Non-equilibrium thermodynamics, maximum entropy production and Earth-system evolution

BY AXEL KLEIDON*

Max-Planck-Institut für Biogeochemie, Hans-Knöll-Strasse 10, 07745 Jena, Germany

The present-day atmosphere is in a unique state far from thermodynamic equilibrium. This uniqueness is for instance reflected in the high concentration of molecular oxygen and the low relative humidity in the atmosphere. Given that the concentration of atmospheric oxygen has likely increased throughout Earth-system history, we can ask whether this trend can be generalized to a trend of Earth-system evolution that is directed away from thermodynamic equilibrium, why we would expect such a trend to take place and what it would imply for Earth-system evolution as a whole. The justification for such a trend could be found in the proposed general principle of maximum entropy production (MEP), which states that non-equilibrium thermodynamic systems maintain steady states at which entropy production is maximized. Here, I justify and demonstrate this application of MEP to the Earth at the planetary scale. I first describe the non-equilibrium thermodynamic nature of Earth-system processes and distinguish processes that drive the system’s state away from equilibrium from those that are directed towards equilibrium. I formulate the interactions among these processes from a thermodynamic perspective and then connect them to a holistic view of the planetary thermodynamic state of the Earth system. In conclusion, non-equilibrium thermodynamics and MEP have the potential to provide a simple and holistic theory of Earth-system functioning. This theory can be used to derive overall evolutionary trends of the Earth’s past, identify the role that life plays in driving thermodynamic states far from equilibrium, identify habitability in other planetary environments and evaluate human impacts on Earth-system functioning.

Keywords: evolution; Gaia hypothesis; habitability; palaeoclimatology; thermodynamics; global change

1. Introduction

The present-day composition of the Earth’s atmosphere reflects a state of the Earth system that is far from thermodynamic equilibrium. This state manifests itself for instance in 21 per cent molecular oxygen and an atmospheric water vapour content that is mostly unsaturated even though about three-quarters of the Earth’s surface is covered with open water. In contrast, in thermodynamic equilibrium we would expect no reactive oxygen in the Earth’s atmosphere, and

*akleidon@bgc-jena.mpg.de

One contribution of 17 to a Theme Issue ‘Patterns in our planet: applications of multi-scale non-equilibrium thermodynamics to Earth-system science’.
Figure 1. A planetary state of maximum entropy production (MEP) on Earth corresponds to one at which incoming solar radiation is absorbed at a maximum possible rate and re-emitted to space at Earth’s radiative temperature. Such a state could result from the contrasting effects of surface temperature on snow/ice extent (decreases with increasing surface temperature) and convective cloud cover (should increase with increasing surface temperature as warmer air can hold more moisture). Adapted from Kleidon (2004).

an atmospheric water vapour content at saturation. This leads to the obvious question of why (and how) Earth’s environmental conditions are maintained in a state far from thermodynamic equilibrium.

This state far from equilibrium was noted more than 40 years ago by James Lovelock (1965) and linked to the profound effect that life has on Earth-system functioning. One of the important effects of biotic activity is that it alters the rates of geochemical reactions. For instance, Schwartzman & Volk (1989) estimated that biotic activity accelerated the rates of silicate rock weathering by one to three orders of magnitude, thereby altering the geological carbon cycle and planetary habitability (also Berner 1997). From this thermodynamic view, Lovelock later developed the controversial Gaia hypothesis (Lovelock 1972a,b, 1975; Lovelock & Margulis 1974). This hypothesis, roughly speaking, conjectures that Earth’s unique thermodynamic state is caused and maintained by life, and that life regulates Earth into a most habitable state that is in atmospheric homeostasis.

Such ‘Gaian’ behaviour of the Earth system can be understood on thermodynamic grounds (Kleidon 2004). The ‘Gaian’ behaviour resulted from three assumptions: (i) that non-equilibrium thermodynamic systems are maintained in states in which entropy production is maximized (proposed principle of maximum entropy production (MEP), see below), (ii) that such a state of MEP exists at the planetary level owing to the contrary effects of surface temperature on clouds and ice cover (figure 1), and (iii) that life adds sufficient degrees of freedom to geochemical reactions, in particular those affecting the carbon cycle and thereby the atmospheric greenhouse forcing. This would then allow the maintenance of optimum surface temperatures that are associated with MEP at the planetary scale.
However, there are clear objections to the Gaia hypothesis and a homeostatic Earth history (e.g. Kirchner 1989). For instance, the reconstructed record of atmospheric oxygen over the lifetime of Earth (Holland 2006) indicates that for the first half of Earth’s history oxygen existed in the atmosphere only in trace gas amounts, and that its concentration has increased over time. The concentrations of other substances, such as carbon dioxide and methane, show similar, quite clear trends throughout Earth’s history (Catling 2005). Rather than claiming regulatory behaviour throughout Earth-system history, we can nevertheless ask whether these trends can be understood and predicted from non-equilibrium thermodynamics.

Such predictability of basic trends in planetary evolution may have its roots in the proposed principle of MEP. This principle states that non-isolated thermodynamic systems maintain steady states at which the internal generation of entropy is maximized for given boundary conditions. The MEP principle has recently gained increased attention (e.g. recent reviews by Ozawa et al. (2003), Martyushev & Seleznev (2006), Kleidon & Schymanski (2008) and Kleidon (2009b) and the book by Kleidon & Lorenz (2005)), but has mostly been applied to purely physical processes within the Earth system, such as turbulence in the atmosphere (e.g. Paltridge 1975, 1978; Grassl 1981; Lorenz et al. 2001; Kleidon et al. 2003). Recent theoretical progress attempts to prove its validity on statistical arguments (Dewar 2003, 2005a, b), particularly the MaxEnt formalism of statistical physics and information theory (Jaynes 1957a, b). This line of justification for MEP would be identical to the assumption of a macroscopic state of maximum probability for isolated thermodynamic systems in equilibrium. That is, MEP describes the macroscopic state of the system that is the most probable given energy and mass-balance constraints.

When we want to apply MEP to the Earth system as a whole, we need to recognize that the Earth-system processes operate and interact at vastly different time scales. Atmospheric turbulence, for instance, takes place on time scales of seconds to days while the geological carbon cycle operates on time scales of hundreds of thousands to millions of years. Despite these drastically different time scales, these two processes interact strongly: atmospheric motion drives the water cycle, which allows for the transport of dissolved rock minerals from the land to the ocean. This solute transport in turn is a critical component of the geological carbon cycle as it supplies the majority of the calcium to the oceans, in which it ultimately binds with carbon dioxide (CO₂) to form limestone, thereby removing CO₂ from the climate system. By shaping the long-term evolution of atmospheric CO₂ concentrations, the geological carbon cycle then feeds back to atmospheric motion through its effects on the atmospheric greenhouse effect.

Because the Earth-system processes operate at contrasting time scales, a state of MEP at the planetary scale would not be achieved instantaneously. Fast processes would maintain MEP associated with these particular processes given the slowly evolving boundary conditions. As slow processes evolve towards MEP at their own pace, the boundary conditions for the fast processes change and so does their associated state of MEP. Hence, the application of MEP to planet Earth as a whole would mainly manifest itself in a trend towards states of higher planetary entropy production. For its application we first need to identify and describe the Earth-system functioning in terms of non-equilibrium thermodynamics, in particular regarding the transformations...
and cycling of energy and mass and how these interact. We can then identify how the Earth’s environmental conditions should look like closer and further away from a state of planetary thermodynamic equilibrium and derive the trends in environmental conditions as the planetary system evolves further from thermodynamic equilibrium.

This consideration should lead us closer to formulate an integrated understanding of the Earth system and its past as well as its future evolution. It would provide a basis to obtain a first-order evolutionary trend of the Earth system and understand regulatory behaviour on thermodynamic grounds. It would allow us to quantify the extent to which the biota drives disequilibrium for processes beyond those shaping atmospheric oxygen concentration, which should allow us to evaluate the habitability of other planetary systems. We can also apply this novel perspective to future global change and evaluate whether human-induced changes, such as altering greenhouse gas concentrations or land cover transformations, bring the Earth system further away or closer to thermodynamic equilibrium.

The goal of this paper is to provide a sketch of how such a quantification could be achieved, to outline thermodynamic trends of the Earth system, and to explore applicability of the MEP principle to the planet. To do so, I first briefly review the basics of non-equilibrium thermodynamics and how they apply to the Earth system in §2 and then illustrate their interconnectedness in §3. I then use this description to outline how the Earth system should look in a state at thermodynamic equilibrium and far away from it, and what the resulting evolutionary trend should look like in §4. I close with a brief summary and conclusion.

2. Characterizing disequilibrium of Earth-system processes

(a) Characterizing thermodynamic disequilibrium

The first and second law are central to thermodynamics. While the first law states the conservation of energy, the second law can be generalized for non-equilibrium systems to an entropy budget (e.g. Kondepudi & Prigogine 1998):

$$\frac{dS}{dt} = \sigma - \text{NEE},$$

(2.1)

where $dS/dt$ is the rate of change of entropy $S$ of the system, $\sigma$ the entropy being produced by irreversible processes within the system, and NEE the net entropy exchange across the system boundary associated with energy and mass fluxes. The second law in the context of equation (2.1) states that $\sigma \geq 0$.

I illustrate the evolution towards a state of thermodynamic equilibrium for isolated systems and the maintenance of a state of disequilibrium using a simple electric circuit (figure 2). In the electric circuit, the resistors play the role of irreversible processes—here, gradients in electric potential are dissipated into heat by the electric current running through the resistor. Capacitors can carry an initial condition in the form of an electric charge. An initial charge provides the free energy that creates a gradient in electric potential and creates the current which depletes the initial charge. Generators provide the electric-free energy source and allow the circuit to dissipate the electric energy into heat in a steady state.
Figure 2. (a) Representation of dissipative processes in terms of an electric circuit network. Resistances \( R_i \) represent the dissipative processes, the charges in the capacitors with capacitances \( C_i \) represent the state away from thermodynamic equilibrium, and the generator with an electromotive force \( E \) (and internal resistance \( R_G \)) represents the process that drives the system out of thermodynamic equilibrium. (b) The change of currents with time for the circuit shown in (a) in qualitative terms. (c) Demonstration of a state of maximum dissipation at resistance \( R_2 \) in steady state associated with variation of \( R_2 \). Dissipation \( D \) and current \( I_2 \) are shown as a solid (dotted) line and in relative units, normalized to their maximum value, respectively. Values of \( R_G = 100 \, \Omega \) and \( E = 100 \, \text{V} \) were used.

This example of an electric circuit allows us to characterize how a non-equilibrium state is expressed (the amount of charge in the capacitors) and how processes can be classified into those that create free energy (the generators) and those that dissipate free energy (the resistors). The electric circuit description is applicable to practically all Earth-system processes. In formal thermodynamic terms, the electromotive force \( E \) associated with the generator and the charge \( Q \) stored in the capacitor are conjugated variables just like the ‘traditional’ thermodynamic variables entropy \( S \) and temperature \( T \), and volume \( V \) and pressure \( p \). Likewise, for mass fluxes, the conjugated variables chemical potential \( \mu \) and particle number \( N \) are relevant, for momentum transfer stress \( \tau_{ij} \) and velocity \( v \).
To link work and dissipation in the electric circuit with entropy production, first note that temperature differences play little role in the discussion, so that entropy production is proportional to dissipation. The work done on the circuit is, for instance, reflected in the electromotive force associated with the generator or by charging the capacitor. Dissipation occurs at the resistances of magnitude $D = R \cdot I^2$, with $I$ and $R$ being the current and resistance, respectively.

This electric circuit description of dissipative processes is used below to describe Earth-system processes. We can then ask what type of loop we deal with within the Earth system in the majority of cases. The first loop in figure 2a consists of a capacitor $C_1$ and a resistance $R_1$. The current in this loop would deplete the initial charge on the capacitor, resulting in a decrease of the current (and charge) with time ($I_1$ in figure 2b). In the second loop, a generator with electromotive force $E$ is added. In this case, the current would increase in time, reach a steady-state value of $I_2 = E/R_2$, and maintain a state away from equilibrium ($I_2$ in figure 2b).

A state of MEP within the electric circuit can be identified in association with the dissipation of resistor $R_2$ (figure 2c) if we assume that there are sufficient degrees of freedom for $R_2$ to adjust to an optimum value. This could, for instance, be the case for a sufficiently complex network of many resistors that provide many alternative ways to dissipate potential gradients (see Zupanovic et al. (2004) for a derivation of Kirchhoff’s loop law from MEP). Given a certain, fixed internal resistance $R_G$ within the generator, the state of MEP then results from the trade-off of a greater value of $R_2$ reducing the current $I$ in the dissipation $D = I^2 \cdot R$ of the resistor $R_2$ (figure 2c). In electrical engineering, this concept is known as the maximum power theorem.

(b) Disequilibrium and Earth-system processes

We can now characterize Earth-system processes in terms of generators, resistors and capacitors (as summarized in table 1). In the following explanation, I attempt to place the focus on the dominant processes and properties that reflect, cause and dissipate disequilibrium. For the purpose of illustration, I will leave out many aspects, so the following description is clearly incomplete, but hopefully captures the most important aspects.

(i) Radiative exchange

Radiation is in equilibrium with matter when the photon composition of the radiation corresponds to the temperature of the matter (in the case of a blackbody). The radiative temperature of the radiation is then the same as the thermodynamic temperature of the matter. Incoming solar radiation carries a photon composition corresponding to the emission temperature of the Sun, which has been ‘diluted’ by the Earth–Sun distance. This composition is far from thermodynamic equilibrium with the temperature that results from full absorption and re-emission of that radiation. Hence, the disequilibrium is reflected in the temperature difference between the emission temperature of the radiation and the temperature at which radiation is absorbed. Emission of radiation acts to bring absorbed radiation back to thermodynamic equilibrium (the resistor), while absorption causes disequilibrium (the generator).

Phil. Trans. R. Soc. A (2010)
Table 1. Earth-system processes and their characterization in terms of properties reflecting a state away from thermodynamic equilibrium, and those processes that drive the state towards or away from thermodynamic equilibrium.

<table>
<thead>
<tr>
<th>process</th>
<th>property reflecting disequilibrium</th>
<th>processes towards equilibrium</th>
<th>processes that drive disequilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>electric circuit</td>
<td>charge in capacitors</td>
<td>dissipation in resistors</td>
<td>electromotive forces associated with generators</td>
</tr>
<tr>
<td>radiative exchange</td>
<td>temperature difference to radiative equilibrium temperature</td>
<td>emission of radiation</td>
<td>absorption of radiation emitted at a higher temperature</td>
</tr>
<tr>
<td>large-scale motion</td>
<td>kinetic energy</td>
<td>frictional dissipation</td>
<td>gradients in pressure and density gravitational pull from the Moon and Sun</td>
</tr>
<tr>
<td>hydrologic cycling</td>
<td>distance of atmospheric water vapour content to saturation</td>
<td>evaporation brings unsaturated air closer to saturation condensation brings supersaturated air closer to saturation updrafts bring water vapour to supersaturation</td>
<td>falling raindrops dissipate potential energy atmospheric motion lifts unweathered material to the surface precipitation adds water to continents</td>
</tr>
<tr>
<td>carbon cycling</td>
<td>difference in CO2 concentration between atmosphere and interior organic carbon topographic gradients</td>
<td>volcanic outgassing respiration erosion depletes potential gradients photosynthesis plate tectonics lifts unweathered material to the surface</td>
<td>mantle convection</td>
</tr>
<tr>
<td>rock cycling</td>
<td>distance to saturation</td>
<td>dissolution brings concentration to saturation</td>
<td>input of unsaturated water by precipitation</td>
</tr>
</tbody>
</table>

(ii) Large-scale motion

The property that reflects the state away from thermodynamic equilibrium is the kinetic energy associated with motion (the charge in the capacitor). Kinetic energy is essentially free energy that is irreversibly converted into heat by frictional dissipation (the resistance). Motion is generated by gradients in pressure and density, mostly caused by gradients in radiative heating and cooling, but also by external factors such as the gravitational pull of the Moon (the generators).
(iii) Hydrologic cycling

Water vapour in the atmosphere is in thermodynamic equilibrium with an open water surface when it is saturated. The charge on the capacitor is represented by the difference to saturation, either in terms of unsaturated air near the surface or in terms of supersaturation aloft. Both evaporation at the surface and condensation at height act to bring the water vapour content closer to saturation (the resistances). What drives the vapour concentration out of equilibrium (the generators) are: (i) updrafts, which bring vapour to supersaturation; and (ii) precipitation, which removes moisture from lifted air that subsequently becomes unsaturated when it descends back down to the surface.

The amount of liquid and solid water in the atmosphere also represents disequilibrium (the capacitor), as these would—in the absence of vertical motion (the generator)—precipitate out and thereby dissipate the potential energy associated with their height in the atmosphere (the resistances).

On land, disequilibrium is represented by the presence of soil water on land at a certain elevation above sea level (the capacitor). Continental runoff depletes this potential energy and dissipates it into heat by friction (the resistance). Precipitation over land acts as the generator that drives the continental water balance out of equilibrium.

(iv) Carbon cycling

Disequilibrium of the carbon cycle is reflected in the difference of atmospheric CO₂ concentration to the CO₂ concentration that would be in equilibrium with the high temperatures of the mantle (the capacitor). This difference results in volcanic outgassing (the resistance). The process that drives this disequilibrium is mantle convection (the generator).

Disequilibrium is also reflected in the amount of organic carbon as this is directly related to the free energy that is stored in the organic compounds (the capacitor). Photosynthesis acts as the generator since it supplies the carbohydrates to build up organic carbon. Respiration is the dissipative process that depletes this disequilibrium (the resistance).

(v) Rock cycling

The cycling of the geochemical elements from continental rocks to the ocean floor and subsequent recycling by plate tectonics involve the transport of sediments and matter in dissolved form. Disequilibrium with respect to sediments is reflected in topographic gradients (the capacitor). Erosion fluxes and sediment transport deplete these gradients (the resistances). The generator is plate tectonics, formation of continental crust and the associated uplift of fresh rock material.

The disequilibrium of material cycling in dissolved form is reflected in the distance to saturation of the dissolved compounds in water (the capacitor). Dissolution of elements from the rock by chemical weathering brings the concentration in water towards saturation (the resistance). The disequilibrium is driven by the input of precipitation, which is usually far from being saturated (the generator).
(c) Example: the global carbon cycle as an electric circuit

I use the global carbon cycle as an example of how such an electric circuit analogy is formulated for material cycles in the Earth system in more detail (figure 3).

The generators that drive the carbon cycle are mantle convection and photosynthesis, as explained above (table 1). Mantle convection acts as the generator that renews the material in contact with the reservoirs of the surface and atmosphere, thereby maintaining the potential gradient between the interior and the atmosphere. Photosynthesis withdraws CO₂ from the atmosphere and converts it into chemical free energy in the form of carbohydrates using solar radiation.

These generators maintain gradients in CO₂ concentration and free energy. These gradients are depleted by the various processes that complete the carbon cycle:

(i) *volcanic outgassing* depletes the gradient between the CO₂ concentrations of the atmosphere and the mantle;

(ii) *respiration* by photosynthesizing tissue, organisms, heterotrophs and fire deplete the free energy stored in the form of carbohydrates, alive and dead biomass (applies to land and ocean biota);

(iii) *diffusion* depletes the CO₂ gradient between the soil and the atmosphere that is maintained through respiration in the soil;

(iv) *diffusion* and *turbulent transport* mix differences in CO₂ concentrations in the atmosphere;

(v) *diffusion* and *mixing* deplete the difference in CO₂ concentration between dissolved CO₂ in rivers and the atmospheric concentration;
air–sea gas exchange deplete the difference in CO$_2$ concentration between atmosphere and oceans;

(vii) precipitation of calcium carbonates depletes the supersaturation of calcium and binds dissolved CO$_2$ in sea water; and

(viii) in the case of unsaturated conditions, dissolution acts to bring the concentration of calcium back to saturation.

As is the case for all irreversible processes, the entropy production by these processes is expressed as the product of a thermodynamic force and flux. The thermodynamic force is expressed in the form of a gradient in chemical potential $\Delta \mu$, which is related to the CO$_2$ concentration, expressed in terms of the corresponding partial pressure (e.g. Kleidon 2009b). The mass fluxes $F$ are generally parametrized in terms of these gradients (i.e. $F = k \cdot \Delta \mu$ with some effective diffusivity $k$). The fluxes have in common that the fluxes $F$ deplete the forces $\Delta \mu$ that drive them. If this trade-off between flux and force can be accomplished in many different ways (e.g. because of different chemical reaction pathways), this would result in a range of plausible values for $k$, and the MEP principle should be applicable to select for the value of $k$ that maximizes entropy production ($\sigma = F \cdot \Delta \mu / T$) in steady state.

3. Interaction between dissipative Earth-system processes

The dissipative processes that were described in the previous section strongly interact. The dominant forms of interaction are summarized in figure 4 and the consequences of these interactions on the overall dissipation of these processes are explained in more detail in the following. The aspect that is emphasized here is the impacts of these effects on gradients and the fluxes that deplete these gradients.

The most important driving force of the Earth system is the incident solar radiation. The difference in orientation of the Earth’s surface to the incoming solar radiation as one moves from the tropics to the poles leads to uneven interception of solar radiation, resulting in gradients in the absorption of solar radiation. This radiative gradient causes differences in heating and cooling, thereby resulting in temperature gradients. Temperature gradients in turn cause differences in density and pressure, which produce forces that set the atmosphere and oceans into motion.

Atmospheric motion is intimately linked with hydrologic cycling (Pauluis & Held 2002a, b). By lifting air masses, water vapour is brought to saturation, condensation and precipitation. The resulting loss of condensed water from the atmosphere results in overall dehumidification, which then drives the evaporation of water from the surface into the atmosphere. Hence, the strength of the hydrologic cycle and the extent to which it operates away from thermodynamic equilibrium are, to a first approximation, directly linked to the strength of atmospheric motion.

The disequilibrium reflected in the hydrologic cycle is associated with the transport of water to land. This net transport of moisture by the atmosphere is returned to the oceans as river discharge, which provides the means to transport sediments and dissolved elements from continental rocks to the oceans. The stronger the hydrologic cycle, the more effectively the products of weathering
and erosion should be transported to the ocean. Since the geological carbon cycle depends on the supply of calcium ions from weathering, this linkage provides a basic connection between the strength of the hydrologic cycle and the carbon cycle (and other geochemical cycles for which the main source are continental rocks).

This chain of energy conversions from radiative gradients to geochemical cycling is not acting solely in one direction, but various interactions take place. Temperature gradients result in uneven emission of terrestrial radiation, thereby altering radiative gradients. Motion transports heat, thereby depleting temperature gradients. Hydrologic cycling contributes to this heat transport by the conversions of water from vapour to liquid to solid and vice versa and the associated uptake and release of latent heat that is involved in these phase changes. Hydrologic cycling also affects cloud and ice cover, both factors being important as they cause reflection and scattering of incident solar radiation to space (as illustrated in figure 1). Hydrologic cycling and geochemical cycling (especially carbon cycling) alter the atmospheric composition, specifically with respect to the important greenhouse gases (water vapour, clouds, CO2). These, in turn, affect radiative gradients.

Biotic activity, in the form of photosynthesis, skips this chain of downgraded energy by using the low entropy nature of sunlight directly to drive photochemical reactions. The free energy generated by photosynthesis then results in an alteration of the rates at which geochemical and hydrological cycling takes place. For instance, root systems of terrestrial vegetation allow for the extraction of water stored deep in the soil that would be unavailable for evapotranspiration in the absence of a transpiring vegetative cover. Decomposition of organic matter in
Table 2. Conditions of the Earth system close to and far from a state of planetary thermodynamic equilibrium. Adapted from Kleidon (2009a).

<table>
<thead>
<tr>
<th>property or flux</th>
<th>close to equilibrium</th>
<th>far from equilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>motion</td>
<td>none</td>
<td>high</td>
</tr>
<tr>
<td>frictional dissipation</td>
<td>none</td>
<td>high</td>
</tr>
<tr>
<td>relative humidity</td>
<td>saturated</td>
<td>low</td>
</tr>
<tr>
<td>cloud cover</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>net evaporation</td>
<td>0</td>
<td>high</td>
</tr>
<tr>
<td>net precipitation</td>
<td>0</td>
<td>high</td>
</tr>
<tr>
<td>continental runoff</td>
<td>0</td>
<td>high</td>
</tr>
<tr>
<td>geochemical cycling</td>
<td>none</td>
<td>high</td>
</tr>
<tr>
<td>atmospheric CO₂</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>greenhouse effect</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>surface temperature</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>ice cover</td>
<td>0</td>
<td>polar, seasonal</td>
</tr>
<tr>
<td>absorption of solar radiation</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>planetary entropy production</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>conditions for life</td>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

Imagine the Earth system to be in a state close to thermodynamic equilibrium. This state would imply the absence of large-scale motion and hence no frictional dissipation. Since atmospheric motion acts as the engine that drives dehumidification of the atmosphere, we would expect the atmosphere to be close to saturation and likely to be quite cloudy. The hydrologic cycle would practically be absent in the sense that net rates of precipitation and evaporation are zero, that is, conditions of local thermodynamic equilibrium ($P \approx E$) would hold nearly everywhere. We would then have no transport of moisture to the continents, hence no continental runoff. With no continental runoff, there would be no flux of sediments and dissolved compounds from land to ocean, hence no geological cycling, which would result in a high atmospheric CO₂ concentration. This, in turn, would imply a strong greenhouse effect and high surface temperatures. High surface temperatures do not allow for the presence of large-scale snow and

4. Application to Earth-system evolution

When we now combine the disequilibrium states of the different Earth-system processes (table 1) with the interactions shown in figure 4, we can derive a first-order integrated view of what the conditions on Earth as a whole should look like in a state closer to planetary thermodynamic equilibrium and how conditions change in an evolution away from thermodynamic equilibrium. This is summarized in table 2.

Imagine the Earth system to be in a state close to thermodynamic equilibrium. This state would imply the absence of large-scale motion and hence no frictional dissipation. Since atmospheric motion acts as the engine that drives dehumidification of the atmosphere, we would expect the atmosphere to be close to saturation and likely to be quite cloudy. The hydrologic cycle would practically be absent in the sense that net rates of precipitation and evaporation are zero, that is, conditions of local thermodynamic equilibrium ($P \approx E$) would hold nearly everywhere. We would then have no transport of moisture to the continents, hence no continental runoff. With no continental runoff, there would be no flux of sediments and dissolved compounds from land to ocean, hence no geological cycling, which would result in a high atmospheric CO₂ concentration. This, in turn, would imply a strong greenhouse effect and high surface temperatures. High surface temperatures do not allow for the presence of large-scale snow and
ice cover at the surface. In total, this state would show high reflectance of solar radiation owing to the high cloud cover, and planetary entropy production would be low. This, of course, directly follows from the starting point of considering a planetary state close to thermodynamic equilibrium. We would find the conditions of Earth to be close to the right end of the $x$-axis in figure 1. The conditions for life would be low because of little availability of solar radiation and low availability of nutrients owing to the lack of geochemical cycling.

The Earth in a state far from equilibrium would look very different. Strong atmospheric motion would result in high rates of frictional dissipation and efficient dehumidification of the atmosphere. In such an atmosphere we would expect to have far fewer clouds that have much shorter lifetimes as the strong upward motions would quickly bring them to condensation and to rain out. We would find a hydrologic cycle characterized by a much greater imbalance of condensation and evaporation (i.e. $|P - E| \gg 0$ locally), hence a stronger, global hydrologic cycle taking place at much larger scale. The atmospheric circulation would transport large amounts of moisture to the continents, resulting in high rates of continental runoff. This, in turn, would allow for relatively large fluxes of sediments and dissolved compounds from land to ocean, and, therefore, in much greater geochemical cycling rates. This would draw down the geological carbon cycle to low concentrations of atmospheric CO$_2$ and a weak greenhouse effect. The weak greenhouse effect would allow for surface temperatures that permit large-scale snow and ice cover. A seasonal snow and ice cover at the poles would enhance the gradient of absorbed solar radiation between the tropics and the poles, thereby further strengthening the large-scale atmospheric circulation. We would now be close to or even at the MEP state illustrated in figure 1. These conditions would allow for high biotic activity because of the greater availability of solar radiation at the surface and high nutrient availability owing to the strong geochemical cycling.

Cold, ‘Snowball-Earth’ like conditions of the Earth (e.g. Evans et al. 1997; Hoffman et al. 1998) that are to the left of the MEP state in figure 1 could only be reached in transient states. Since radiative gradients would not be at their maximum value, the strength of the atmospheric circulation would be weaker than at the MEP state, allowing for less hydrologic and geochemical cycling. This, in turn, would not allow the maintenance of low CO$_2$ concentrations in the atmosphere in steady state. Hence, atmospheric CO$_2$ would slowly increase, the surface would warm, and the system would transit back to the MEP state.

Using these two extreme thermodynamic states we can now get a comprehensive picture of how Earth’s environment should have changed in the past from the assumption that the Earth system has mainly evolved away from thermodynamic equilibrium. The early Earth, being in a state closer to thermodynamic equilibrium, would show high concentration of atmospheric CO$_2$, high temperatures, high cloud cover and little dissipation. Through time, the Earth would evolve towards states of higher geochemical cycling, lower concentration of greenhouse gases, lower cloud cover and colder surface temperatures. This trend towards lower greenhouse gas concentrations qualitatively matches reconstructions (Catling 2005).

In this picture, life would represent the means by which alternative and more complex geochemical reaction pathways are added to the system thus resulting in more degrees of freedom associated with geochemical cycles.
Higher degrees of freedom allow the system to evolve to states of higher entropy production in which more free energy is generated and subsequently dissipated. The interaction of life adding degrees of freedom to the geochemistry on the one hand, but environmental conditions of the Earth system constraining biogeochemical reaction rates on the other could then result in a positive feedback loop: if a change in environmental conditions takes place that allows for new geochemical reaction pathways to be ‘profitable’ (i.e. they generate the free energy necessary to maintain the means by which this pathway is achieved), the associated geochemical process would achieve a state of higher entropy production. The altered rate would affect environmental conditions, possibly providing environmental conditions for even more new pathways to be ‘profitable’, which would enhance entropy production even more, thereby establishing the positive feedback loop. This feedback would be bound by certain energetic thresholds that would place a constraint on the extent to which new degrees of freedom would be ‘profitable’.

In principle, such trends and the feedback could be tested by comparison with palaeontological records and by process-based modelling. At a qualitative level, the trend to higher entropy production of biotic processes that would be associated with a biosphere of lower entropy and more degrees of freedom could be linked to the general trend towards higher biodiversity through time, as for instance reflected in the record of the vascular plants on land (Niklas et al. 1983), and the existence of energetic thresholds could be linked to major evolutionary transitions at the planetary level (Lenton et al. 2004). Testing this feedback with process-based models would require a representation of the biosphere that includes biodiversity dynamics and a thermodynamic evaluation of biotic effects on Earth-system functioning.

What has not been included in the discussion here is that the extent of continental areas is also an internal variable that links processes at the surface with those in the interior. The presence of continents implies topographic gradients (i.e. gradients in gravitational potential) that are depleted by the erosion flux, resulting again in a trade-off between thermodynamic flux (erosion flux) and force (topographic gradient). In this context, Rosing et al. (2006) speculated that the presence of life has contributed significantly to the rise of continents by supplying additional free energy to geochemical cycles. This coupling to processes in the Earth’s interior would need to be included in a full treatment of a planetary state of thermodynamic equilibrium.

5. Summary and conclusions

In this paper I have provided a qualitative sketch of how non-equilibrium thermodynamics and maximum entropy production should be applicable to the Earth system as a whole, and how this should result in some basic thermodynamic trends in its evolutionary history. Such thermodynamic considerations have the potential to provide us with a unifying theory of Earth-system functioning, which should allow us to much better constrain and reconstruct the past history of the Earth system as well as future trajectories owing to human-induced changes and therefore be of invaluable help.
Obviously, at this point what has been described here is just simple, qualitative reasoning that seems reasonable under the assumption that MEP is valid. To move forward, the theoretical foundation needs to be established more firmly, and further examples are needed to assess the validity of MEP and its range of application. In this respect, some examples of how MEP is applied to specific cases in environmental and ecological systems are to be found in a forthcoming special issue on this topic of *Philosophical Transactions B* (edited by A. Kleidon, Y. Malhi and P. Cox).

The author thanks the Helmholtz-Gemeinschaft as this research has been supported by the Helmholtz Association through the research alliance ‘Planetary Evolution and Life’.

References


