1

Thermodynamics and the Earth system

1.1 A thermodynamic basis for Earth system science

The Earth is a vastly complex system. This complexity is reflected in the broad range of processes that it entails, from the solar radiative forcing to the highly dynamic circulatory patterns in the atmosphere, ocean and the interior, to high level and diversity of metabolic activity of life, and to human activities. The complexity is further enhanced by strong interactions by which processes alter their own drivers. Atmospheric motion, for instance, transports such vast amounts of heat that it alters the radiative exchange with space. The activity of the Earth's biosphere, the sum of all living organisms, has strongly altered the chemical composition of the atmosphere, as for instance reflected in its high abundance of molecular oxygen, resulting in altered physical and chemical conditions. And finally, human activity over the last century has released such large amounts of buried organic carbon by its industrial activities that it has substantially altered the global carbon cycle resulting in enhanced concentrations of carbon dioxide in the atmosphere and global climate change. With such complexities in mind, it would seem almost impossible to make robust predictions of magnitudes, the strength of interactions, and the overall evolutionary direction of the Earth system as a whole in order to get a robust, physical understanding of how the whole Earth system functions and responds to change.

Yet there is a range of fundamental, practical, and relevant questions that require such a robust understanding. What determines, for instance, the strength of the atmospheric circulation and its ability to transport and mix heat and mass? The answer to this question would help us to make better predictions of the magnitude of climate system processes and how these would respond to perturbations and change. Does the climate system, and the planet as a whole, regulate its climatic state to some particular reference level? Is climate even regulated to a point that is most suitable to life, because of the presence of life, as proposed by the Gaia hypothesis (Lovelock 1972b,a; Lovelock and Margulis 1974)? If this is so, how would human activity play into such a planetary regulation? A better understanding of these questions would provide information about the role of the biosphere at the planetary scale and the factors that shape planetary habitability. What are the limits to human activity, for instance, limits related to food production and the availability of renewable energy and how do these relate to the functioning of the planet? What are the associated human impacts on the system, and can these impacts tip the planet off some edge into an inhabitable state? The answer to this question would help us to develop better scenarios of a sustainable future. What these questions have in common is that they require a perspective on the whole Earth system, as the limits that these questions involve are of a physical nature, and ultimately relate to the planetary forcing represented by solar radiation and the cooling of the interior Earth.

The goal of this book is to provide a fundamental basis rooted in physics that allows to approach these questions. This book will show that a central component for this basis is described by thermodynamics, a fundamental theory in physics that deals with conversions of energy and their direction towards states of higher entropy. It is particularly the latter aspect, known as the second law of thermodynamics, that has intrigued many scientists, that provides a fundamental direction for processes, and that sets fundamental limits. Energy and entropy are very basic quantities that apply to practically all processes, from radiation to metabolic and human activity. Thermodynamics thus provides a way to formulate all Earth system processes in comparable quantities, thus providing a general accounting basis.

Yet, it is not just the thermodynamic formulation that is important. Equally important is to place the thermodynamic formulation into a systems perspective of the whole Earth system. This combination allows us to link processes to their ultimate driver of solar radiation and interior heat through sequences of conversions. These conversions are associated with converting energy of different forms, and the laws of thermodynamics constrain these conversions by thermodynamic limits. When thermodynamic limits are then applied to these sequences, we can quantify the limits on their rates and on the interactions that result from these processes. As we will see in the book, these limits provide basic and robust estimates for a range of Earth system processes that compare well with observations. This is not because these processes are organized in a simple, predictable way, but rather likely because they are so complex and have evolved so far that they operate near their thermodynamic limit. Even if not all processes may necessarily operate near their limit, it nevertheless provides us with an evolutionary "target" that can be used to interpret evolutionary dynamics, and it can be used to yield relatively simple and transparent estimates for the magnitude of processes that typically require only a mere minimum of empirical knowledge. This view may then already provide sufficient information to understand Earth system processes, how they respond to change, how life and human activity fit in and what a sustainable future may look like.

In the remaining part of this chapter, the basics of thermodynamics are described qualitatively, with the details being described in Chapter 3. It is illustrated how systems are maintained in an ordered state by fluxes of different entropy going through the system, and how this relates to living organisms and the Earth system as a whole. It is then explained how these basic components of thermodynamics result in limits to the ability to perform work, and how this work feeds sequences of energy conversions of different processes. At the end of this chapter, an overall view of the thermodynamic Earth system is given and it is described how this view is partitioned into the different chapters of the book.

1.2 Thermodynamics in a nutshell

Thermodynamics sets the rules for the conversions of energy from one form into another and sets the general direction into which these conversions take place. These two aspects are described by the first and second law of thermodynamics. The first law essentially states the conservation of energy. When energy is converted from one form into another, overall no energy is lost or gained. As energy is converted, its concentration, or reversely, its dispersal is altered. The extent of energy dispersal during a conversion is described by entropy, with more dispersed forms of energy corresponding to a higher entropy. The second law states that energy is, overall, increasingly being dispersed. This dispersal of energy is, for instance, reflected when a heated object such as a hot cup of coffee cools down and approaches the temperature of its surroundings. Here, the first law would tell us that the heat given off by the object is added to the heat content of the surroundings so that the total energy of the object and its surroundings is conserved. The dispersal of heat that is associated with the cooling of the object, however, is not captured by the first law. This tendency is rather the manifestation of a profound direction of nature to spread energy, mass, and other physical attributes into uniform concentrations. Such states of uniformity are described in thermodynamics as states of maximum entropy, or thermodynamic equilibrium. The natural direction of processes to spread energy and to increase entropy is described by the second law of thermodynamics.

These two laws of thermodynamics, the conservation of energy, and the increase in entropy, are so general, that Albert Einstein once said that (Klein 1967):

> "[a] theory is more impressive the greater the simplicity of its premises, the more different are the kinds of things it relates, and the more extended its range of applicability. Therefore, the deep impression which classical thermodynamics made on me. It is the only physical

theory of universal content, which I am convinced, that within the framework of applicability of its basic concepts will never be overthrown."

The second law in particular sets such a profound direction for physical processes that it has been labeled the "arrow of time" (Eddington 1928). The increase in entropy tells us something quite specific about how we would expect the dynamics within systems to take shape. No matter how complicated a system is, how large it is, or how many types of processes and constituents it involves, the overall dynamics that take place within the system must obey the laws of thermodynamics. Overall, energy needs to be conserved, and entropy needs to increase. How this increase in entropy is accomplished within a system is non-trivial as it also needs to consider how exchange fluxes across the system alter the entropy of the system.

One important characteristic of Earth system processes is that they typically operate far from thermodynamic equilibrium as gradients and fluxes are maintained within the system. This disequilibrium can be maintained in a steady state, in which the mean properties of the system do not change in time, without violating the laws of thermodynamics. The actual formulation of these laws is somewhat different and needs to account for the exchanges between the system and its surroundings across the system boundaries. These exchanges do not only exchange energy, mass, or other physical quantities, they also exchange entropy. When energy is added to warm places, and the same amount of energy is removed from cold places, the total energy within the system does not change, but a gradient in temperature is being maintained, reflecting disequilibrium. As this energy was added and removed at different temperatures, the exchange of entropy does not cancel out, but results in a net export of the entropy that is produced within the system so that the system can be maintained in a state of thermodynamic disequilibrium. In the application of the second law, this exchange of entropy needs to be taken into account. This exchange with the surroundings has important consequences for the state of a system: It allows a system to maintain a state away from thermodynamic equilibrium, and the entropy exchanges across the boundary reflects important information on the extent to which a system is maintained in a state of thermodynamic disequilibrium.

To illustrate this critical point in more detail, let us consider the two systems shown in Fig. 1.1. The top row of this figure shows a "system A" in which no exchange with its surroundings takes place. In such a setting, an initial internal difference in temperature would fade in time. Expressed differently, the processes within the system are directed to deplete this temperature difference. The total energy of the system during this redistribution of heat remains unchanged, but its distribution has changed. This latter aspect reflects the increase of entropy within the system. In the final state of a uniform temperature distribution, energy is distributed most uniformly within the system, the entropy is at a maximum, and this state corresponds to a state of thermodynamic equilibrium. This situation is akin

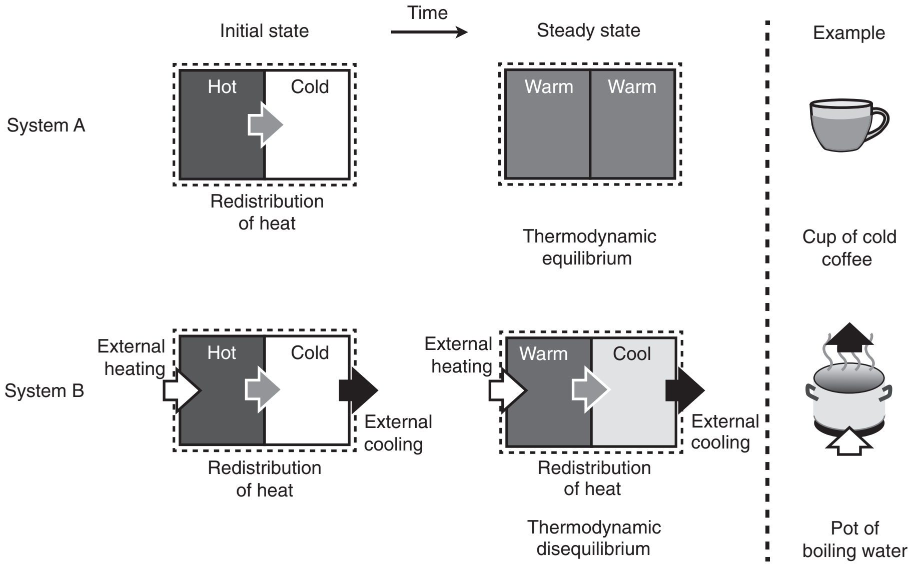

Figure 1.1 Two different types of systems (A: an "isolated" system; B: a "non-isolated" system) develop from an initial state to a final, steady state. An everyday example for such systems are given on the right.

to a hot cup of coffee in a room that would cool down to the temperature of the surroundings.

The bottom of Fig. 1.1 shows a "system B" in which heat is added to one side of the system, while it is cooled at the other side of the system. This setting is comparable to a pot of water on a stovetop that is heated from below, and cooled from above. Just as in the situation described earlier, the processes within the system are directed to deplete the temperature difference within the system, attempting to spread the heat uniformly within the system. However, since energy is continuously added and removed at different parts of the system boundary, the system is maintained in a state of disequilibrium. This disequilibrium manifests itself in the temperature difference that is being maintained within the system and that we can observe, but is also reflected in the dynamics that take place in the system. In the case of the pot of water, these dynamics are simply the convecting motion of boiling water within the pot.

To sum up, the two systems shown in Fig. 1.1 may look the same in terms of the amounts of heat that they contain, but they differ significantly in terms of the internal dynamics and their thermodynamic state. System A describes a system in which the final, steady state is a static state of thermodynamic equilibrium. The properties of the system do not change in time, there is no exchange in energy or entropy, and the system does not show any dynamics. System B also reaches a steady state in which the properties of the system do not change in time. However, its steady state is characterized by disequilibrium and reflects dynamics associated with the heat flux within the system. These dynamics are maintained by exchange fluxes, and the trend to deplete the temperature gradient within the system is mirrored by the entropy exchange of the system with its surroundings. This latter aspect, that the dynamics within the system are directed to deplete gradients, is not a consequence of energy conservation, but of the second law of thermodynamics. This steady state in which fluxes and gradients are being maintained in system B reflect thermodynamic disequilibrium and this state is being maintained by the entropy exchange of the system with its surroundings.

These considerations apply to the Earth system as well. As the Earth is a thermodynamic system that is maintained in a state with fluxes and gradients like system B, this thermodynamic view suggests that the dynamics that we can observe within the Earth system result as a consequence of the second law as well. A necessary foundation to implement these considerations is the formulation of the dynamics entirely in terms of energy and entropy exchange. This is not just captured by the Earth's energy balance, which is the common starting point in climatology. It also requires that all other processes are described in energetic terms, ranging from atmospheric motion to geochemical cycling, biotic and human activity. Furthermore, it requires a description of the entropy fluxes that are associated with these dynamics, which are rarely considered in Earth system science, yet central if we want to interpret the dynamics of the Earth system in terms of the second law.

1.3 Disequilibrium, life, and Earth

Entropy considerations are central when we want to understand how disequilibrium is being maintained and how highly complex phenomena such as life or the Earth system as a whole do not violate the second law, at each and every process and at the scale of the whole planet. In his seminal book on "What is life," Erwin Schrödinger (1944) described that a living organism satisfies the second law by consuming low-entropy food and producing high-entropy waste (Fig. 1.2a). Averaged over some time period, the food uptake by the organism roughly balances its waste, so that there is no net gain or loss of mass. The mass flux in itself does not contain the relevant information that would tell us that the organism is alive. That the influx of mass balances the outflux of mass is simply a consequence of the overall conservation of mass. The relevant information comes from the fact that the influx of mass is of a different constitution than the outflux. It is this difference in constitution that is captured by entropy, and it is this difference in entropy in the exchange of the living organism with its environment that allows the living organism to extract energy to run its metabolism. This metabolic energy is then dissipated and released as heat.

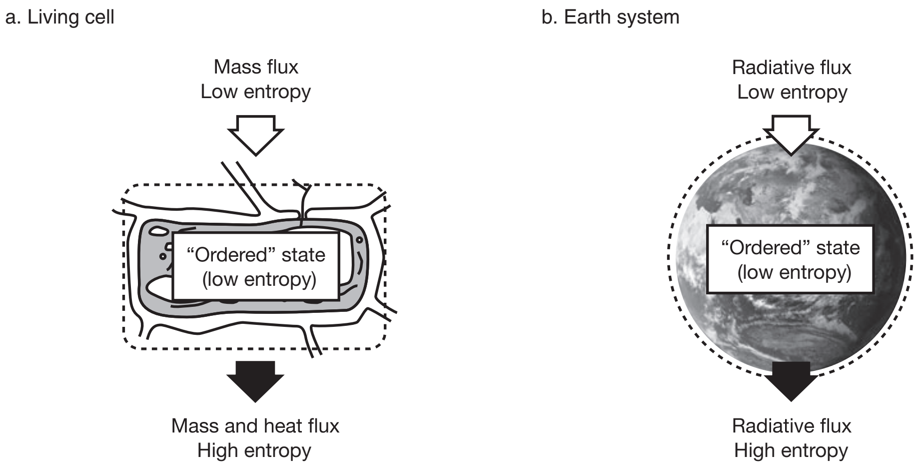

Figure 1.2 (a) A living cell and (b) the whole Earth system as examples of dissipative systems that are maintained far from thermodynamic equilibrium by the exchange of entropy. Source of Earth image: NASA.

A living organism is just one example of a so-called dissipative structure, a term introduced by Ilja Prigogine, a chemist who extensively worked on non-equilibrium thermodynamics and structure formation. His many accomplishments include textbooks on non-equilibrium thermodynamics (Prigogine 1962; Kondepudi and Prigogine 1998), and his major contributions are summarized in an article published on the occasion of the Nobel prize in Prigogine (1978). There are many more examples for such dissipative structures, ranging from the patterns that form in chemical reactions, convection cells in fluids, living organisms, and, ultimately, the whole Earth system. These structures have in common that they maintain states far away from thermodynamic equilibrium, and that these states are being maintained by the entropy exchange with the surroundings, just like the simple system B that was depicted in Fig. 1.1.

In the Earth system (Fig. 1.2b), the entropy exchange of all dissipative structures is ultimately integrated at the planetary scale to the entropy exchange with space. Just like a living organism is maintained by the entropy exchange associated with the fluxes of "food" and "waste," so is the activity of the whole Earth system being maintained by entropy exchange. The similarity between a living cell and the whole Earth as being dissipative structures was noted by James Lovelock and Lynn Margulis. It was this thermodynamic consideration that led Lovelock and Margulis to view the Earth system as a "superorganism" and to formulate the controversial Gaia hypothesis that compared the functioning of the Earth system to a living organism (Lovelock and Margulis 1974).

For the Earth system, the vast majority of entropy exchange is accomplished by the radiative exchange with space (Fig. 1.2b). The "food" of the Earth system is solar radiation, which has a low entropy because the radiative energy is composed of relatively short wavelengths, an aspect we get back to in Chapter 6. The "waste" of the Earth system is exported by the radiation emitted to space, which is radiation of relatively long wavelength and thus has a high entropy. As in the case of the living cell, we deal with a system in which the planetary energy balance is balanced, so that the total radiative energy absorbed by the Earth system roughly balances the total radiative energy emitted to space. The relevant difference is not contained in the energy fluxes, but rather in the associated fluxes of planetary entropy exchange. It is this planetary entropy exchange that allows the maintenance of thermodynamic disequilibrium as well as all the dissipative structures and activities that take place on Earth. Such a system with dissipative structures and net entropy exchange we refer to as a dissipative system, and to the processes that take place within the system and cause an increase in entropy as dissipative activity.

While living organisms and the Earth system both constitute dissipative systems, they do not act independently. In a system's view, the dissipative structures of living organisms feed on the boundary conditions set by the Earth system, yet the products of living organisms will affect their surroundings, and thereby alter the planetary system. This connection between life and the Earth system was formulated very nicely by Ludwig Boltzmann (1844–1906), a physicist who set much of the statistical foundation of thermodynamics. He expressed this connection as (Boltzmann 1886):

"The general struggle for existence of living organisms is therefore not a struggle for the basic materials – these materials are abundantly available for organisms in air, water and soil – nor for energy, which is abundant in form of heat in any body, albeit unfortunately unavailable, but a struggle for entropy, which through the transformations of energy from the hot sun to the cold Earth becomes available. To fully explore this transformation, plants spread their leaves in unimaginable extent and force the solar energy to perform chemical synthesis in yet unexplored ways before this energy sinks to the temperatures of the Earth's surface. The products of this chemical kitchen forms the object of struggle for the animate world."

Note that Boltzmann formulated this relationship before much of the details involved in photosynthesis were discovered. This description of living organisms places biotic activity into a planetary context. Yet, before we describe these interdependencies and their implications in greater detail, we first explore what else thermodynamics can tell us about dynamics, interactions, and evolution.

1.4 Thermodynamic limits

So far, we described the first and second law of thermodynamics and how these provide a constraint and a direction for Earth system processes. When these two

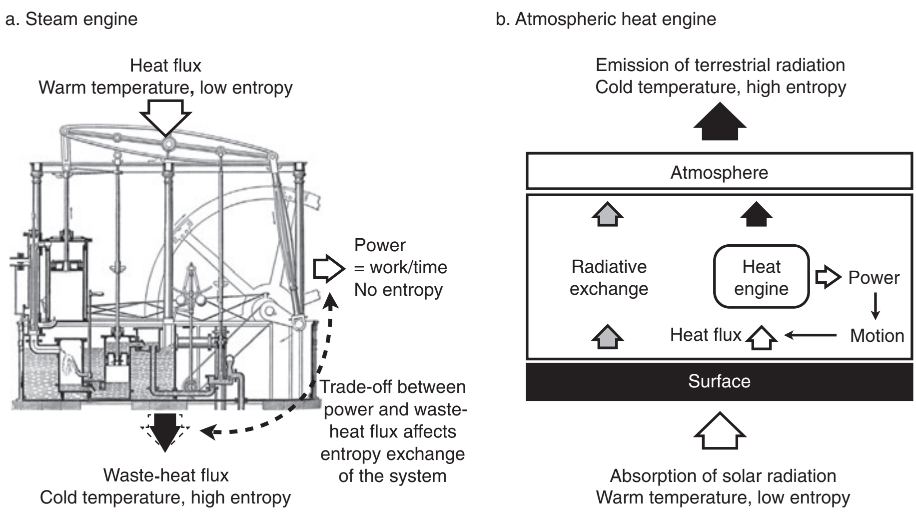

Figure 1.3 (a) Illustration of a heat engine using Watt's original steam engine as a template. Only a fraction of the heat flux into the engine can be converted to mechanical power because of the condition imposed by the second law. (b) An illustration of the surface-atmosphere system in which a heat engine generates motion from radiative heating of the surface and the cooling of the atmosphere. Steam engine image from Meyer (1886), available on www.wikipedia.org.

laws are being combined, they yield a fundamental limit on how much work can be performed, or, more generally, how much energy can be converted from one form into another. This limit constrains, for instance, how much of the radiative heating of the Earth's surface can be converted into the kinetic energy associated with atmospheric motion. The two laws thus set firm constraints on the magnitude of the dynamics that can take place within the Earth system as a result of the planetary forcing.

The limit for converting heat into mechanical work is known as the Carnot limit, named after the engineer Sadi Carnot (1796–1832), who was one of the pioneers in developing thermodynamics at the times at which steam engines were invented. The limit is illustrated in the following using the steam engine shown in Fig. 1.3a. This steam engine operates by the addition of heat at a high temperature, typically by the combustion of a fuel. This is shown in the figure by the white arrow at the top. The first law states that this addition of heat is balanced by the removal of "waste" heat through its exhaust at a colder temperature, shown by the black arrow at the bottom, and by the mechanical work (the white arrow on the right). The second law requires that the engine cannot decrease the overall entropy, but that it at best remains unchanged. This leads to a constraint on the entropy exchange of the engine, which is accomplished by the heat exchange at different temperatures.

The waste-heat flux plays a critical role, as it exports heat at a colder temperature and exports higher entropy to the surroundings of the engine. When the engine performs work, this must come at the expense of a reduced waste-heat flux. This follows from the first law. With a reduced waste-heat flux, less entropy is being exported, so that the net entropy exchange of the engine decreases the more work is being performed by the engine. The Carnot limit marks the point at which there is no net entropy exchange by the system, which is the absolute limit that is permitted by the second law. The work performed by the system then results in the generation of motion, that is, kinetic energy, or the lifting of mass against gravity, that is, potential energy. Hence, the Carnot limit describes the limit to energy conversions from heat to another form.

As already imagined by Carnot, this limit does not just apply to steam engines, but to "natural" engines that drive Earth system processes. In his book "Reflections on the motive power of fire" (Carnot 1824), he writes that:

> "The vast movements which take place on the Earth are ... due to heat. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors, the currents of water which channel the surface of the globe, and of which man has thus far employed but a small portion. Even earthquakes and volcanic eruptions are the result of heat."

Carnot's writing suggests already back in the early nineteenth century that one should view these Earth system processes as if these operate like steam engines, or, more generally, heat engines, hypothetical devices that convert thermal energy into physical work. These engines would be subjected to the same Carnot limit as a heat engine in how much work these could perform and thus result in the dynamics that we can observe.

When formulating thermodynamic limits for Earth system processes, there are a few critical differences to the setup of a Carnot heat engine that need to be considered. An example of this application to atmospheric convection is shown in Fig. 1.3b. Firstly, the generation of atmospheric motion can be viewed as the result of an atmospheric heat engine that is constrained by the heating and cooling associated with radiative exchange. The setup of the Carnot limit does not, however, account for the fluxes of energy that bypass the engine. This aspect becomes important for the atmosphere, where radiation transports energy from the heated surface to the cold atmosphere and which cannot be utilized by the convective heat engine of the atmosphere. This bypass results in unavoidable entropy production within the system which can only be reduced, but not avoided. Secondly, the mechanical work that is derived from the engine actually feeds back upon itself. Atmospheric motion, for instance, results from this work being performed, yet atmospheric motion is associated with the heat flux that drives the engine. In other words, there are internal