Review of the book

THERMODYNAMICS

A Dynamical Systems Approach

by Wassim M. Haddad, VijaySekhar Cellaboina, and Sergey Nersesov

Princeton Series in Applied Mathematics,
Princeton University Press, 2005.

Reviewed by Jan C. Willems
ESAT, K.U. Leuven, B-3001 Leuven, Belgium
www.esat.kuleuven.ac.be/∼jwillems
Jan.Willems@esat.kuleuven.ac.be

March 8, 2006

1 Introduction

Thermodynamics, the science of heat and work and hot and cold, puts forward a number of principles that have far reaching consequences in physics and engineering. Central to thermodynamics are two laws. The first law states that energy is conserved. Energy can be transformed from one form to another, but it cannot be destroyed nor can it be created. There are many equivalent statements of the second law. The most common one is that the increase of entropy is larger than the heat delivered to the system divided by the temperature.

These laws bring good news and bad news. The first law is a comforting thought in an economy of ever increasing energy bills. The second law is, for sure, not a laughing matter. One consequence is that we cannot let a physical system interact with its environment and make it go through a time history that brings both the system and the environment in the same condition at the end as they had in the beginning. Another consequence of the second law is that in a system that does not exchange heat with its environment, entropy is forever increasing. Accordingly, in the words of Kelvin, the universe is destined to come to a state of eternal rest. This consequence of the second law has come to be known as *the heat death of the universe*.

The book under review starts off with a number of quotes about thermodynamics. One is by Einstein:

> *Thermodynamics is the only physical theory of a universal nature of which I am convinced that it will never be overthrown.*

Another is by Eddington:

> *The law that entropy increases — the second law of thermodynamics — holds, I think, the supreme position among the laws of Nature.*

From an engineering point of view, the laws of thermodynamics have far reaching consequences. For example, it is not possible to simply transport heat from one place to another. We cannot achieve refrigeration by cooling one room and heating another. This transformation, unfortunately, requires intervention of another energy source, at home typically electricity. Another consequence is that, notwithstanding the law of conservation of energy, not all forms of energy are equally valuable, with heat being the 'lowest' form. As a result, it is unavoidable that electrical power generation stations that burn oil or gas or carbon or nuclear fuel to produce electrical power, must also produce waste in the form of heat. They usually dump this heat into the environment, often causing unpleasant side effects for fauna and flora. The inability to transform also this waste heat into electrical energy is not a matter of unwillingness or of inefficiency, but an unavoidable consequence of the laws of thermodynamics.

From the pedagogical point of view, thermodynamics is a disaster. As the authors rightly state in the introduction, many aspects are *"riddled with inconsistencies"*. They quote V.I. Arnold, who concedes that *"every mathematician knows it is impossible to understand an elementary course in thermodynamics"*. Nobody has eulogized this confusion more colorfully than the late Clifford Truesdell. On page 6 of his book *The Tragicomical History of Thermodynamics 1822-1854* (Springer Verlag, 1980), he calls thermodynamics *"a dismal swamp of obscurity"*. Elsewhere, in despair of trying to make sense of the writings of some local heros as De Groot, Mazur, Casimir, and Prigogine, Truesdell suspects that there is *"something rotten in the (thermodynamic) state of the Low Countries"* (see page 134 of *Rational Thermodynamics,* McGraw-Hill, 1969).

The following seem to be stumbling blocks.

(i) The notion of entropy that enters in the second law. It is not a directly measurable physical quantity, contrary to temperature or pressure or volume. It somehow needs to be deduced from the laws of the system. Given the physical laws of a system, what is it then equal to? What is its domain? Is it uniquely defined?

(ii) The strange use of derivatives. As Truesdell notes, in this domain even derivatives look different, and one can find statements like

$$\frac{\delta^{\text{rev}} Q}{dV} = T \left(\frac{\partial S}{\partial V} \right)_P .$$

Such notation poses challenges, especially to eager students who have just passed a course on 'Functions of Many Variables'.

(iii) The many vaguely defined terms and functions, as 'entropy', 'enthalpy', 'Gibbs free energy', 'Helmholtz free energy', 'extensive' and 'intensive', 'reversible' and 'irreversible', etc., etc.

(iv) The tradition of invoking probability theory at random moments in an argumentation. Once one is thoroughly confused, one is invariably presented with a justification based on statistical mechanics. This in keeping with the basic debating principle that the most effective way of 'explaining' something that is badly understood is by invoking something that is even worse understood. When the going gets though, the though get going.

(v) The penchant for the big idea. The second law is often called the 'most metaphysical of all physical laws'. This has allowed thermodynamics to be used as support by the left as well as by the right, by believers as well as by non-believers, by creationists as well as by evolution theorists, and, I suspect, that intelligent designers will also find arguments in thermodynamics for their point of view. And when Shannon chose to use the term 'entropy' for 'amount of information', this was like poring oil on Maxwell's demon's eternal fire.

The book under review uses a rigorous mathematical format to thermodynamics. The logical line is refreshingly clear. The basic setting is the input/state/output formulation of dynamical systems theory, combined with interconnection laws among subsystems (called compartments). The construction of the internal energy and the entropy is solidly founded on the theory of dissipative systems and storage functions. Stability results invariably use rigorous Lyapunov theory arguments. Throughout, a definition/lemma/proposition/theorem/proof/corollary format is adopted. No statistical arguments are used. The difficulties referred to above are absent.

There have been previous attempts to give thermodynamics a solid mathematical underpinning. One notable program that set this as the goal is the work by the school of Noll, Coleman, Gurtin, e.a., documented in a series of publications in the journal *Archive for Rational Mechanics and Analysis* in the 1960's and 1970's (see also Truesdell's 1969 book referred to above). The present book is similar in philosophy. But, since it is based on input/state/output representations of dynamical systems, it has altogether a different flavor.

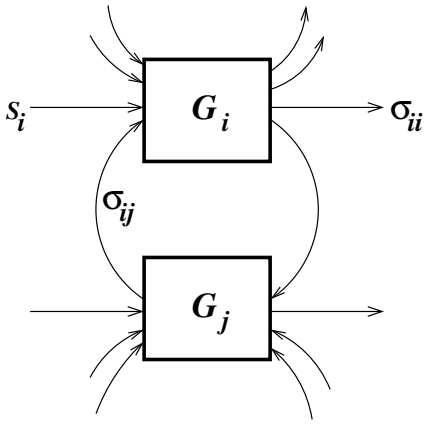
2 Contents

The book consists of eight chapters. Chapter 1, the introduction, sets the stage. It contains a historical introduction which discusses classical thermodynamics as laid out through the work of Carnot, Clausius, Kelvin, Planck, Gibbs, and Carathéodory. The

authors are, rightfully so, very sceptical of the coherence of classical thermodynamics. They then present their central thesis: that a state space formulation of dynamics, combined with interconnected nonlinear compartmental systems ensures a consistent model for heat and energy flow.

Chapter 2 is a mathematical introduction. It first introduces nomenclature around dynamical systems ('flows') of the form $\frac{d}{dt}z(t) = w\big(z(t)\big)$. Of special interest are state spaces $\mathbb{R}^q_+$ that are non-negative orthants of finite dimensional vector spaces. Various Lyapunov stability theorems are given. The authors then turn to concepts surrounding input/state/output systems. This is followed by a number of abstract concepts, as reversibility and recoverability. These concepts pertain to general dynamical flows and appear to be original. In later chapters, these notions are used in an effective way in the context of thermodynamics. They finally turn to volume preserving flows and Poincaré's recurrence theorem. All this is introduced on a very general level (the state space, for example, is assumed to be a Banach space, but the dynamics are fully nonlinear). The notation is somewhat heavy, and many of the definitions take at least 10 lines to write down. It is not an easy chapter to read. It is not uncommon for applied mathematics books to get the mathematical background out of the way before getting into the main subject matter. However, it is, in my opinion, never a reader-friendly idea to frontload a book with a chapter devoted to intricate mathematical concepts and notation.

The third chapter, entitled *A Systems Foundation for Thermodynamics*, is the core of the monograph. In it, the basic mechanism how thermodynamic subsystems are viewed to interact among each other and with the environment is explained. This set-up is shown in the figure below.



The system is composed of a (finite) number of interacting subsystems, called compartments: $G_1, \ldots, G_i, \ldots, G_j, \ldots, G_q$. The subsystem G_i receives heat from the environment at rate S_i (the sign determines whether heat flows in or out). Further, the system G_i receives heat from the system G_j at a rate $\sigma_{ij} \geq 0$, and dissipates heat at a rate $\sigma_{ii} \geq 0$. The subsystem G_i has internal energy $E_i \geq 0$. It is assumed that this energy $E_i \in \mathbb{R}_+$ is the state the i-th compartment, leading to the overall state

$E := (E_1, \ldots, E_i, \ldots, E_j, \ldots, E_q)$. The heat flow rates σ_{ij} are all assumed to be functions of the full state E. This leads to the following system of differential equations describing the interconnected system

$$\frac{d}{dt}E_i = \sum_{j=1, j\neq i}^{q} \sigma_{ij}(E) - \sum_{j=1, j\neq i}^{q} \sigma_{ji}(E) - \sigma_{ii}(E) + S_i.$$

These equations are coupled, because E involves all the subsystem energies E_i.

I was slightly confused by some elements in this basic model. For one thing, the role of the dissipation term σ_{ii} escapes me. I would have simply replaced S_i by $S_i - \sigma_{ii}$. The fact that the authors have both terms separately in the equations may have to do with input/output thinking, which, I have argued repeatedly, is all too often simply not suitable for the analysis of physical systems. I will come back to this point in subsection 3.4. In the same vein, I did not see the need to introduce both σ_{ij} and σ_{ji}. It is unclear to me why, if heat flows between the compartments i and j, we should see this as the difference of a non-negative flow from i to j minus a non-negative flow from j to i. Another element which I found confusing is the fact that the heat flows σ_{ij} were assumed to be a function of the whole state vector $E := (E_1, \ldots, E_i, \ldots, E_j, \ldots, E_q)$. I realize that this gains generality, but it seems to me that true compartmental thinking would come up with $\sigma_{ij}(E_i, E_j)$. Related to this is my lack of understanding of what is called *diffusion* on page 38. I find it hard to understand how a heat flow from compartment j in state E_j to compartment i in state E_i could be of the form $\sigma_{ij}(E_j)$, not involving E_i. But, all by all, these shortcomings are minor, and the model used is in the end quite convincing.

After having set up their model, the authors prove that the interconnected system is conservative with storage function $\sum_{i=1}^{q} E_i$ and supply rate $\sum_{i=1}^{q} (S_i - \sigma_{ii}(E))$, meaning that along solutions of the dynamical equations, there holds

$$\frac{d}{dt}\sum_{i=1}^{q} E_i = \sum_{i=1}^{q} (S_i - \sigma_{ii}(E)).$$

Under some additional reasonable assumptions (called 'axioms' in the book) on the σ_{ij}'s, it is further proven that for cyclic processes: $E(t_{\text{initial}}) = E(t_{\text{terminal}})$, there holds

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \sum_{i=1}^{q} \frac{S_i - \sigma_{ii}}{E_i + \mathrm{c}}\, dt \leq 0.$$

Here c is an arbitrary positive constant that expresses the fact that energy is only defined up to an additive constant. Next, the theory of dissipative systems is used to show the existence of a state function, $E \mapsto \mathfrak{S}(E)$, the entropy function, satisfying

$$\frac{d}{dt}\mathfrak{S}(E) \geq \sum_{i=1}^{q} \frac{S_i - \sigma_{ii}(E)}{E_i + \mathrm{c}}.$$

Subsequently, it is proven, by considering the available and the required entropy, and using some intricate clever analysis, that the above inequality defines the entropy uniquely up to an addtive constant, and that it is given by

$$\mathfrak{S}(E) = \sum_{i=1}^{q} \mathfrak{S}_{\mathrm{i}}(E_i)$$

with

$$\mathfrak{S}_{\mathrm{i}}(E_i) = \ln(E_i + \mathrm{c}).$$

It is then concluded that $E_i + \mathrm{c}$ is the temperature of the i-th compartment, and hence that

$$T_i = \frac{d\mathfrak{S}_i}{dE_i}.$$

Subsequently, the authors introduce a new notion related to entropy, called the *ectropy*. This concept is original to this book and is a true dual to entropy in the sense that entropy increases in an adiabatic (no heat exchange with the environment) regime if and only if ectropy decreases in that regime. Unlike entropy, however, ectropy seems to be a natural candidate (quadratic) Lyapunov function for analyzing stability and obtaining energy equipartition of thermodynamic systems using Lyapunov and invariant set theory.

In the remainder of the chapter, the authors demonstrate that their thermodynamic system has the desired qualitative properties. They prove that both the entropy and the ectropy are continuous (which is not automatic, and the proof makes very effective use of systems thinking, since it uses local controllability). They prove stability and asymptotic energy equipartition (nice!), and discuss irreversibility and the arrow of time. A nice result here is that their system satisfies Gibbs' principle, that states that in order for a state to be an equilibrium in an isolated system, it is necessary and sufficient that motions that do not alter the energy, should not increase the entropy. Finally, they discuss the feedback interconnection of two thermodynamic systems. The chapter ends with a discussion of the monotonicity of the energy function during transient solutions.

The fourth chapter of the book is a refinement of the third. In the interconnected system of the third chapter, the energies of the subsystems are equal to their temperatures. In the fourth chapter, this assumption is relaxed, and it is assumed that the energies are proportional to the temperatures, the proportionality constant being equal to the *specific heat* of the subsystem. An analysis, similar to the one performed in chapter 3, now leads to expressions for the entropy of the form

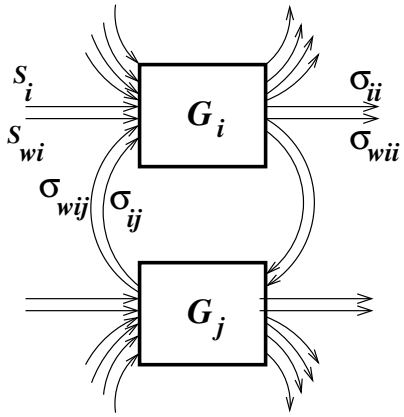
$$\mathfrak{S}(E) = \sum_{i=1}^{q} \mathfrak{S}_{\mathrm{i}}(E_i)$$

with

$$\mathfrak{S}_{\mathrm{i}}(E_i) = \frac{1}{\beta_i} \ln(\beta_i E_i + \mathrm{c}).$$

with the β_i the reciprocal of the specific heat of the i-th compartment. This structure is then applied to a closed system in which each of the individual compartments consists of an ideal gas separated by diathermal walls (walls through which energy, but matter can diffuse). They then recover the essential features of Boltzmann thermodynamics in a deterministic setting.

Until now, purely heat transfer phenomena were studied. Work (e.g. mechanical or electrical work) did not enter the analysis. In chapter 5, the compartmental system of chapter 3 is generalized to incorporate mechanical work in the form of changes in the volume of each of the compartments (see figure).



The equations now become slightly more complex, and involve, in addition to differential equations for the change of energy, also differential equations for the change of volume of each of the compartments. This means that the basic equation of chapter 3 is now complemented with the equation

$$\frac{d}{dt}V_i = \frac{d_{wi}(E,V) - S_{wi}}{E_i + \mathrm{c}}V_i$$

where d_{wi} denotes the rate of work done by the i-th subsystem on the environment, and S_{wi} the rate of work done by the environment on the i-th subsystem. This leads to classical expression for the rate of work done on the i-th subsystem equal to $P_i\frac{d}{dt}V_i$, with the pressure equal to $\frac{E_i+\mathrm{c}}{V_i}$. They prove conservation of energy, the existence of an internal energy function, the existence of a unique entropy function, and the second law. The entropy is now, up to a constant, given by

$$\mathfrak{S}(E,V) = \sum_{i=1}^{q} \mathfrak{S}_{\mathrm{i}}(E_i,V_i)$$

with

$$\mathfrak{S}_{\mathrm{i}}(E_i,V_i) = \ln(E_i+\mathrm{c}) + \ln V_i.$$

The presence of both heat transfer as well as work done on and by the environment of the thermodynamic system now allows to investigate the full range of thermodynamic phenomena, as far as the limitations in transforming heat into work are

concerned. In particular, they prove the equivalence of the Kelvin-Planck statement and the Clausius statement of the second law. The Kelvin-Planck formulation states that a process that completely transforms heat into work is impossible. The Clausius formulation states that a process whose only final result is to transport heat from a lower to a higher temperature is impossible. The equivalence of these statements are proven through the analysis of the efficiency of a Carnot cycle, that is, a cyclic process consisting of four regimes: beginning, from an initial state, with an adiabatic (no heat transfer with the environment) regime, followed by an isothermal (constant temperature) one, followed by again an adiabatic one, and then again an isothermal one, bringing the system back to the initial state.

In the next chapter, the system of chapter 3 is analyzed under the assumption that the dynamical equations are linear, leading to the differential equations

$$\frac{d}{dt}E = WE - DE + S,$$

with $E \in \mathbb{R}^q_+$ the vector of energy states, and $S \in \mathbb{R}^q_+$ the vector of heat supplies, $W \in \mathbb{R}^{q\times q}$ the matrix expressing the rate of heat transfer between the compartments, and $D \in \mathbb{R}^{q\times q}_+$ the diagonal matrix expressing the rate of heat dissipation. The analysis now leads to the theory of non-negative matrices, and special attention is paid to the case of strong coupling between the subsystems, i.e. when $\sigma_{ij} \to \infty$ in an appropriate sense.

In chapter 7, the system of chapter 3 is generalized to the case in which there are an infinite number of subsystems, parameterized by a spatial variable $x \in \mathfrak{V}$, with $\mathfrak{V}$ a compact connected subset of a finite dimensional real vector space with a smooth boundary. This leads to continuum thermodynamics, partial differential operators, and integral expressions over $\mathfrak{V}$ for the energy and entropy functions.

Chapter 8 contains the conclusions. It is of interest the list the main conclusions the authors draw from their work. In the context of the model from chapter 3, they reiterate the main postulates that went into their model:

(i) if the energies in connected subsystems are equal, energy exchange between these subsystems is not possible,

(ii) energy flows from subsystems with higher energy content to subsystems with lower energy content.

Using these assumptions the following conclusions were arrived at, and proven, using a rigorous theorem/proof format.

(i) Conservation of energy.

(ii) The energy in an isolated system is constant.

(iii) In an adiabatic regime, the entropy is nondecreasing,

(iv) and therefore tends to a maximum.

(v) In an isolated system, the energy tends to equipartition.

(vi) Although the total energy in an adiabatic regime is conserved, the usable energy is diffused.

(vii) A state is an equilibrium state of an isolated system if and only if states of equal energy do not have a larger entropy.

(viii) The entropy corresponding to zero temperature can be taken to be zero.

These conclusions are nicely summarized as follows.

1st Law: *You can't win, you can only break even.*

2nd Law: *You can break even only at absolute zero.*

3rd Law: *You can't reach absolute zero.*

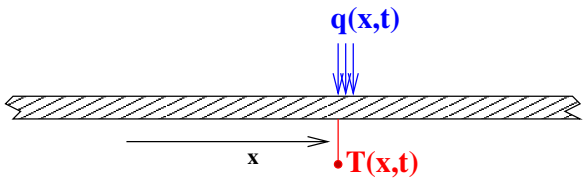
3 Some remarks from a personal perspective

This book review gives me an occasion to put forward a few personal views on systems theory and modeling of physical systems on the one hand, and thermodynamics and its relation to dissipative systems on the other hand.

3.1 The second law

The second law of thermodynamics is often presented as a sort of mystery. Of course, it is a deep law, with far reaching consequences, but it is not an enigma. And it certainly is of no help to introduce probability theory in order to explain something which in the end holds in our deterministic world. To the contrary, probability is bound to obfuscate the situation.

My own favorite example to illustrate the fact that there is something in nature beyond conservation of energy, is the exceedingly well-known diffusion equation model for heat transport in a uniform bar (see figure).



Using Fourier's law of heat conduction, or simple intuition, it is readily seen that the relation between the rate of heat exchanged with the environment (x is space, t

is time), $q(x,t)$ (chosen > 0 when heat is absorbed by the bar), and the temperature, $T(x,t)$, is given by the PDE (we assume that the units have been chosen to make the relevant constants are equal to 1)

$$\frac{\partial}{\partial t}T = \frac{\partial^2}{\partial x^2}T + q$$

Once we accept this equation as a description of reality, we can quickly arrive at a statement like the second law, as follows.

Assume that $T(\cdot,t)$ and $q(\cdot,t)$ have compact support for all $t \in \mathbb{R}$. It is easily seen that for all such (T,q) that satisfy the PDE, there holds:

$$\frac{d}{dt}\int_{-\infty}^{+\infty} T(x,t)\,dx = \int_{-\infty}^{+\infty} q(x,t)\,dx.$$

The right hand side is the power delivered to the bar at time t. Therefore $\int_{-\infty}^{+\infty} T(x,t)\,dx$ satisfies the requirement to be the stored energy. It is readily shown that it is the unique time function whose derivative along solutions of the PDE equals $\int_{-\infty}^{+\infty} q(x,t)\,dx$. Therefore *it is* the stored energy.

It requires only a little bit more effort to show that (T,q) also satisfies

$$\frac{d}{dt}\int_{-\infty}^{+\infty} \ln T(x,t)\,dx = \int_{-\infty}^{+\infty} \left(\frac{1}{T(x,t)}\frac{\partial}{\partial x}T(x,t)\right)^2 dx + \int_{-\infty}^{+\infty} \frac{q(x,t)}{T(x,t)}\,dx.$$

Whence

$$\frac{d}{dt}\int_{-\infty}^{+\infty} \ln T(x,t)\,dx \geq \int_{-\infty}^{+\infty} \frac{q(x,t)}{T(x,t)}\,dx.$$

Therefore $\int_{-\infty}^{+\infty} \ln T(x,t)\,dx$ satisfies the requirement to be the entropy. It can also be shown that it is the unique function whose time derivative is $\geq \int_{-\infty}^{+\infty} \frac{q(x,t)}{T(x,t)}\,dx$. Therefore the entropy *must be* $\int_{-\infty}^{+\infty} \ln T(x,t)\,dx$.

Now assume that we take the heated bar through a tortuous history starting at time t_{initial} in a temperature distribution $T(\cdot,t_{\text{initial}})$ and ending at time $t_{\text{terminal}} > t_{\text{initial}}$ in the same temperature distribution $T(\cdot,t_{\text{terminal}}) = T(\cdot,t_{\text{initial}})$. During the time interval $[t_{\text{initial}},t_{\text{terminal}}]$, all sorts of things could happen. At some time t and at some place x, $q(x,t)$ could be positive, at another place and the same time it could be negative, at another time and the same place it could be zero, etc. But, whatever happens, there will hold

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \left(\int_{-\infty}^{+\infty} q(x,t)\,dx\right) dt = 0,$$

and

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \left(\int_{-\infty}^{+\infty} \frac{q(x,t)}{T(x,t)}\,dx\right) dt \leq 0.$$

Now, it is easy to see that these two relations combined imply that

$$\max_{x\in\mathbb{R},t\in[t_{\text{initial}},t_{\text{terminal}}]}\{T(x,t)\mid q(x,t)>0\} \\ \geq \min_{x\in\mathbb{R},t\in[t_{\text{initial}},t_{\text{terminal}}]}\{T(x,t)\mid q(x,t)<0\}.$$

This is Clausius' version of the second law. It is appealing since it does not involve the entropy. But it does require to understand that $T(\cdot,t)$ is the state.

The equality

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}}\Big(\int_{-\infty}^{+\infty} q(x,t)\,dx\Big)\,dt = 0$$

states that the net effect of the (T,q) history is to transport exactly the same amount of heat from places and times where it is delivered by the environment to the bar to places and times where it is delivered by the bar to the environment. But the inequality

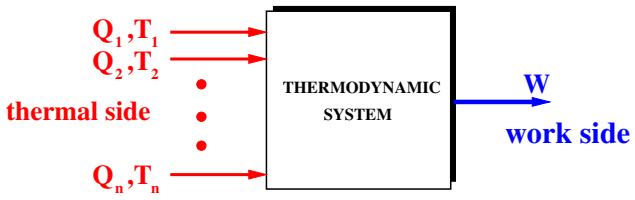
$$\max_{x\in\mathbb{R},t\in[t_{\text{initial}},t_{\text{terminal}}]}\{T(x,t)\mid q(x,t)>0\} \\ \geq \min_{x\in\mathbb{R},t\in[t_{\text{initial}},t_{\text{terminal}}]}\{T(x,t)\mid q(x,t)<0\}$$

cautions that the coldest point where heat flows into the bar cannot have a higher temperature that the hottest point where heat flows out of the bar. In other words, the bar cannot be used to transport heat from cold to hot.

3.2 Dissipative systems

The book is basically concerned with rather concrete physical systems, say interacting gasses or materials, or interconnected systems with simple subsystems and specific interactions. In these situations, the authors show how to construct the internal energy and the entropy uniquely. However, one of the main messages of thermodynamics is its generality: the laws apply just as well to something like a simple ideal gas, as to a complicated combination of electrical components, mechanical devices, and chemical reactions, as to the efficiency of a power station involving burners, boilers, turbines, condensers, generators, etc. Perhaps an abstract discussion in terms of 'blackboxes' could have helped in bringing out this generality.

Consider, as an abstract view, the situation described in the figure below.



This thermodynamic system has two sides. On the heat side, there are many terminals (for simplicity, we assume a finite number, n, of such terminals). Along the i-th terminal, heat is supplied to the thermodynamic system at a rate Q_i, with temperature

T_i, and at the work terminal, work is performed at a rate W. The arrow on the heat and work terminals signifies the positive direction of the heat flow: heat flow is counted positive if it flows from the environment into the system. Work is counted positive when it flows out of the system into the environment. Consequently, at any time, any of the Q_i's or the W could be positive, negative or zero. These arrows have nothing to do with inputs and outputs, as they are understood in systems theory. The chosen convention stems from the fact that one likes to think of a thermodynamic engine as a machine that transforms heat into work. But in a typical situation such an engine also has cooling terminals, where the heat flows out. The heat terminals could be places where an exothermal chemical reaction takes place, or where heat is supplied by transporting mass in or out, or where heat supplied through a heating coil, etc. The important assumption is that heat is always supplied at a particular temperature. It seems to be a physical law that *heat flow goes along with a temperature*. There cannot be one without the other.

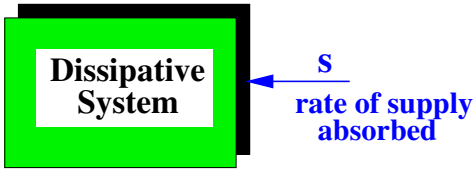
A typical thermodynamic engine will also have many work terminals, where work is done in the form of mechanical or electrical work, etc. However, in order to formulate the first and second law of thermodynamics, we do not need to distinguish between the different work terminals, and so, for simplicity, we have lumped them all into one. This lumping cannot be done on the thermal side, because of the required pairing of heat flow with temperature.

The internal dynamics of the thermodynamic system result in the fact that a family of trajectories

$$t \in \mathbb{R} \mapsto \big(W(t), Q_1(t), T_1(t), Q_2(t), T_2(t), \cdots, Q_n(t), T_n(t)\big) \in \mathbb{R} \times (\mathbb{R} \times \mathbb{R}_+)^n$$

are compatible with the laws of the engine. The totality of all such time trajectories is called the *behavior* of the engine. We denote it by $\mathscr{B}_{\text{thermodynamic}}$. Now, the first and second law make some universal statements about this behavior. Whatever the internal mechanism of the engine that leads to $\mathscr{B}_{\text{thermodynamic}}$, it will have to satisfy certain restrictions. Otherwise the dynamics that led to the behavior are a physical impossibility.

These restrictions are of course, the first and second law of thermodynamics. But it is not a trivial matter how to formulate them. As is often the case in mathematics, one can formulate a number of more or less equivalent versions of these laws, versions which can be shown to be equivalent under certain reasonable, but not compelling, conditions. In order to articulate these difficulties, it is best to backtrack even further, to the context of dissipative and conservative systems.



Consider the system shown in the figure above. Assume that it exchanges a real

valued quantity with its environment, at a rate s, counted positive when it flows into the system. This quantity is called the *supply rate*. The laws of the system allow a family of possible trajectories $s : \mathbb{R} \to \mathbb{R}$. Denote the set of all trajectories that are compatible with the laws of the system by $\mathscr{B}$. We also assume that the laws of the system do not change in time, i.e. that the system is time-invariant, formally that $s(\cdot) \in \mathscr{B}$ implies $s(\cdot + t) \in \mathscr{B}$ for all $t \in \mathbb{R}$.

When would we wish to call $\mathscr{B}$ dissipative? The answer is not evident: either we may want to impose restrictions directly on the behavior $\mathscr{B}$, or we may want to postulate the existence of a storage function (we will soon explain what we mean by a storage function), or something else.

A logical definition is obtained by putting restrictions on the *periodic* responses (only). Thus we arrive at the following definition. $\mathscr{B}$ is said to be *dissipative* if $s \in \mathscr{B}$ and s periodic, implies

$$\int_0^T s(t)\,dt \geq 0, \quad \text{where } T \text{ is the period.}$$

It is said to be *conservative* if instead

$$\int_0^T s(t)\,dt = 0.$$

The interpretation of the inequality is clear: in a dissipative system no supply is gained in a cyclic motion. In a conservative system, the account is balanced: all the supply that went in, came out.

We now turn to the storage function. This is defined as follows. Start from the behavior $\mathscr{B}$. Associate with it an extended behavior $\mathscr{B}_{\text{extended}}$ consisting of a time-invariant family of functions $(s,V) : \mathbb{R} \to \mathbb{R}^2$, such that after projection on the s variable, we get $\mathscr{B}$ back, i.e.

$$\mathscr{B} = \{s \mid \text{there exist } V \text{ such that } (s,V) \in \mathscr{B}_{\text{extended}}\}.$$

Call V a *storage function* if the *dissipation inequality*

$$V(t_{\text{terminal}}) - V(t_{\text{initial}}) \leq \int_{t_{\text{initial}}}^{t_{\text{terminal}}} s(t)\,dt$$

holds for all $(s,V) \in \mathscr{B}_{\text{extended}}$ and for all $t_{\text{initial}} < t_{\text{terminal}}$. In other words, the difference of the initial storage minus the final storage cannot exceed the supply absorbed from the initial to the final time.

Note that a quick unburdened application of the dissipation inequality along a periodic motion suggests that the existence of a storage function implies dissipativity. The converse seems more difficult, since it requires a clever construction of the extended system $\mathscr{B}_{\text{extended}}$. But, indeed, under mild conditions, it can be shown that a system *$\mathscr{B}$ is dissipative iff there exists a storage function that satisfies the dissipation inequality*, and *conservative iff there exists a storage function that satisfies the dissipation*

inequality with equality. Further, the storage function is in an essential way not unique (more than up to a additive constant) in the dissipative case, while in the conservative case, it is unique.

It would take us too far to spell out in this book review the mild conditions under which these statements hold. They have to do with

(i) *controllability*, ensuring the existence of 'enough' periodic trajectories, and

(ii) *observability* (of V from s).

These conditions may be termed mild, but they are by no means compelling.

Using these notions, we come to a formulation of the laws of thermodynamics as they apply to the abstract system introduced in the beginning of this subsection. The formulations ask for conservativity and dissipativity of $\mathscr{B}_{\text{thermodynamic}}$, as follows.

1. $\mathscr{B}_{\text{thermodynamic}}$ is conservative with respect to the supply rate $\sum_{i=1}^{n} Q_i - W$.

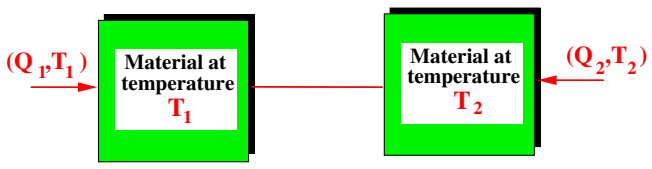
2. $\mathscr{B}_{\text{thermodynamic}}$ is dissipative with respect to the supply rate $-\sum_{i=1}^{n} \frac{Q_i}{T_i}$.

The associated storage functions are respectively the *internal energy* and the *negative of the entropy*.

We reiterate that as far as the definitions are concerned, the statements in terms of periodic trajectories are but one choice. One could equally well focus on storage functions in the definition, with perhaps with more restrictions imposed on them than we have done. Or we could assume an equilibrium, and focus on trajectorier from and to this equilibrium. But there are also formulations possible that exploit the fact that in thermodynamics two related dissipation laws are considered at once, etc.

3.3 Interconnected systems

One of the important features of dissipativity is its behavior under interconnection. We illustrate this by means of a very simple example of a system of the type considered in chapter 3 of the monograph. Consider two interconnected vessels, filled with some material, shown in the figure below.



The vessels, respectively at temperature T_1 and T_2, receive heat from the environment at these temperatures T_1 and T_2, and rates Q_1 and Q_2. In addition there is heat diffusion from vessel 1 to vessel 2 at rate $T_1 - T_2$ (this may hence be positive or negative). Under reasonable and intuitive assumptions (and choice of physical constants),

the relation between the T_1, T_2, Q_1, Q_2 is

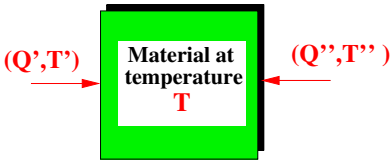
$$\frac{d}{dt}T_1 = Q_1 - T_1 + T_2, \quad \frac{d}{dt}T_2 = Q_2 + T_1 - T_2.$$

A simple calculation shows that

$$\begin{aligned} \frac{d}{dt}(T_1 + T_2) &= Q_1 + Q_2, \\ \frac{d}{dt}(\ln T_1 + \ln T_2) &= \frac{Q_1}{T_1} + \frac{Q_2}{T_2} + \left(\frac{1}{T_1} - \frac{1}{T_2}\right)^2 \geq \frac{Q_1}{T_1} + \frac{Q_2}{T_2}. \end{aligned}$$

This shows that the interconnected system obeys the first and second law with internal energy $T_1 + T_2$ and entropy $\ln T_1 + \ln T_2$.

The question occurs if we can view this system as the interconnection of two systems, both satisfying by themselves the laws of thermodynamics, and which, with the appropriate interconnection constraints, lead to the interconnected system. This leads us to consider the vessel shown in the figure below.



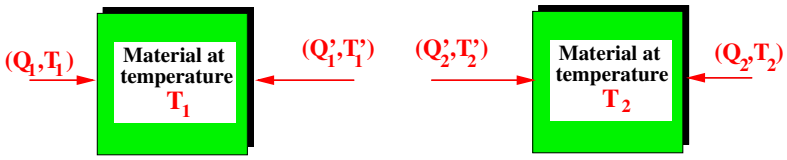
Under what relations between Q', T', Q'', T'', and T does this system obey the laws of thermodynamics? If we take as dynamics

$$\frac{d}{dt}T = Q' + Q'',$$

we obtain right away the first law with internal energy T. But the second law is not automatically satisfied. It is, if we take $T' \geq T$ and $T'' \geq T$, for then

$$\frac{d}{dt}\ln T = \frac{Q' + Q''}{T} \geq \frac{Q'}{T'} + \frac{Q''}{T''}.$$

The condition $T' \geq T$, $T'' \geq T$ means that we can bring in hot heat, but we cannot bring in cold heat. This simple example, which in fact done with a single terminal, points to the essence of thermodynamics. Bringing in heat from the outside does not violate the law of energy conservation. But, it is impossible to bring in heat at a temperature that is colder than the temperature of the system. This violates the second law.



We may obtain the original system by interconnecting the vessels as shown in the figure above. Take the relations between the variables Q_1, T_1, Q_1', T_1' of the first vessel to be

$$\begin{aligned} \frac{d}{dt}T_1 &= Q_1 + Q_1', \\ T_1' &= T_1 \qquad \text{for } Q_1' \leq 0, \\ T_1' &= T_1 - Q_1' \qquad \text{for } Q_1' \geq 0. \end{aligned}$$

Similarly, take the relations between the variables Q_2, T_2, Q_2', T_2' of the second vessel to be

$$\begin{aligned} \frac{d}{dt}T_2 &= Q_2 + Q_2', \\ T_2' &= T_2 \qquad \text{for } Q_2' \leq 0, \\ T_2' &= T_2 - Q_2' \qquad \text{for } Q_2' \geq 0. \end{aligned}$$

Next, verify that the interconnection laws

$$T_1' = T_2' \quad \text{and} \quad Q_1' + Q_2' = 0$$

lead to the correct equations

$$\frac{d}{dt}T_1 = Q_1 - T_1 + T_2, \quad \frac{d}{dt}T_2 = Q_2 + T_1 - T_2.$$

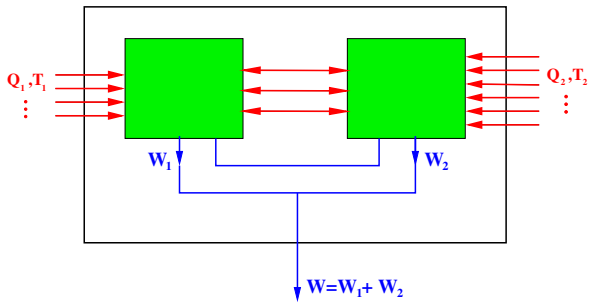
From the earlier analyis, we can conclude that both vessels satisfy the first and second law. This yields

$$\begin{aligned} \frac{d}{dt}T_1 = Q_1 + Q_1', \quad \frac{d}{dt}\ln T_1 \geq \frac{Q_1}{T_1} + \frac{Q_1'}{T_1'} \\ \frac{d}{dt}T_2 = Q_2 + Q_2', \quad \frac{d}{dt}\ln T_2 \geq \frac{Q_2}{T_2} + \frac{Q_2'}{T_2'}. \end{aligned}$$

Adding and using the interconnection constraints yields

$$\frac{d}{dt}(T_1 + T_2) = Q_1 + Q_2, \quad \frac{d}{dt}(\ln T_1 + \ln T_2) \geq \frac{Q_1}{T_1} + \frac{Q_2}{T_2}.$$

These are the first and second law as they pertain to the original interconnected system.



Conditions that an interconnection of systems (see the figure above) that satisfy the laws of thermodynamics yields an interconnected system that also satisfies the laws of thermodynamics, are readily obtained. The basic constraints are similar to those which we were put in evidence in the simple example given above. For each interconnected thermal terminal the interconnection constraints should imply equal temperatures and that the heat flows sum to zero. The work interconnections should imply that the work performed by the interconnected system is equal sum of the work performed by the subsystems. There could, of course, be all kinds of mechanical interactions between the subsystems, but these should be neutral, meaning work in one equal work out the other.

It is easy to see that this leads to an interconnected system with as total internal energy and entropy, the sum of the internal energies and entropies of the subsystems. In thermodynamic parlance, this states that the energy and entropy are *extensive quantities*, just as volume and mass and charge, in contrast with *intensive quantities*, as temperature and voltage and position. This extensive property of entropy has important consequences as far as the calculation of the entropy is concerned. Zooming in on simple subsystems often even yields uniqueness of the entropy function, by tearing the interconnected system into subsystems, each of which have a unique entropy function.

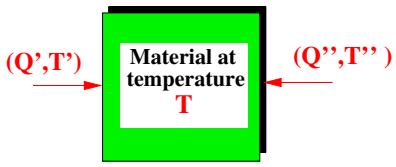
3.4 Inputs, outputs, and states

Throughout the 20-th century, mainstream system theory has been developed in an input/output mode of thinking. Starting with the work of Heaviside, via the impedance description of circuit theorists, to the stimulus/response view of Wiener, generations of system theorists have been trained to think of a system as an input/output map. This point of view is still very prevalent in, for example, system identification, where the statement that a system is an input/output map is commonplace. Of course, the fact that usually their own first example contradicts this statement, does not bother these writers: to state something that is almost correct, not to say wrong, rather that something that is correct, is usually regarded as good pedagogy.

Of course, a system is patently not an input/output map. For all but the most simplistic examples, the output also depends on the initial conditions, and it is often the response to the initial conditions that is of main concern. The fact that initial conditions in the form of state variables are automatically incorporated in state models is for sure one of the main reasons of their deep influence in the field. As such, I consider Kalman's input/state/output framework to be the first model structure that is adequate for the analysis of a reasonably general class of physical systems. The authors of the monograph under review are clearly adherents of this point of view. And indeed, it is the use of the input/state/output setting that has enabled them to present their rigorous theory of thermodynamics.

Nevertheless, the input/output partition of the variables of interest is often hard to maintain from a physical point of view. Why should it be a unviversal fact that some variables act as causes, and some as effects? The input/output picture may be

appropriate for signal processing, but a physical system is not a signal processing. A law of physics states that certain outcomes are compatible, that certain values of physical variables can occur simultaneously, but not that one causes the other.

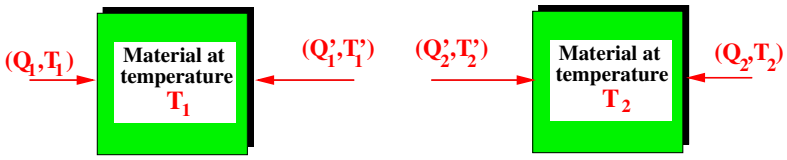


Consider, as an example, the simple system discussed in subsection 3.3, shown again in the figure above. With the dynamics

$$\frac{d}{dt}T = Q' + Q'', \;\; T' \geq T, \;\; T'' \geq T,$$

this defines, as we have seen, a thermodynamic system. The external system variables are Q', T', Q'', T''. They are all, in a sense, free inputs, but only to some extent. T' and T'' are constrained not to be smaller than T, Q' cannot make T become larger than T'', etc. The question of what causes should not be posed. We arrive at the same conclusion when we consider the heated bar discussed in subsection 3.1. Of course, the orthodox systems theory point of view is to consider $q(\cdot,t)$ the input, and $T(\cdot,t)$ the output and the state. But in some region along the bar, the heat radiation could be caused by the temperature. In physical systems, there are certain variables which the model aims at, but there is no point is insisting on a partition of these variables in inputs and outputs.

The drawback of input/output thinking comes forward very pointedly when considering interconnected systems. The view that interconnections should be modelled as an input-to-output assignment is contradicted by almost all physical examples. Consider again the system discussed in subsection 3.3.



As we have seen, the interconnection law that governs the interconnection of the two vessels shown above is

$$T_1' = T_2', \qquad Q_1' + Q_2' = 0.$$

So if, for some reason, we have decide to consider T_1' an input and Q_1' an output for the first system, and, likewise, by symmetry considerations, T_2' an input and Q_2' an output for the second system, we see that the interconnection law demands equating two inputs and putting the sum of two outputs equal to zero. Exactly what is forbidden in the usual input/output thinking. It turns out that this situation, equating similar

variables (pressures, positions, voltages, etc.) and putting the sum of similar variables (flows, forces, currents, etc.) equal to zero, is the rule in physical interconnections, and the input-to-output assignment is the exception. Interconnection of systems physical systems means variable sharing, not signal transmission.

In the subsection 3.2, we have seen that, in order to discuss dissipative systems, is very reasonable to consider the behavior that consists of the possible supply rate trajectories $s : \mathbb{R} \to \mathbb{R}$. The question of how supply flows in and out leads to a 'no frills' definition of dissipativity. Obviously, the question if s is an input or an output is absurd. In its very essence, the situation in thermodynamics is precisely this one: it is a theory that studies the behavior defined by all trajectories

$$(s_1, s_2) : \mathbb{R} \to \mathbb{R}^2,$$

with $s_1 = \sum_{i=1}^n Q_i - W$, and $s_2 = \sum_{i=1}^n \frac{Q_i}{T_i}$, and the W's and (Q_i, T_i)'s constrained by $\mathscr{B}_{\text{thermodynamic}}$. Asking if s_1 or s_2 is an input or an output is again absurd.

We have also seen in subsection 3.2 that it is not necessary to introduce a state in order to discuss storage functions. But, it is a good question to ask whether there always exists a storage function that is state function.

4 Conclusions

Thermodynamics is, by its very essence, a theory of *open systems*[1]. It puts limitations to the way in which physical systems are able to exchange energy and heat with their environment. Notwithstanding the fact that systems and control theory has grown into *the* field that deals with open systems in a fundamental way, there have been very few publications that deal with thermodynamics from a modern systems theory perspective. The monograph under review appears to be the first book to do so. As such, it is a most welcome contribution.

Thermodynamics is also a theory of *interconnected systems*. An essential feature is that if we combine simple physical systems that individually satisfy the laws of thermodynamics, we obtain a more complex system that also obeys these laws. This is a recurrent theme in this book. The authors invariably call what I would call an interconnected system a 'large scale system', a forgivable, albeit somewhat irritating, concession to sponsored research themes, I suppose.

[1]Thermodynamicists and systems and control theorists differ in what they mean by open and closed. In thermodynamics, it is common to call systems that exchange matter and energy with their environment, *open*. Systems that exchange energy but not matter, are called *closed*, and those that exchange neither energy nor matter are called *isolated*. In systems and control theory, on the other hand, a closed system is, very roughly speaking, one whose past trajectory defines the future trajectory uniquely. Closed systems can be described by a flow $\frac{d}{dt}x = f(x)$, combined with, perhaps, an output equation. This is more akin to what thermodynamicists call an isolated system. In the conventional input/output thinking, a closed system is one that evolves without inputs, while an open system is one that is influenced by external inputs. We use 'open' and 'closed' in the systems and control sense.

Hence, both *open* and *connected*, the features that make systems theory into a discipline of its own, are key elements of this book. As such, this monograph makes a very substantial contribution to the field. Not only by the originality of the topic, the approach, and the results, but also by the systems point of view as an approach to thermodynamics.

In my opinion, the main shortcoming of this monograph is the lack of concrete physical examples. Of course, most readers will have no difficulty to construct some, but I do not think that this should have been left to the readers to do. I believe that the basic set-up in chapter 3 could haven been much clarified by considering for example ideal gasses in their proverbial vessels, each governed by $PV = RNT$, with P the pressure, V the volume, T the temperature, N the number of moles of the gas, and R the universal gas constant. By letting these vessels be in thermal contact with each other and with the environment, one would have had a nice concrete example of the situation covered in chapters 3 and 4. By letting the vessels also be in mechanical contact, influencing each other's volumes and pressures, one could have obtained a good example of how to visualize the situation covered in chapter 5. For chapter 7, heat diffusion in a (uniform and non-uniform) bar would have been a good example.

The book takes the orthodox pedagogical approach in explaining the laws of thermodynamics by going from the simple to the complex: first, heat transport in a finite number of compartments with identical substances, then heat transport with non-identical substances, then heat transport combined with work, and finally an infinite number of compartments. I would have topped this off with a fully abstract discussion of thermodynmaics in the context of dissipative systems and interconnections, along the lines of what I pointed to in sections 3.2 and 3.3.

This book is also a courageous one. It comes in a time that research is dominated by impact factors, citation analysis, and what have you. As such, a book that is not along the beaten path of the newest research themes, and deals with a classical poorly understood subject, is very welcome. The authors of this monograph should be recommended in their aim to explain an important domain as thermodynamics from a systems theory point of view to the community.