Closure as a scientific concept and its application to ecosystem ecology and the science of the biosphere

H. Morowitz a, J.P. Allen b, M. Nelson b,c,* , A. Alling d

a George Mason University, Krasnow Institute, Mail Stop 2A1, Fairfax, VA 22030, USA
b Biosphere Technologies, 7 Silver Hills Road, Santa Fe, NM 87508, USA
c Institute of Ecotechnics, 24 Old Gloucester Street, London WC1 3AL, UK
d Biosphere Foundation, 9 Silver Hills Road, Santa Fe, NM 87508, USA

Received 10 August 2004; received in revised form 22 June 2005; accepted 22 June 2005

Abstract

Closure is a key concept in the physical sciences that has infrequently been used in ecology. The paper reviews closure to the flow of matter and energy (adiabatic walls) and closure to the flow of matter (diathermal walls). A system with rigid adiabatic walls will degrade eventually to chemical equilibrium, a state of maximum entropy. A third type of closure involves semi-permeable walls permitting the flow of one or more types of chemicals. These closure concepts were important to the development of classical thermodynamics and statistical mechanics in the 19th and 20th centuries. “Equilibrium” is often used to describe a time independent steady state. This usage leads to confusion, because equilibrium has such a precise meaning in thermal physics. All living systems are far-from-equilibrium and life cannot persist without the flow of energy. The Earth is an almost materially closed system. Only a small amount of cosmic matter is captured by the Earth’s gravitational field and only a small fraction of lighter elements escape that field. The Earth receives photon flux from the sun and generates thermal energy from the planetary decay of radioisotopes. A hypothesis can be advanced that the planetary biosphere exists in part because of material closure due to gravitation. In the science of ecology partial material closure has been introduced in limnology and island ecology. This has advanced biogeographical theory and systems ecology. The development in the past half century of first balanced aquaria and terrariums, and then partially materially closed microcosms and mesocosms has also greatly aided the development of ecology as an experimental rather than merely descriptive science. All the above systems are open atmospherically, and often have some water and nutrient inputs. The development of truly materially closed man-made systems offers further scope for the development of experimental ecology. The paper reviews and defines the various types of closed ecological systems: Class 1: natural planetary biospheres (like the Earth’s); and Class 2: man-made systems which range from laboratory microbial ecospheres to ones capable of human life support: Controlled Environmental Life Support Systems (CELSS such as are being developed by NASA and the European Space Agency), Closed Ecological Systems (such as Bios-3 at the Institute of Biophysics in Krasnoyarsk, Russia and the Biosphere 2 Test Module) to mini-biospheric systems with a complexity of internal ecosystems (e.g., Biosphere 2 and the Closed Ecology Experimental Facility, CEEF, in Japan).

2005 COSPAR. Published by Elsevier Ltd. All rights reserved.

Keywords: Closure; Biospheres; Closed ecological systems; Materionally closed; Energy flux; Biosphere 2; Diathermal; Adiabatic; Equilibrium; Entropy

1. Introduction: closure as a key scientific construct

Crucial to the development of 18th and 19th century physics was the concept of a system or portion of the universe which formed the focus of study. To formulate laws of physics requires, at the very least, the domain of space we are talking about. Thus, the classical physicist
would say: “Consider a cube of dimensions $d\times dy\times dz$ located at $x, y, z$”. This cube is thus bounded by an imaginary surface, and we study the flow of matter and energy in and out of the system. Such a construct lies at the base of classical hydrodynamics, diffusion theory and the Fourier theory of heat flow (Page, 1935).

The dynamical theories we have just mentioned are inadequate to thermodynamics, which must deal with equilibrium, a notion that requires a greater degree of isolation than an imaginary surface. Barriers or walls are introduced in thermodynamics to isolate the system of interest from its surroundings in a variety of ways (Morowitz, 1978). The most restrictive type of wall is impermeable to the flow of heat and matter and is designated as adiabatic. A second type of wall that allows the flow of heat but not matter is designated as diathermal (Fig. 1).

Classical thermodynamics allows a number of types of closure using these two types of walls, which are either considered to be rigid, or which allows the walls to move, thus introducing pressure-volume work as a component of energy. The most restrictive type of closure is that of surrounding a system by immovable rigid adiabatic walls. Such a system does not allow the flow of matter or energy and does not allow the volume to change. Such a system will approach equilibrium, a state fully determined by the internal energy, $U$, the volume, $V$, and the mole numbers, $n_i$. Strictly speaking, these are the mole numbers of the atomic constituents.

An adiabatically isolated system will decay to equilibrium, or to entropy maximum. The rate at which this happens is outside the domain of thermodynamics and requires more kinetic methods of study. Continuous life is, of course, impossible in an adiabatic system, as energy flow is required to maintain biota.

Consider the following thought experiment. Suppose we were to place a living mouse in a rigid adiabatic system. The animal’s metabolism would use up the oxygen and the animal would die of anoxia. The enzymes of microbes would participate in the breakdown of the rodent. In time the enzymes would break down due to thermal denaturation, and the microbes, even the anaerobic ones, would die due to food depletion and related factors. Even after this happened the system would continue to decay and move toward the chemical equilibrium point as a universal attractor. Eventually, it would consist of CO$_2$, H$_2$O, CH$_4$, N$_2$ and similar small molecules that characterize the equilibrium state for a system of this atomic composition and total energy, which is fixed by this type of isolation. Equilibrium of this type will probably take millennia or longer. In dealing with isothermal systems, we shall see how to speed it up.

Next consider a system surrounded by diathermal walls. To make this problem well-defined, we have to place it in a large, theoretically infinite thermal reservoir. This is a less restrictive type of isolation than the adiabatic case because heat can flow between system and reservoir. A system so isolated is called isothermal, and if the reservoir is time invariant, the system will come to equilibrium in exactly analogous fashion to the adiabatic system.

To return to the previous example: if the mouse is isothermally isolated at the same temperature as the final temperature as in the previous adiabatic case, then the equilibrium states will be the same for both systems. To speed up equilibrium, the system can first be immersed in a very high temperature reservoir and then cooled slowly to the final temperature.

If one of the walls of the system is movable, the interior pressure moves the wall and does what is designated as PV work, removing that amount of energy from the system. For classical thermodynamics, all of the reservoirs are time independent equilibrium systems.

One other type of wall is introduced for purposes of chemical thermodynamics. It is rigid and permits both the flow of heat and one chemical species of molecules. Such a wall is semi-permeable and opens the system to the flow of matter as well as energy. Separate isothermal equilibrium reservoirs of constant chemical potential are needed for such a system, one for each chemical species. This type of wall tends to be an abstraction, but can be approached by materials such as a rigid sheet of palladium–silver alloy, which permits the flow of hydrogen, but nothing else (Fig. 1).

The three types of walls – adiabatic, diathermal and semi-permeable – form the basis of the micro canonical ensemble, the canonical ensemble, and the grand canonical ensemble of the statistical mechanics as formulated by Josiah Gibbs (1988) (Fig. 1). Closure is deeply related to the formulations of thermal physics, a point that is insufficiently stressed in developing the role of closure as a major construct in related sciences.

![Fig. 1. Types of closure: (a) complete closure-system closed to the flow of energy and matter; (b) system open to the flow of energy, but closed materially; (c) system open to the flow of energy and matter. $T$ represents energy, and the symbol $\mu$ (chemical potential) indicates open to the flow of matter; (d) closed ecological systems: open to information ($I$) and energy ($T$) but essentially closed to material exchange.](image-url)
2. Equilibrium systems and steady state systems

Next consider what is usually called irreversible thermodynamics, which in its present embodiment is near to equilibrium thermodynamics (Morowitz, 1978). For this kind of analysis, a system must be minimally in contact with two reservoirs, a source and a sink, so that the steady state involves a flow of matter or energy. The simplest example of such a system is one in contact with two thermal reservoirs of temperatures \( T_1 \) and \( T_2 \) through diathermal walls. Suppose the system contains gaseous neon. When the steady state is achieved there is a temperature gradient, a concentration gradient, and a steady flow of energy between the two reservoirs.

The difference between steady state equilibrium states should be noted. Both involve time independent variables, but the former involves a flux through the systems and multiple reservoirs. In the case noted, there is a constant entropy generation due to the flow from hotter to colder reservoirs. The term equilibrium is used in many fields to describe time independent steady states. This usage is confusing because equilibrium has such a precise meaning in classical thermodynamics and statistical mechanics. We would suggest that the word equilibrium be restricted to those fields, and steady state applied in other cases.

The same type of closure used to describe near to equilibrium systems can be extended to far from equilibrium systems, but the treatments become much more complex, and in most cases they are at present insoluble. Consider, for example, a system closed to the flow of matter but open to the flow of energy, having one reservoir at 5750 K and the second at 3 K. Consider that the internal system starts with a mixture of \( \text{H}_2\text{O} \), \( \text{CO}_2 \) and \( \text{N}_2 \). As the system ages, there will be a temperature and concentration gradient between the two reservoirs but the chemical description becomes very complex. The high temperature will lead to free radical formation, and the short wavelength end of the 5750 K black body spectrum will drive many photochemical and radiochemical reactions. Thus all molecules of CHNO are possible in such a system.

Does the system described above have a unique steady state? The answer is probably no, but is at present unknown. For an equilibrium system, the singularity of the final state occurs as one of the postulates of the theory; indeed, it lies at the root of the idea of internal energy, being a state function. No such postulate exists for very far from equilibrium systems and the required kinetic methods are very complex. Nevertheless, the system just described is a model of the Sun–Earth system and therefore of great importance in thinking about the biosphere. Considering what is being modelled, it seems very unlikely that the relatively simple system discussed above has a unique steady state.

3. The Earth as a materially closed system

Consider the Earth. To a first order approximation, it is closed to the flow of matter, although meteors, comets and planetary debris arrive episodically, and hydrogen and helium escape (“leak”) from the upper atmosphere. In recent years an occasional satellite is also lost from the Earth’s gravitational field. There is a constant diurnal inflow of solar energy, plus the thermal energy from radioactive decay in the Earth’s interior, and a constant outflow of energy to the 3 K cold of outer space. It is this energy flux, with a small additional component driven by reduced materials released at volcanoes and subduction trenches, which drives the present day biosphere. Thus, within the limits of the small mass fluxes, the Earth may be regarded as an essentially closed system, meaning it is closed to the flow of matter and open to the flow of energy (Fig. 2). Open systems are open to the flux of matter and energy from both equilibrium or non-equilibrium sources and sinks. Open systems with one restriction on closure may have certain interesting properties as the following two biological examples demonstrate. In microbiology, one often starts an experiment with a flask of sterile nutrient medium stoppered with a cotton plug. One then introduces a pure culture or cultures into the flask under sterile conditions and places the flask on a shaker in an incubator. Such a system is isothermal, open to exchange of atmospheric gas or gases produced in the system, but closed to the flow of biota. Much of current microbiology depends on this type of “gnotobiotic closure”.

A somewhat different type of restricted closure is a semi-permeable membrane such as is used in osmotic pressure experiments. Here the closure is at a molecular level but nevertheless causes a major macroscopic difference in the pressure of the two sides. The passage of solvent, but not solute, through the membrane is clearly a type of closure.

Fig. 2. Illustrations of closed ecological systems, closed materially and open to energy and information. These living systems are complex far-from-equilibrium systems, capable of self-organization and evolution: (a) Biosphere 2, Oracle, Arizona, the largest man-made closed ecological system research facility yet constructed; (b) planet Earth's biosphere is the only presently known example of a naturally closed ecological system.
The Maxwell Demon in the mid-1800s introduced a new type of closure (Loff and Rex, 1990). The “demon” controlled a door between two closed systems, opening it and closing it to control flow of molecules between the two sides. This operation can lead to the temperature or pressure differences between the two compartments. In modern terminology, this introduces the notion of the flow of information and closure with respect to information flow. Thus a system may have valves, switches and robot arms which may in principle be controlled from the outside. This involves signals which must be either matter or energy, but the material or energy flows can be made vanishingly small with respect to the flows controlled by the information.

The biospheric far-from-equilibrium closure involves starting with a biotic system, one containing living organisms, closing it to the flow of matter, leaving it open to the flow of energy from source to sink and the flow of information, and then letting the system develop in time.

Precursors to such systems were balanced aquaria and terrariums, although these systems were open to the atmosphere with gas exchange. As objects of ecological interest, it is important to maintain these systems over long time periods. Materially closed systems involving living biota have been designated as natural biospheres (e.g., Earth’s) and man-made systems which range from laboratory ecospheres, controlled and closed ecological life support systems and man-made biospheres.

4. The value of limited closure, e.g., lakes, islands, watersheds and mesocosms, to ecological science

Closure and clearly elucidated boundaries have been an important if largely implicit concept in the advancement of ecological theory. Much progress in ecology resulted from the work of limnologists, such as Hutchinson, who found that the study of lakes lends itself to the detailed examination of internal processes of its natural boundaries. Indeed Hutchinson refers to lakes as “more or less closed systems” (Hutchinson, 1957) as Forbes had earlier studied them as “microcosms” (Forbes, 1887). In a similar fashion, the ecological studies of islands, important in the development of biogeographical theory, evolutionary radiation and recolonization processes (MacArthur and Wilson, 1967) is facilitated because of their separation from interaction with other ecosystems.

Field enclosures or exclosures bounded by some kind of fence or other barrier, and where input and outputs are controlled and reduced by not eliminated entirely are known as mesocosms (Odum, 1984). Since they are subjected to the natural environment, are big enough to contain macro organisms, and can be replicated, such mesocosms provide excellent experimental systems for testing the effects on the entire ecosystem of toxic substances, other perturbations, or the addition or removal of species. One type of specific closure used in ecology of open systems is exclosure. An example from Volcanoes National Park on the island of Hawaii illustrates this type of restricted closure. In a fairly high rainfall region, a wire fence was placed around a large area and it was ascertained that no feral mammals were in the fenced area. The fence is open to flow of matter and energy; the open top also permits such fluxes as airborne particles and birds as well as atmospheric exchange. In a few years the area inside the fence has an ecosystem description totally different from its surroundings. It appears to be in a successional stage towards becoming a forest. The only things kept out of the fence were large feral grazers especially goats and cattle. This modest biotic closure totally alters the ecosystem.

The creation of ecological microcosms and mesocosms provided new tools for the ecologists to study fundamental ecosystem processes in miniature (Bormann and Odum, 1993). Many of these systems exhibit a high degree of internal, material recycling while they are open to atmospheric interactions, loss of water through evapotranspiration, and some minute inputs of nutrients. Indeed, this trend of utilizing discrete boundaries is used by system ecologists to elucidate the properties of ecosystems (Odum, 1971, 1983).

A powerful long-term study using these methods has focussed on the energetic and material inputs and outputs of a watershed, Hubbard Brook (Bormann and Likens, 1979) which was selected because its geological substrate is impervious to percolation of groundwater. The Hubbard Brook Ecosystem Study examined how this small watershed worked, providing information on ecosystem dynamics, and included experimental manipulations of ecosystems (McIntosh, 1985). Long-term experimental watershed studies at Coweeta in North Carolina are noteworthy for the comparisons of discrete watersheds that are subjected to differing forestry and land-use practices (Swank and Crossley, 1988).

Thus the development in the past thirty years of new types of man-made, experimental closed biotic and ecological systems continues a tradition deeply imbedded in the history of ecological science. These vary in size, degree of material closure and in their ecological complexity (Fig. 2). They all, however, are artificial and therefore can be manipulated and therefore can be manipulated experimentally. They approximate a diathermal closure with a small material leak and permit a flux of energy and information with their surroundings (Cooke, 1971; Shepelev, 1972; Nelson, 1997). These new types of objects offer another avenue of approach for the elucidation of fundamental ecological processes.
5. Very-far-from-equilibrium closed systems

5.1. Class 1: natural biospheres, Earth

The Earth’s biosphere is a naturally materially closed system which is informationally and energetically open to the cosmos. The biosphere is driven largely by sunlight and receives most of its energy as radiant energy, although some energy is produced geothermally from the Earth’s interior and this heat is important in driving plate tectonics and other geological processes. The biosphere is thus not isolated from the solar system. However, with the exception of cosmic debris that falls into our atmosphere and of hydrogen and other light gases that escape from our upper atmosphere to outer space, both of which are almost negligible quantities compared with the Earth’s mass, the biosphere is materially closed.

This material closure of a planet may be an essential factor in allowing a biosphere to evolve. It is believed that the early Earth lost most of its light gases, especially hydrogen and helium, during its period of coalescence. During that period the Earth also received significant material inputs from the large pool of meteors, comets from the young Solar System, and other materials which contributed significant quantities of elements now considered vital to Earth life. It was crucial for the evolution of the early biosphere that Earth’s gravitational field (that ensures its tight material closure) was able to hold the H₂O and CO₂ outgassed by the intense volcanic activity of the young Earth. From the materials outgassed, the Earth was able to constitute its first atmosphere and later, after cooling of its surface, precipitate water to form the oceans. The ability of a planet to hold sufficient atmosphere may well be a crucial factor in whether a planetary biosphere can evolve on that planet.

One of the scientific constraints to more rapid progress in understanding how our Earth’s biosphere functions is that it cannot be manipulated as an experimental system. One of the important opportunities that artificial, closed ecological systems makes available is the possibility of a range of experiments on the effects of biospheric conditions.

5.2. Class 2: experimental closed ecological systems

5.2.1. Laboratory ecospheres

The first class includes systems termed “materially closed ecospheres”. These are small systems which cannot support humans or indeed large vertebrates or mammals, but which are materially closed, and energetically and informationally open. Examples of such systems are 100 ml to 5 litre flasks that contain different communities of aquatic or marine communities which can be sealed off from our atmosphere and which are energetically driven by inputs of indirect sunlight or artificial light. In addition, they can be observed and monitored which facilitates informational exchange. With adequate light input and microbial diversity, these ecospheres continue to persist and self-organize. Such laboratory ecospheres were pioneered by Folsome at the University of Hawaii, and subsequently by Taub at the University of Washington, Maguire at the University of Texas and Hanson at the Jet Propulsion Laboratory, and others (Folsome and Hanson, 1986; Hanson, 1982). It is important to distinguish between two types of laboratory microcosm: (1) the defined type, seeded from pure cultures and (2) the derived type, multiple-seeded from naturally diverse environmental samples.

5.2.2. Controlled environmental life support systems

The second class of engineered (artificial) systems are those that can support humans and require sophisticated technology to maintain and operate. These are the Controlled Environmental Life Support Systems (CELSS). These are apparatus that are designed to at least partially biologically regenerate the air, water and food for humans. They provide an experimental setting useful in the study of simple agricultural systems and individual crops, and have been primarily used for growing algae and/or food crops. One of the largest such facility is the 72 m³ Breadboard Plant Growth Chamber at NASA’s Kennedy Space Center (Wheeler et al., 1996) and the more recent Bio-Plex at NASA’s Johnson Space Center (Pickering and Edeen, 1998). European efforts in CELSS include the MELISSA project at the University of Barcelona and at ESTC in Noordwijk, the Netherlands. Research in CELSS is motivated by the attempt to provide bioregenerative life support to lessen the requirements of stored supplies and resupply from Earth. In most envisioned space applications of CELSS technology, only part of the required food, air and water will be bioregenerated, the rest will be supplemented by physico-chemical technologies, storage of supplies and export of wastes. If these systems were able to achieve complete bioregeneration, i.e., complete material closure, they would join the third type of man-made closed ecological system.

5.2.3. Closed ecological life support systems

Closed ecological systems are capable of supplying and regenerating the air, food, water and waste required for human life support by recycling metabolic wastes and completing essential nutrient cycles. The first such system was the 315 m³ Bios-3 facility at the Institute of Biophysics, Krasnoyarsk, Russia. This system operated for a series of experiments primarily from 1970 to 1982, and supported 2–3 crew members for experiments of as long as 4–6 months (Terskov et al., 1979; Gitelson et al., 2003). While some food was imported, and solid wastes exported, the degree of closure approached that of total material closure. The second closed ecological life support system was the 480 m³ Biosphere 2 Test Module (Alling et al., 1996) and the more recent Bio-Plex at NASA’s Johnson Space Center (Pickering and Edeen, 1998).
It was designed to be able to fully support one human for up to two months, although the longest closure experiment was 21 days in 1989. Unlike Bios-3, which was powered by artificial lights; the Biosphere 2 Test Module was open to sunlight. Both systems relied on electricity for mechanical work/computer control and had sophisticated technologies for waste and air purification: a catalytic converter was used in Bios-3; while the Biosphere 2 Test Module used soil bed reactors (Frye and Hodges, 1990) and a constructed wetland for wastewater regeneration (Nelson, 1998). A third Closed Ecological Life Support System was constructed in 2000 in Santa Fe, New Mexico at the headquarters of the Biosphere Consortium. This “Laboratory Biosphere” is a 1200–1400 ft³ steel container with a variable expansion chamber. It has a 5.5 m² (59 ft²) soil planting bed, artificial lighting, and a water recycling system (Dempster et al., 2004).

5.2.4. Man-made biospheres

The fourth class of artificial closed ecological systems are man-made “biospheres”. They differ from closed ecological life support systems in that they contain more than one type of internal ecosystem (analogous to Earth’s biomes) and may thus contain sufficient ecological complexity for long-term persistence. Biosphere 2, in Oracle, Arizona, was the first attempt to create such a biospheric closed ecological system (Allen, 1991; Alling et al., 1993; Nelson and Dempster, 1996) (Fig. 2). The Closed Ecology Experiment Facility (CEEF) in Rokkasho, a small township in prefecture Aomori, Japan, currently being completed is the second such system (Nitta, 2001).

Biosphere 2 and CEEF illustrate that differing approaches may be used in such mini-biospheric systems. Biosphere 2 was designed for support of a crew of eight people, with an intensive agricultural area and human habitat. In addition, ecosystems patterned on five major tropical biomes were included to increase ecological complexity and buffering capacity; and to enhance the use of the facility as a laboratory for global ecology (Allen, 2001). The atmosphere and water cycles are shared between the anthropogenic and “wilderness” biomes. The facility was species-packed, to allow for the process of ecological self-organization. The overall model was “top-down” although a great deal of detailed reductionist science and engineering was required to make the facility extremely air-tight (less than 10% annual air exchange) and to complement natural ecological functions. CEEF is designed to support two people at maximum and while it contains a number of internal components (e.g., food production; domestic animal production; an oceanic biome; living quarters and waste treatment technologies); the system is designed to carefully measure the inputs and outputs from each component.

6. Conclusion

The advent of these first closed ecological systems and man-made biospheric systems opens up a wealth of potential avenues of ecological research. These areas include biogeochemical cycles, portions of which are accelerated by the small reservoir sizes and the biotic density of such systems; the dynamics of small populations; techniques of ecosystem restoration; the response of organisms and biotic communities to environmental parameters which differ from those found in their natural range; the sustainability of agricultural and other human technological practices; and the development of biological recycling and purification technologies. In addition, the inherent integration in human life support closed ecological systems of natural, ecological processes and technological ones designed to assist/augment and sometimes replace natural mechanisms (e.g., mechanical wave generators for the Biosphere 2 ocean and coral reef; nutrient scrubbers and heat exchange/wind generators) are highly relevant as testbeds and models for the intelligent integration of technology with our global biosphere (e.g., Kelly, 1994). Closed ecological systems offer a range of powerful new tools to make ecology a truly experimental science. Such systems will be necessary for the ultimate habitation of other planets.

References

H. Morowitz et al. / Advances in Space Research 36 (2005) 1305–1311

Hanson, J. Workshop on closed system ecology Summary report. Jet Propulsion Laboratory Publication #82-64, Pasadena, CA, 1982.