Oxford University Press, Oxford

Dawkins R. (1987), The Blind Watchmaker, Norton Co., New York

Dopfer K. (Ed.) (2005), The Evolutionary Foundations of Economics, Cambridge University Press, Cambridge (UK)

Douma S., H. Schreuder (2008), Economic Approaches to Organizations, Fourth Edition, Pearson Education, Harlow (UK)

Eigen M. (1971), Self organization of Matter and the Development of Biological Macromolecules, Die Naturwissenschaften, 58(10), 465-523

Eigen M., P. Schuster (1977), The Hypercycle, A Principle of Natural Self-Organization, Part A: Emergence of the Hypercycle, Die Naturwissenschaften, 64 (11), 541-565

Eigen M., P. Schuster (1978a), The Hypercycle, A principle of Natural Self-Organization, Part B: The abstract Hypercycle, Die Naturwissenschaften 65, 7-41

Eigen M., P. Schuster (1978b), The Hypercycle, A principle of Molecular Self-Organization, Part C: The Realistic Hypercycle, Die Naturwissenschaften, 65(7), 341-369

Elton E.J., M.J. Gruber (1984), Modern Portfolio Theory and Investment Analysis, 2nd Edition, John Wiley & Sons, New York

Fleming A. (1929), On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzæ, Br. J. Exp. Pathol. 10 (31): 226–236

Glansdorff P., I. Prigogine (1971), Thermodynamic Theory of Structure, Stability and Fluctuations, Wiley Interscience, New York

Gleick J. (1988), Chaos: Making a New Science, Penguin, New York

Glendenning N.K. (2004), After the Beginning, Imperial College Press an World Scientific Publishing, London.

Gould S.J. (2004), The Evolution of Life on Earth, Scientific American, October 2004.

Gould S.J. (2001), The Book of Life-An illustrated History of the Evolution of Life on Earth, W.W. Norton & Company, New York.

Hannan M.T. and J. Freeman (1989) Organizational Ecology*,* Harvard University Press, Cambridge, MA
Hellyer, M, (2003), The Scientific Revolution-The Essential Readings, Blackwell Publishing, Oxford.
Hespanha J.P. (2009), Linear Systems Theory, University Press of California. Columbia, CA
Hill T.L. (1960), An introduction to Statistical Thermodynamics, Addison Wesley, Menlo Park (CA)
Hirshleifer J. (1977), Economics form a biological viewpoint, Journal of Law and Economics, 20 (1), 1-52
Huxley T. H. (1863), Evidence as to Man's place in nature, Williams & Norwood, London
Jensen M.C. (1972), Capital Markets: Theory and Evidence, Bell Journal of Economics and Management Science, 3(2), 357-398
Kooreman H.J., J.A. Roels, "Biotechnology". In: EIRMA Workshop Technology '99, EIRMA Conference Papers Vol. 38, p. 235-252 (1989).
Kuhn T.S. (1962), The Structure of Scientific Revolutions, University of Chicago Press, Chicago
Lee M.W. (1955), Economic Fluctuations, Richard D. Irwin, Homewood, IL
Lenski G., P. Nolan (2005). Human Societies: An Introduction to Macrosociology, Paradigm Publishers, Boulder (CO)
Lewin R., and R.A. Foley (2004), Principles of Human Evolution (2nd Edition), Blackwell Publishing, Oxford
Maddison A. (2007), Contours of the World Economy, 1-2030 AD, Oxford University Press, New York
Mayr E. (1985), The Growth of Biological Thought, Blackwell Publishing, Oxford
Markwell D. (2006), John Maynard Keynes and International Relations: Economic Paths to War and Peace, Oxford University Press, New York
Miller S.L. and L.E. Orgel (1974), The Origin of Life on Earth, Prentice Hall, Englewood Cliffs, N.J.
Monod J. (1971), Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology, Alfred A. Knopf, New York
More C. (2000), Understanding the Industrial Revolution, Routledge, New York
Nelson R.R, S.G. Winter (1982), An Evolutionary Theory of Economic Change, Belknap Press of Harvard University Press, London
Nelson R.R. (1987), Understanding Technical Change as an Evolutionary Process, North Holland, Amsterdam
Nicolis G., I. Prigogine (1977), Self-Organization in Non-Equilibrium Systems, J. Wiley Sons, New York
Nicolis G., I. Prigogine (1977a), Self-Organization in Nonequilibrium Systems, Wiley&Sons, New York, p. 459-461
Onsager L. (1931a), Phys. Rev. 37 (4), 405,
Onsager L. (1931b), Phys. Rev.38 (12), 2265
Pauling L. (1960), The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry (Third edition), Cornell University Press, Ithaca, NY
Peace W. (2004), Leslie A. White: Evolution and Revolution in Anthropology , University of Nebraska Press, Lincoln (NE)
Plosser C.I. (1989), Understanding Real Business Cycles, Journal of Economic Perspectives, 3, 51-77
Popper K. (1963), Conjectures and Refutations: The Growth of Scientific Knowledge, Routledge, Londen
Porter M.E. (1981), The Contribution of Industrial Organizations to Strategic Management,

Academy of Management Review, 6, 609-620
Porter M.E. (1980), Competitive Strategy: Techniques for Analyzing Industries and Competitors, Free Press, New York
Porter M.E. (1985), The Competitive Advantage, Free Press, New York
Prahalad C.K. and G. Hamel (1990), The Core Competences of the Corporation, Harvard Business Review, 68 (3), 79-93
Prigogine I. (1980), From Being to Becoming, W.H. Freeman, San Francisco
Prigogine I. and I. Stengers (1984), Order out of Chaos, Bantam Books, Toronto
Roels J.A. (1980), Application of Macroscopic principles to microbial metabolism. Biotechnology and Bioengineering, 22:2457–2514.
Roels J.A. (1983), Energetics and Kinetics in Biotechnology, Elsevier Biomedical Press, Amsterdam
Roels, J.A. (2010), Information Asymmetries and the Creation of Economic Value, IOS Press, Amsterdam
Romer P. (1990), Endogenous Technological Change, Journal of Political Economy 98, S71-S102
Ross S. (2008), The Industrial Revolution, Evans Brothers, London.
Schrödinger, E. (1945), What is Life? Cambridge University Press, New York
Sharpe, W.F. (1970), Portfolio Theory and Capital Markets, Mc Graw-Hill, New York
Shannon, C.E., W. Weaver (1949), The Mathematical Theory of Communication, University of Illinois Press, Urbane (IL)
Simon H.A. (2005), In K. Dopfer (Ed.), The Evolutionary Foundation of Economics, Cambridge University Press, Cambridge (UK), Chapter 4
Smith A. (1776), An Inquiry into the Nature and Causes of the Wealth of Nations, W. Straham and T. Cadell, London
Solow R. (1956), A Contribution to the Theory of Economic Growth, Quarterly Journal of Economics 70, 65-94
Van den Berg M. A., et al. (2008), Genome Sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*, Nature Biotechnology, 26 (10), 1161-1168
Van der Beek C.P., and J.A. Roels (1984), Penicillin Production: biotechnology at its best, Antonie van Leeuwenhoek, 50, 625-639
Watson J.D., F.H.C. Crick (1953), The Structure of DNA, Cold Spring Harbor Symp. Quant. Biol., 18, 123-131.
Wicander, R., J.S. Monroe (2009), Historical Ecology: Evolution of earth and life trough time. Brooks/Cole, Belmont (CA)
Williamson O.E. (1975), Markets and Hierarchies: Analysis and Antitrust Implications, Free Press, New York