Book Reviews

Thermodynamics: A Dynamical Systems Approach—W. M. Haddad, V. S. Chellaboina, and S. Nersesov (Princeton, NJ: Princeton Univ. Press, Princeton Series in Applied Mathematics, 2005). *Reviewed by Jan C. Willems*

I. 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 coal 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 humbly 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* (New York: Springer Verlag, 1980), he calls thermodynamics "*a dismal swamp of obscurity*." Elsewhere, Truesdell, in despair of trying to make sense of the writings of some local heros as De Groot, Mazur, Casimir, and Prigogine, suspects that there is "*something rotten in the (thermodynamic) state of the Low Countries*" (see page 134 of *Rational Thermodynamics*, New York: 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 it a function of, i.e., what is its domain? Is it uniquely defined?

ii) The strange use of derivatives, with differentiation often applied to functions whose domain is unspecified, or with respect to variables that do not belong to the domain. As Truesdell notes, in thermodynamics even derivatives look different, and statements like

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

are not uncommon. 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.

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 tough, the tough 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 nonbelievers, 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[1] chose to use the term "entropy" for "amount of information," this was like pouring 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

The reviewer is with K. U. Leuven, B-3001 Leuven, Belgium (e-mail: Jan.Willems@esat.kuleuven.be).

Digital Object Identifier 10.1109/TAC.2006.878567

[1]My greatest concern was what to call it. I thought of calling it "information," but the word was overly used, so I decided to call it "uncertainty." John von Neumann had a better idea, he told me,"*You should call it entropy, for two reasons. In the first place, your uncertainty function goes by that name in statistical mechanics. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage*." (Claude Shannon, as quoted in M. Tribus and E. C. McIrvine, Energy and Information, *Scientific American*, 224, Sep. 1971, pp. 178–184.)

0018-9286/$20.00 © 2006 IEEE

compartments). The construction of the internal energy and the entropy is solidly founded on the theory of (cyclo-)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.

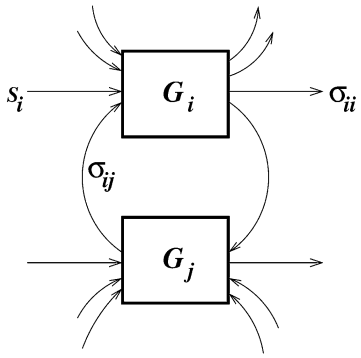
There have been previous attempts to give thermodynamics a solid mathematical underpinning. One notable program that set this as the goal is the work by the school of Noll, Coleman, Gurtin, *et al.* documented in a series of publications in the journal *Archive for Rational Mechanics and Analysis* in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

II. 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. The nomenclature around dynamical systems ("flows") of the form $(d/dt)z(t) = w(z(t))$ is given. Of special interest are state spaces $\mathbb{R}^q_+$ that are non-negative orthants of finite dimensional vector spaces. Various Lyapunov stability theorems are proven. 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 spaces, for example, are assumed to be a Banach space, but the dynamics are fully nonlinear). The notation is somewhat heavy, and many of the definitions are rather involved. 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 proceeding with the main subject matter. However, it is, in my opinion, never a reader-friendly idea to frontload a book with a chapter devoted to 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 setup 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 ith 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 a little bit puzzled by some elements in this basic model. For example, I would have simply replaced S_i by $S_i - \sigma_{ii}$. The fact that the authors have both terms separately in the equations is used in their stability proofs. However, I feel that it is a bit artificial from the physical point of view. It may also have to do with their input/output point of view, since the separation allows to think of S_i as an input, and of σ_{ii} as an output. I will come back to this point in Subsection 3-D. In the same vein, I did not appreciate 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 a bit 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 have stuck with $\sigma_{ij}(E_i, E_j)$. 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

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

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 (cyclo-)dissipative systems is used to show the existence of a state function$E \mapsto \mathfrak{S}(E)$, the entropy function, satisfying

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

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 additive constant, and that it is given by

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

with

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

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

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

The authors also 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. An interesting result here is that their system satisfies Gibbs' principle, which 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 an analysis of the monotonicity of the energy function during transient motions.

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

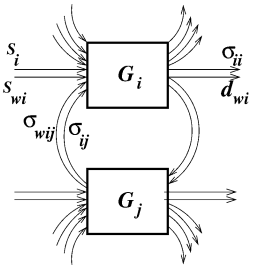
$$\mathfrak{S}(E) = \sum_{i=1}^{q} \mathfrak{S}_i(E_i)$$

with

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

with the β_i the reciprocal of the specific heat of the ith 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 can, but matter cannot 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 the figure below).



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 ith subsystem on the environment, and S_{wi} the rate of work done by the environment on the ith subsystem. This leads to classical expression for the rate of work done on the ith subsystem equal to $P_i(d/dt)V_i$, with the pressure equal to $(E_i + \mathrm{c}/V_i)$. 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

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

with

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

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 to 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.

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) 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 cannot win, you can only break even.
- 2nd Law: You can break even only at absolute zero.
- 3rd Law: You cannot reach absolute zero.

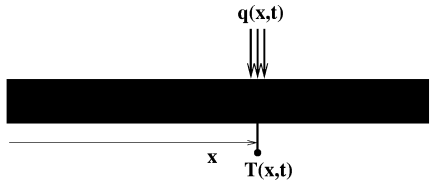
III. Remarks

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 dissipative systems and its relation to thermodynamics on the other hand.

A. The Second Law

The second law of thermodynamics is often presented as some sort of mystery. Surely, it is a deep law, with far reaching consequences, but it is not an enigma. And it certainly is of no help to introduce the presumed probabilistic behavior of the micromicroworld in order to explain things which in the end hold in our deterministic macroscopic world. To the contrary, also here, 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 the figure below).



Using Fourier's law of heat conduction, or simple intuition, it is readily seen that a reasonable model for 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

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

where ρ is (proportional to) the specific heat coefficient of the material, and γ the heat diffusion coefficient. 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 the length of the bar is L, that the temperature at the ends is fixed to T_0, and there is no heat transport at the ends. This leads to the boundary conditions

$$T(\cdot,0) = T(\cdot,L) = T_0, \quad \frac{\partial}{\partial x}T(\cdot,0) = \frac{\partial}{\partial x}T(\cdot,L) = 0.$$

Assume that the units have been chosen such that $\rho = 1, \gamma = 1, L = 1$.

It is easily seen that for all $(T,q) : \mathbb{R}\times[0,1] \to \mathbb{R}_+ \times \mathbb{R}$ that satisfy the PDE and boundary conditions, there holds

$$\frac{d}{dt}\int_0^1 T(x,t)\,\mathrm{dx} = \int_0^1 q(x,t)\,dx.$$

The right-hand side $\int_0^1 q(x,t)\,\mathrm{dx}$ is the power delivered to the bar at time t. Therefore, $\int_0^1 (T(x,t)\,\mathrm{dx}$ satisfies the requirement to be the stored energy. It is readily shown that it is, up to an additive constant, the unique time function whose derivative along solutions equals $\int_0^1 q(x,t)\,\mathrm{dx}$. Therefore, *it is* the stored energy.

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

$$\frac{d}{dt}\int_0^1 \ln T(x,t)dx = \int_0^1 \left(\frac{1}{T(x,t)}\frac{\partial}{\partial x}T(x,t)\right)^2 dx + \int_0^1 \frac{q(x,t)}{T(x,t)}dx.$$

Where

$$\frac{d}{dt}\int_0^1 \ln T(x,t)\,dx \geq \int_0^1 \frac{q(x,t)}{T(x,t)}\,dx.$$

Therefore, $\int_0^1 \ln T(x,t)dx$ satisfies the requirement to be the entropy. It can also be shown that it is, up to an additive constant, the unique

function whose time derivative is $\geq \int_0^1 (q(x,t)/T(x,t))dx$. Therefore $\int_0^1 \ln T(x,t)dx$ *must be* the entropy.

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 along the bar, $q(x,t)$ 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

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \left(\int_0^1 q(x,t)dx \right) dt = 0,$$

and

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \left(\int_0^1 \frac{q(x,t)}{T(x,t)}dx \right) dt \leq 0.$$

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

$$\max_{x\in[0,1], t\in[t_{\text{initial}}, t_{\text{terminal}}]} \{T(x,t) | q(x,t) > 0\} \geq \min_{x\in[0,1], t\in[t_{\text{initial}}, t_{\text{terminal}}]} \{T(x,t) \mid q(x,t) < 0\}.$$

This is Clausius' version of the second law. It is appealing since it does not involve the entropy.

The equality

$$\int_{t_{\text{initial}}}^{t_{\text{terminal}}} \left(\int_0^1 q(x,t)dx \right) dt = 0$$

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. Energy, heat in this example, is merely redistributed. However, the inequality

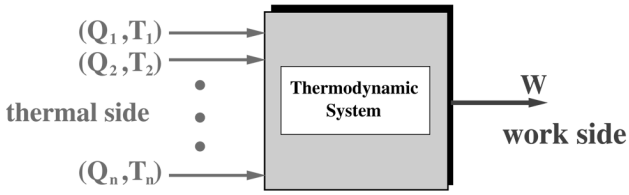
$$\max_{x\in[0,1], t\in[t_{\text{initial}}, t_{\text{terminal}}]} \{T(x,t) | q(x,t) > 0\} \geq \min_{x\in[0,1], t\in[t_{\text{initial}}, t_{\text{terminal}}]} \{T(x,t) \mid q(x,t) < 0\}$$

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.

B. 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 or a uniform material, as to a complicated combination of electrical apparatus, mechanical devices, thermal components, 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 following figure.



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 ith such 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 arrows on the heat and work terminals indicate 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. For example, a plant that burns coal to boil water into steam under pressure that spins a turbine that drives an electric generator that produces electrical power. However, in a typical situation such an engine also has cooling terminals, where the heat flows out. In fact, thermodynamics obliges heat to flow out at some places.

The heat terminals could be places where an exothermal chemical reaction takes place, or where heat is supplied by transporting mass in and 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 only a certain family of trajectories

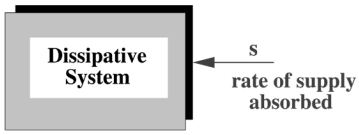
$$t \in \mathbb{R} \mapsto (W(t), Q_1(t), T_1(t), Q_2(t), T_2(t), \ldots, Q_n(t), T_n(t)) \in \mathbb{R} \times (\mathbb{R} \times \mathbb{R}_+)^n$$

is 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 $\mathcal{B}_{\text{thermodynamic}}$.

It may not be a sinecure to come up with a representation of $\mathcal{B}_{\text{thermodynamic}}$ in the form of, say, a system of differential equations. But the laws of thermodynamics allow us to make some universal statements about this behavior. Whatever the internal mechanism of the engine that leads to $\mathcal{B}_{\text{thermodynamic}}$ is, it will have to satisfy certain universal 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. However, it is not a trivial matter how to formulate them. As is often the case in mathematics, one can formulate a number of versions of these laws, versions that can be shown to be more or less 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 below.



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 $s : \mathbb{R} \to \mathbb{R}$, expressing how the system exchanges supply with its environment. Denote the set of all trajectories that are compatible with the laws of the system by $\mathcal{B}$. We also assume that the laws of the system do not change in time, i.e., that the system is time-invariant, formally that $s(\cdot) \in \mathcal{B}$ implies $s(\cdot + t) \in \mathcal{B}$ for all $t \in \mathbb{R}$.

When would we wish to call $\mathcal{B}$ dissipative? The answer is not evident: either we may want to impose restrictions directly on the behavior $\mathcal{B}$, or we may want to postulate the existence of a storage (we will soon explain what we mean by a storage), or something else. What restriction does dissipativeness impose on $\mathcal{B}$? A logical definition is obtained by putting restrictions on the *periodic* responses (only). Thus, we arrive at the following definition. $\mathcal{B}$ is said to be *dissipative* if $s \in \mathcal{B}$ and s periodic imply

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

where T 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 can ever be 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. This is defined as follows. Start from the behavior $\mathcal{B}$. Associate with it an extended behavior $\mathcal{B}_{\text{extended}}$ consisting of a time-invariant family of maps $(s, V) : \mathbb{R} \to \mathbb{R}^2$, such that after projection on the s variable, we get $\mathcal{B}$ back, i.e.,

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

Call V a *storage* 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 $(s, V) \in \mathcal{B}_{\text{extended}}$ 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 during the time interval starting at the initial time and extending to the terminal time.

Note that a quick unburdened application of the dissipation inequality along a periodic motion (s, V) suggests that the existence of a storage implies dissipativity. The converse seems more difficult, since it requires a clever construction of the extended behavior $\mathcal{B}_{\text{extended}}$. And indeed, under mild conditions, it can be shown that a system *$\mathcal{B}$ is dissipative iff there exists an extended behavior $\mathcal{B}_{\text{extended}}$ with a storage that satisfies the dissipation inequality,* and *conservative iff there exists a storage that satisfies the dissipation inequality with equality*. Further, for all but very simple systems, the storage is in an essential way not unique (more than up to an 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 this equivalence holds. They have to do with

i) *controllability*, ensuring the existence of 'enough' periodic trajectories, so that periodic motions become representative of the whole behavior $\mathcal{B}$, and
ii) *observability* (of V from s, or something like that, so that properties of $\mathcal{B}$ can be lifted to $\mathcal{B}_{\text{extended}}$).

These conditions may be termed mild, but they are not compelling.

Using these notions of dissipative and conservative, 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 $\mathcal{B}_{\text{thermodynamic}}$, as follows.

i) $\mathcal{B}_{\text{thermodynamic}}$ is conservative with respect to the supply rate $(\sum_{i=1}^n Q_i) - W$.
ii) $\mathcal{B}_{\text{thermodynamic}}$ is dissipative with respect to the supply rate $-\sum_{i=1}^n (Q_i/T_i)$.

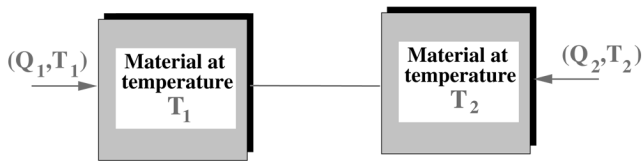
The associated storages are respectively the *internal energy* and the *negative of the entropy*.

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 the storage in the very definition of dissipativeness, with perhaps more restrictions imposed on them than we have done. Or we could assume an equilibrium, and focus on trajectories from and to this equilibrium. Or one could depart from a set of "observables," including or implying the supply rate, and demand dissipation along periodic motions of these observables. However, there are also formulations possible that exploit the fact that in thermodynamics two related supply rates, $(\sum_{i=1}^n Q_i) - W$ and $-\sum_{i=1}^n (Q_i/T_i)$, are considered both at once, etc.

Note that what I have just called a "dissipative" system is what the book under review, and I elsewhere, call a "cyclo-dissipative" system. The difference in the end has to do with the question if we want the storage to be non-negative (more precisely, bounded from below). Is a storage necessarily bounded from below? It seems not. For instance, since entropy is often a logarithm, we will not have boundedness from above or below for entropy. How about energy? Electrical engineers think of energy as something that is non-negative, and rightfully object to calling a negative inductor "passive," even though it looks passive in a periodic regime. However, in mechanics, energy is often not bounded from below (consider, for example, a particle or a planet orbiting in an inverse square law potential field). When the theory of dissipative systems is employed in stability analysis, non-negativity of the storage is natural. But sign definiteness of the storage is a subtle matter from a physical point of view, and should certainly not universally be adopted.

C. 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 an interconnection of two systems of the type considered above, in the spirit of what the authors dealt with in Chapter 3 of the monograph. Consider two interconnected vessels, as shown in the figure below.



The vessels, respectively at temperatures 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 proportional to $T_1 - T_2$ (this may, hence, be positive or negative). Under reasonable and intuitive assumptions, the relation between the T_1, T_2, Q_1, Q_2 can be taken to be

$$\rho_1 \frac{d}{dt} T_1 = Q_1 - \gamma(T_1 - T_2)$$
$$\rho_2 \frac{d}{dt} T_2 = Q_2 - \gamma(T_2 - T_1)$$

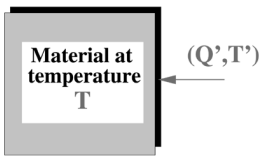
with ρ_1, ρ_2 are (proportional to) the specific heat coefficients of the material in the vessels, and γ is the heat diffusion coefficient between the vessels. A simple calculation shows that

$$\frac{d}{dt}(\rho_1 T_1 + \rho_2 T_2) = Q_1 + Q_2,$$
$$\frac{d}{dt}(\rho_1 \ln T_1 + \rho_2 \ln T_2) = \frac{Q_1}{T_1} + \frac{Q_2}{T_2} + \gamma T_1 T_2 \left(\frac{1}{T_1} - \frac{1}{T_2}\right)^2 \geq \frac{Q_1}{T_1} + \frac{Q_2}{T_2}.$$

This shows that the interconnected system obeys the first and second law with internal energy $(\rho_1 T_1 + \rho_2 T_2)$ and entropy $(\rho_1 \ln T_1 + \rho_2 \ln T_2)$.

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, correspond to the interconnected system.

Before getting into interconnections, consider the simple system shown below.



This system has only one heat terminal, that feeds into a vessel at temperature T. Assume that the dynamical equation is

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

where ρ is (proportional to) the specific heat coefficient of the material in the vessel.

Is this a thermodynamic system? $E = \rho T$ leads to the first law with internal energy ρT, unique up to an additive constant. It requires a bit more effort to show that the second law imposes to the condition

$$[\![Q' \geq 0 \quad \text{and} \quad T' \geq T]\!] \quad \text{or} \quad [\![Q' \leq 0 \quad \text{and} \quad T' \leq T]\!]$$

and leads to $S = \rho \ln T$ for the entropy (again unique up to an additive constant). For then

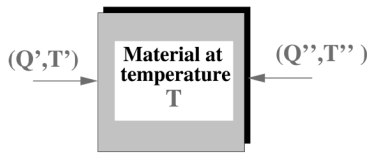
$$\frac{d}{dt} S = \rho \frac{d}{dt} \ln T = \frac{Q'}{T} \geq \frac{Q'}{T'}.$$

The complete dynamical equations are, therefore

$$\rho \frac{d}{dt} T = Q',$$
$$[\![Q' \geq 0 \quad \text{and} \quad T' \geq T]\!] \quad \text{or} \quad [\![Q' \leq 0 \quad \text{and} \quad T' \leq T]\!].$$

So, in addition to an equation expressing the rate of change of the temperature, the second law imposes the impossibility to transport heat from cold to hot. This simple example points to the essence of thermodynamics. Bringing in heat from the outside at any temperature does not violate the law of energy conservation. However, it is impossible to bring in heat at a temperature that is colder than the temperature of the vessel. This violates the second law.

In order to deal with the interconnection, consider first the vessel shown in the figure below, a simple generalization to two heat terminals of the vessel with one heat terminal considered below.

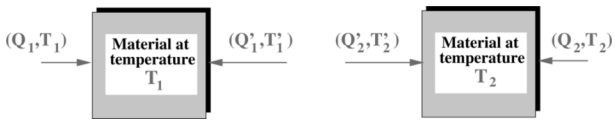


Take as dynamics

$$\rho \frac{d}{dt} T = Q' + Q''$$
$$[\![Q' \geq 0 \quad \text{and} \quad T' \geq T]\!] \quad \text{or} \quad [\![Q' \leq 0 \quad \text{and} \quad T' \leq T]$$
$$[\![Q'' \geq 0 \quad \text{and} \quad T'' \geq T]\!] \quad \text{or} \quad [\![Q'' \leq 0 \quad \text{and} \quad T'' \leq T]\!].$$

Then, with internal energy $E = \rho T$ and entropy $\rho \ln T$, the first and second laws follow.

We obtain the original interconnected system by interconnecting two such vessels as shown in the figure below.



Take the relations between the variables Q_1, T_1, Q_1', T_1' of the first vessel to be

$$\rho_1 \frac{d}{dt} T_1 = Q_1 + Q_1'$$
$$[\![Q_1' \geq 0 \quad \text{and} \quad Q_1' = \gamma(T_1' - T_1)]\!] \quad \text{or}$$
$$[\![Q_1' \leq 0 \quad \text{and} \quad T_1' = T_1]\!].$$

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

$$\rho_2 \frac{d}{dt} T_2 = Q_2 + Q_2'$$
$$[\![Q_2' \geq 0 \quad \text{and} \quad Q_2' = \gamma(T_2' - T_2)]\!] \quad \text{or}$$
$$[\![Q_2' \leq 0 \quad \text{and} \quad T_2' = T_2]\!].$$

Next, verify that the interconnection laws

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

lead, after elimination of T_1', T_2', Q_1', Q_2', to the correct equations

$$\rho_1 \frac{d}{dt} T_1 = Q_1 - \gamma(T_1 - T_2), \quad \rho_2 \frac{d}{dt} T_2 = Q_2 - \gamma(T_2 - T_1)$$

for the interconnected system.

From the earlier analysis, we may conclude that both vessels individually satisfy the first and second law. This yields

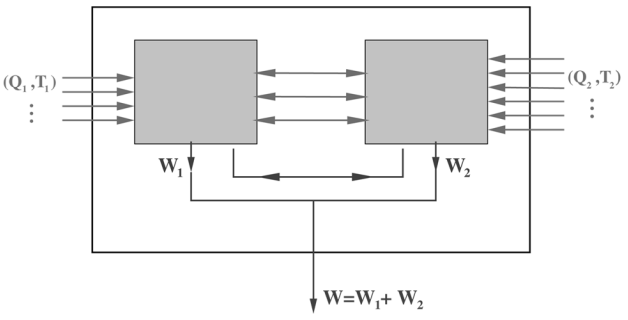
$$\begin{aligned} \rho_1 \frac{d}{dt} T_1 &= Q_1 + Q_1' \\ \frac{d}{dt} \rho_1 \ln T_1 &\geq \frac{Q_1}{T_1} + \frac{Q_1'}{T_1'} \\ \rho_2 \frac{d}{dt} T_2 &= Q_2 + Q_2' \\ \frac{d}{dt} \rho_2 \ln T_2 &\geq \frac{Q_2}{T_2} + \frac{Q_2'}{T_2'}. \end{aligned}$$

Adding and using the interconnection constraints yields

$$\begin{aligned} \frac{d}{dt}(\rho_1 T_1 + \rho_2 T_2) &= Q_1 + Q_2, \\ \frac{d}{dt}(\rho_1 \ln T_1 + \rho_2 \ln T_2) &\geq \frac{Q_1}{T_1} + \frac{Q_2}{T_2}. \end{aligned}$$

These are the first and second law as they pertain to the original interconnected system. The basic principle is that the interconnected system obeys the laws of thermodynamics because the subsystems do. The internal energy and the entropy are the sum of the internal energies and entropies of the subsystems.

This is in fact a fully general principle. Conditions for an interconnection of systems (see the figure below, but view this, hierarchically, as a complex interconnection of numerous subsystems), to satisfy the laws of thermodynamics if the interconnected systems do, are readily obtained.



The basic constraints are similar to those that were put in evidence in the simple example given above. For each interconnected thermal terminal the interconnection constraints should imply equal that the temperatures are equal and that the heat flows sum to zero. The work interconnections should imply that the rate of work performed by the interconnected system is equal to the sum of the rates of 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 into one equals work out of 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*, quantities that add: the entropy of the whole is the sum of the entropies of the parts, just as volume and mass and charge, in contrast to *intensive quantities*, quantities that do not add, as temperature and voltage and position. This extensivity of entropy can have 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 simpler subsystems, each of which has a unique entropy function.

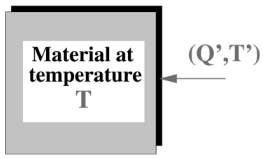
D. Inputs, Outputs, and States

Throughout the 20th century, mainstream systems 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 cybernetic stimulus/response view of Wiener, generations of systems 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, 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 regard Kalman's input–state–output framework to be the first model structure that is adequate for the dynamical description 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 universal 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 processor. A law of physics states that certain outcomes are compatible, that certain combinations of values of physical variables can occur simultaneously, but not that one variable causes another.

Consider, as an example, the simple system discussed in Subsection 3-C, with only one heat terminal, shown again in the figure below.



The dynamical equations consist of the differential equation

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

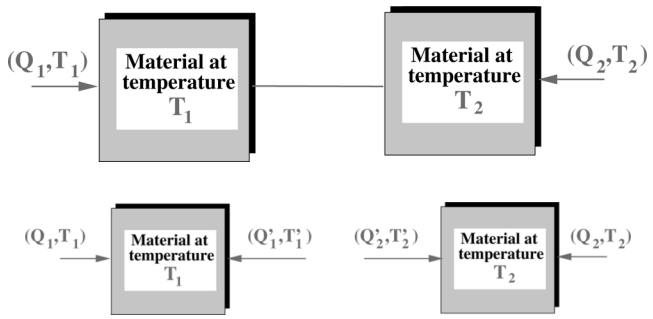
combined with

$$[\![Q' \geq 0 \quad \text{and} \quad T' \geq T]\!] \quad \text{or} \quad [\![Q' \leq 0 \quad \text{and} \quad T' \leq T]\!]$$

This defines, as we have seen, a thermodynamic system. The external system variables are Q', T'. They are both, in a sense, free inputs, but only to some extent. $Q' \geq 0$ cannot let T become larger than T', $T \geq T'$ implies $Q' \leq 0$, etc. The question of what causes what should

not be posed. This exceedingly simple example shows that the laws of thermodynamics are at odds with input/output thinking. In physical systems, there are certain related variables which the model aims at, but there is no point is insisting on a partition of these variables in inputs and outputs, causes and effects.

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 once again the system discussed in Subsection 3-C, viewed as the interconnection of two systems.



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

$$T_1' = T_2' \quad Q_1' + Q_2' = 0.$$

So if, for some reason, we have decided 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 physical systems means variable sharing, not signal transmission.

In Subsection 3-B, we have seen that, in order to discuss dissipative systems, it is very reasonable to consider the behavior that consists of all possible supply rate trajectories $s : \mathbb{R} \to \mathbb{R}$. The question of how supply flows in and out leads to our "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 $s_2 = \sum_{i=1}^n (Q_i/T_i)$, and the W's and (Q_i, T_i)'s constrained by $\mathcal{B}_{\text{thermodynamic}}$. Asking if s_1 or s_2 is an input or an output is again absurd.

We have also seen in Subsection 3-B that it is not necessary to introduce a state in order to discuss the storage. However, it is a good question to ask whether, if there exists a storage at all, there always exists a storage that is a state function. For physical systems, state ("memory") is a much more fundamental concept than input ("cause") or output ("effect").

IV. Conclusion

Thermodynamics is, by its very essence, a theory of *open systems*.[2] It puts limitations to the way in which physical systems are able to exchange energy and heat with their environment. "Flows" are totally incapable of dealing with thermodynamics. 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 discuss 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 aspect 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.

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 approach and the results, but also by the systems point of view as the basis for thermodynamics.

In my opinion, a 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. I believe that the basic set-up in chapter 3 could haven been clarified by considering classical examples, as 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 their 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 nonuniform) 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 thermodynamics in the context of dissipative systems and interconnections, along the lines of what I pointed to in Sections 3-B and 3-C.

This book is a scholarly one. It 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 trumpeted newest research themes, and deals with a classical poorly understood, but exceedingly important subject, is very welcome. The authors of this monograph should be commended in their aim to explain an important domain as thermodynamics from a systems theory point of view to the community.

[2]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 $(d/dt)x = f(x)$, 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.