

TOOLS OF FINITE-TIME THERMODYNAMICS

Bjarne Andresen

Niels Bohr Institute, University of Copenhagen
Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

Abstract Finite-time thermodynamics is the extension of traditional reversible thermodynamics to include the extra requirement that the process in question goes to completion in a specified finite length of time. As such it is by definition a branch of irreversible thermodynamics, but unlike most other versions of irreversible thermodynamics, finite-time thermodynamics does not require or assume any knowledge about the microscopics of the processes, since the irreversibilities are described by macroscopic constants such as friction coefficients, heat conductances, reaction rates and the like. Some concepts of reversible thermodynamics, such as potentials and availability, generalize nicely to finite time, others are completely new, e.g. endoreversibility and thermodynamic length. The basic ideas of finite-time thermodynamics are reviewed and several of its procedures presented, emphasizing the importance of power and rate of entropy production. Finally, its impact on the global optimization algorithm simulated annealing is outlined.

1. Introduction

1.1 Motivation

From its infancy over 150 years ago, thermodynamics has provided limits on work or heat exchanged during real processes. The first problem treated in a systematic way was how much work a steam engine can produce from the burning of one ton of coal. With true scientific generalization Sadi Carnot concluded that any engine taking in heat from a hot reservoir at temperature T_H has to deposit some of that heat in a cold reservoir (e.g. the surroundings), whose temperature we call T_L ; the largest fraction of the heat which can be converted into work is

$$\eta_C = 1 - \frac{T_L}{T_H}, \quad (1.1)$$

traditionally known as the Carnot efficiency. This expression contains the two basic ingredients of a thermodynamic limit: a) it applies to *any* process converting heat into work; and b) it is an *absolute* limit, i.e. no process, however ingenious, can do better.

As thermodynamic theory developed, emphasis changed from process variables like work and heat exchanged to state variables like entropy and chemical potential. A bridge between the two are the thermodynamic work potentials, such as enthalpy H for isobaric processes or the Gibbs free energy G for isothermal, isobaric processes. They are defined such that their changes provide upper bounds on the work a process can supply or lower bounds on the work required to drive a process. Gibbs introduced the concept of ‘available work’ as the maximum work that can be extracted from a system allowed to go from a constrained, internally equilibrated state to a state in equilibrium with its surroundings. This quantity is used more and more frequently in engineering contexts (Keenan 1941;

Gaggioli 1980) under the names ‘availability’ in the U.S. and ‘exergy’ in Europe. For a system relaxing to an ambient temperature T_0 , pressure P_0 , and chemical potentials μ_{0i} it is given by

$$A = U + P_0V - T_0S - \sum_i \mu_{0i}N_i \quad (1.2)$$

and is thus not a state function in the usual sense of depending only on variables of the system; the availability depends on the intensive variables of the environment as well.

Such criteria of merit have long been common currency for thermodynamic studies in physics, chemistry, and engineering. They all share one characteristic: The ideal to which any real process is compared is a *reversible process*. Stated in a different way, *traditional thermodynamics is a theory about equilibrium states and about limits on process variables for transformations from one equilibrium state to another*. Nowhere does time enter the formulation, so these limits must be the lossless, reversible processes which proceed infinitely slowly and thus take infinite length of time to complete. However, referring back to the original question addressed by Carnot, who is interested in an engine which operates infinitely slowly (and thus produces zero power) — or any other process with zero rate of operation, for that matter?

In order to obtain more realistic limits to the performance of real processes *finite-time thermodynamics is designed as the extension of traditional thermodynamics to deal with processes which have explicit time or rate dependences*. These constraints, of course, imply a certain amount of loss, or entropy production, which is at the heart of the question posed above.

1.2 Early developments

In the course of developing finite-time thermodynamics we discovered that a few isolated papers already had considered different aspects of processes operating at nonzero rates. The first of these was the important work of Tolman and Fine (1948) who put the Second Law of thermodynamics into equality form,

$$W = \Delta A - T_0 \int_{t_i}^{t_f} \dot{S}_{\text{tot}} dt \quad (1.3)$$

by subtracting the work equivalent of the entropy produced during the process from the reversible work, i.e. the decrease of system availability, as defined in eq. (1.2). The superscript dot indicates rate, and the integral limits are the initial and final times of the process. This is a quantification of the ‘price of haste’.

Another model has evolved into a classic paradigm of systems operating in finite time. This is the model of Curzon and Ahlborn (1975), a Carnot engine with the simple constraint that it be linked to its surroundings through *finite* heat conductances. Figure 1 illustrates the slightly more general endoreversible system with the triangle signifying any reversible engine. (The term endoreversible means ‘internally reversible’, i.e. all irreversibilities reside in the coupling of flows to the surroundings. In this case that means resistance to heat transfer and possibly friction.) It turns out that the results derived by Curzon and Ahlborn explicitly for an interior Carnot engine are equally valid for a general endoreversible system. The maximum efficiency of their engine is of course $\eta_C = 1 - T_L/T_H$, obtained at zero rate so that losses across the resistors vanish, but these authors showed that, when the system operates to produce *maximum power*, the efficiency of the engine is only

$$\eta_w = 1 - \sqrt{\frac{T_L}{T_H}}. \quad (1.4)$$

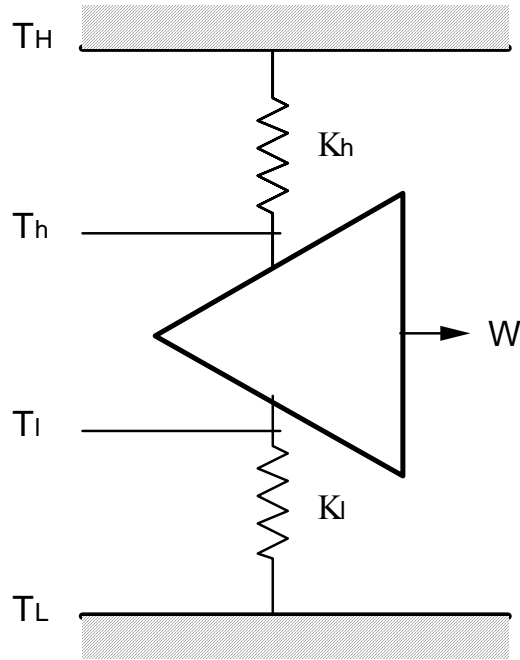


Figure 1. An endoreversible engine has all its losses associated with its coupling to the environment, there are no internal irreversibilities. This is illustrated here as resistances in the flows of heat to and from the working device indicated by a triangle. These unavoidable resistances cause the engine proper to work across a smaller temperature interval, $[T_h; T_l]$ than that between the reservoirs, $[T_H; T_L]$, one which depends on the rate of operation.

Besides the simplicity of the expression it is remarkable that it does not contain the value of the heat conductances. A closer analysis of this expression and its relationship to the effects of finite size heat reservoirs is presented by Gordon (1989). The initial development of finite-time thermodynamics was primarily inspired by Curzon and Ahlborn's paper.

In the next section we will see how it is possible to derive bounds on finite-time processes without knowing their detailed time paths. Following that we derive the optimal paths for a few examples.

2. Performance bound without path

The smallest amount of information one can ask for concerning the performance of a system is a single number, e.g. the work or heat exchanged during the process, its efficiency, or any other figure of merit. In most cases this can be calculated without knowledge of the detailed path followed and is then computationally much simpler to obtain.

2.1 Generalized potentials

In traditional thermodynamics potentials are used to describe the ability of a system to perform some kind of work under given constraints. These constraints are usually the constancy of some state variables like pressure, volume, temperature, entropy, chemical potential, particle number, etc. Under such conditions the decrease in thermodynamic potential P from state i to state f is equal to the amount of work that is produced when a reversible process carries out the transition, and hence is the upper bound to the amount of work produced by any other process,

$$W \leq W_{\text{rev}} = P_i - P_f. \quad (2.1)$$

In this section we will show that the constraints need not simply be the constancy of some state variable, and that the potentials may be generalized to contain constraints involving time (Salamon et al. 1977). The procedure will be a straight forward extension of the Legendre transformations (Hermann 1973) used in traditional thermodynamics (Callen 1985; Tisza 1966), and we will start with such an example.

In a reversible process heat and work can be expressed as inexact differentials,

$$dQ = TdS, \quad dW = PdV, \quad (2.2)$$

i.e. they cannot by themselves be integrated, further constraints defining the integration path are required. Such a constraint could be that the process is isobaric, $dP = 0$. One can then add a suitable integrating zero-term, $x dP$ to make dW an exact differential. The obvious choice is $x = V$,

$$dW = PdV = PdV + VdP = d(PV), \quad (2.3)$$

such that the isobaric work potential becomes $P_W = PV$. Similarly the isobaric heat potential $P_Q = U + PV$ is obtained from

$$dQ = TdS = dU + PdV = dU + PdV + VdP = d(U + PV), \quad (2.4)$$

where the First Law of thermodynamics

$$dU = TdS - PdV \quad (2.5)$$

has been used. Table 1 shows the results of this procedure for the classical examples.

Table 1. The classical thermodynamic potentials for the process variables of work $dW = PdV$ and heat $dQ = TdS$.

Process type	Zero along process	Integrating term	Work potential	Heat potential
Isobaric	dP	VdP	PV	$U + PV$
Isothermal	dT	SdT	$TS - U$	TS
Isochoric	dV	$-PdV$	0	U
Isentropic	dS	$-TdS$	$-U$	0

Now, the constraints need not be the constancy of one of the state variables. Consider a balloon with constant surface tension α . In equilibrium with an external pressure P_{ex} such a sphere of radius r has an internal pressure

$$P = P_{ex} + \frac{2\alpha}{r}, \quad (2.6)$$

which can be rearranged into

$$(P - P_{\text{ex}})V^{1/3} = 2\alpha \left(\frac{4\pi}{3}\right)^{1/3}. \quad (2.7)$$

Since the right hand side of this equation is a constant, this means that $(P - P_{\text{ex}})V^{1/3}$ is an integral of motion for the fluid inside the balloon. We can therefore add a suitable amount of $d[(P - P_{\text{ex}})V^{1/3}]$ ($=0$) to dW to make it exact,

$$\begin{aligned} dW &= PdV = PdV + \frac{3}{2} V^{2/3} d[(P - P_{\text{ex}})V^{1/3}] \\ &= d\left[\frac{1}{2} V(3V - P_{\text{ex}})\right]. \end{aligned} \quad (2.8)$$

Thus the work done by the coupled system balloon + fluid is given by the decrease in the potential $P_W = \frac{1}{2} V(3V - P_{\text{ex}})$, regardless of path followed.

In its most general form the Legendre transformation can be used to calculate a potential P_B for the arbitrary process variable B , expressible as a path integral in terms of generalized forces f_i and displacements x_i ,

$$B = \sum_i \int f_i dx_i = \int \mathbf{f} \cdot d\mathbf{x}. \quad (2.9)$$

B will usually be work, and vector notation is used for compactness. To find P_B , one adds to $\mathbf{f} \cdot d\mathbf{x}$ an integrating term $\mathbf{g} \cdot d\mathbf{y}$, where $d\mathbf{y}$ is necessarily zero as a result of the constraints defining the process. Note that $d\mathbf{y} = \mathbf{0}$ may involve time and could come from a condition in the form of a differential equation as well as from the more familiar thermodynamic condition of a constant variable, as used in the example above. Hence the differential form $d\mathbf{y} = \mathbf{0}$ is used rather than the integrated form $\mathbf{y} = \text{constant}$, since \mathbf{y} itself may not exist. The mathematical problem of finding P_B has two steps, finding a function \mathbf{g} which makes $d\omega = \mathbf{f} \cdot d\mathbf{x} + \mathbf{g} \cdot d\mathbf{y}$ an exact differential dP_B , and then integrating to get P_B itself. The first step involves the Cauchy-Riemann condition that $d\omega$ has equal cross derivatives with respect to the free state variables, e.g. a and b :

$$\frac{\partial}{\partial b} \left[\mathbf{f} \cdot \left(\frac{\partial \mathbf{x}}{\partial a}\right)_b + \mathbf{g} \cdot \left(\frac{\partial \mathbf{y}}{\partial a}\right)_b \right] = \frac{\partial}{\partial a} \left[\mathbf{f} \cdot \left(\frac{\partial \mathbf{x}}{\partial b}\right)_a + \mathbf{g} \cdot \left(\frac{\partial \mathbf{y}}{\partial b}\right)_a \right] \quad (2.10)$$

or

$$\left(\frac{\partial \mathbf{g}}{\partial a}\right)_b \cdot \left(\frac{\partial \mathbf{y}}{\partial b}\right)_a - \left(\frac{\partial \mathbf{g}}{\partial b}\right)_a \cdot \left(\frac{\partial \mathbf{y}}{\partial a}\right)_b = \left(\frac{\partial \mathbf{f}}{\partial b}\right)_a \cdot \left(\frac{\partial \mathbf{x}}{\partial a}\right)_b - \left(\frac{\partial \mathbf{f}}{\partial a}\right)_b \cdot \left(\frac{\partial \mathbf{x}}{\partial b}\right)_a. \quad (2.11)$$

With \mathbf{f} , $d\mathbf{x}$, and $d\mathbf{y}$ known, this is the equation from which \mathbf{g} may be obtained. In the usual case of $\mathbf{f} = P$, $d\mathbf{x} = dV$, $a = V$, and $b = P$, the right hand side of eq. (2.11) simplifies to 1. The second step in finding P_B , the integration of dP_B , is, of course, only unique within a constant of the motion; i.e. two methods of integration may yield two different potentials P_B and P_B' , but their variations will always be the same, $\Delta P_B = \Delta P_B'$.

Whereas for reversible processes there is no question that thermodynamic potentials exist, because the processes can always be reversed or go by way of an arbitrary third state, this is not obvious for generalized potentials with built-in time dependence and possible loss terms. In generalizing the Legendre transformation above, we have implicitly assumed the existence of a potential or, equivalently, a solution to eq. (2.11). The conditions for existence are (Salamon et al. 1977):

- a) the process is quasistatic, and
- b) the process variable can be expressed as a path integral [eq. (2.9)].

The quasistaticity is equivalent to saying that the relaxation times of the system are negligible compared to the time scale of the process. In other words, the system is not required to be in equilibrium with its surroundings at all times, but all state variables must be defined (make sense) at a countable pointset along the path. The second condition is similar in content, because this seemingly trivial condition is violated when one tries to describe the extraction of work from a system at a rate faster than that system can equilibrate internally (think of a combustion process). In such a case thermodynamic variables lose their meaning, and one must go to a definition of work in terms of energy transfer at the microscopic level, which sometimes can be too complex to be useful. Nevertheless, work and availability can be defined for some systems, such as simple lasers whose operation depends on changes in populations of specific quantum states (Geusic et al. 1967). Another approach relies on the information-theoretic ‘maximum entropy’ formalism (see e.g. Levine and Tribus 1979).

2.2 Finite-time availability

One of the more powerful results in finite-time thermodynamics is the definition of a finite-time availability (Andresen et al. 1983). As mentioned in Sect. 1.1, the traditional availability A of a system in contact with given surroundings is a state function with the quality that the decrease in its value in going from state i to state f is the maximum (and hence reversible) work that can be extracted during that process. The finite-time availability A retains this property and simply adds that the process is restricted to operate (go to completion) during time $\tau = t_f - t_i$. Then

$$A = W_{\max}(\tau) = \max \left[A(t_i) - A(t_f) - T_0 \int_{t_i}^{t_f} \dot{S}_{\text{tot}} dt \right], \quad (2.12)$$

where the last equality uses the Tolman-Fine (1948) form of the Second Law of thermodynamics, eq. (1.3).

The extension may seem trivial, but the principal content lies in the way the maximization is carried out, or rather restricted. It must be carried out within the constraints imposed on the process, temporal or otherwise. These constraints in effect define a generic model which constitutes the confines within which we expect to be able to modify our real system in order to improve its performance. The constraints in such a generic model should not be excessively detailed, but only contain the essential loss terms and limiting factors in the process, otherwise the calculations will become unwieldy. Of course, if one is going to use finite-time availability to compare the performance of two different processes, they must be represented by the same generic model. Otherwise a path allowed in one may not be available to the other process, and such a restriction always costs performance. Losses are not always detrimental to the performance of a system if they open up new pathways — actually some processes depend on irreversibilities for their very existence (e.g. Wheatley et al. 1983).

In addition to the above considerations, the maximum search in eq.(2.12) can either be constrained to exactly reach a given final state at time t_f (the initial state is always considered known), in which case ΔA is fixed, and the optimization becomes one of minimizing the entropy production, or also the final state may be included in the optimization, in which case A must be evaluated by optimal control. If the final state is specified, a solution may not exist if τ is too short, since only a certain set of states can be reached from a given initial state in time τ . In addition, the finite-time availability does not necessarily have ΔA as its limit for very long times, because the system may contain internal relaxation processes which remain irreversible even for very slow operation. If there is a direct heat

leak from the system reservoir to the surroundings, then a long process time may even reduce A to zero.

As defined, the finite-time availability is as general as the traditional availability, i.e. it can be applied to *any* thermodynamic process. Andresen et al. (1983) report an optimal control calculation of the finite-time availability for a work-producing system with competing internal relaxation which can be interpreted as heat engines, internal molecular degrees of freedom, a hydraulic system, or a chemical reaction, simply by changing nomenclature.

2.3 Non-endoreversible processes

In all fairness it should be said that not all real processes are well described by an endoreversible model. Some operations, like large air conditioning systems, have internal losses that far outweigh losses in the couplings to the surroundings (Gordon and Ng 1994, 2000). In those cases obviously the Curzon-Ahlborn procedure and its extensions yield unrealistic results. However, it is possible to turn the models around and still get nice simple efficiency expressions with predictive power for where operational problems are located (Gordon and Ng 2000).

2.4 Thermodynamic length

In an effort to develop a more direct and transparent way of calculating all the usual partial derivatives in traditional thermodynamics Weinhold (1975abcd, 1978) proposed using vector products between vectors in the abstract space of equilibrium states of a system, represented by all its extensive variables X_i . The products were defined relative to

$$\mathbf{M}_U = \left\{ \frac{\partial^2 U}{\partial X_i \partial X_j} \right\} \quad (2.13)$$

as the metric, where U is the internal energy. However, second derivatives are usually identified as curvatures and, as such, should be interpreted as availabilities in Gibbs space (U as a function of all the other extensive variables). This lead us to seek another interpretation of pathlengths calculated with Weinhold's metric, now called thermodynamic lengths, and we (Salamon et al. 1980a) found that they always are changes in some molecular velocities, depending, of course, on the constraints of the process (isobaric, isochoric, etc.). The geometry is illustrated in Fig. 2.

Subsequently Salamon and Berry (1983) found a connection between the thermodynamic length along a process path and the (reversible) availability lost in the process. Specifically, if the system moves via states of local thermodynamic equilibrium from an initial equilibrium state i to a final equilibrium state f in time τ , then the dissipated availability $-\Delta A$ to leading order in the rate is bounded by the square of the distance (i.e. length of the shortest path) from i to f times ε/τ , where ε is a mean relaxation time of the system. If the process is endoreversible, the bound can be strengthened to

$$-\Delta A \geq \frac{L^2 \varepsilon}{\tau}, \quad (2.14)$$

where L is the length of the *traversed* path from i to f . Equality is achieved at constant thermodynamic speed $v = dL/dt$ corresponding to a temperature evolution given by (Nulton and Salamon 1988)

$$\frac{dT}{dt} = -\frac{vT}{\varepsilon\sqrt{C}}, \quad (2.15)$$

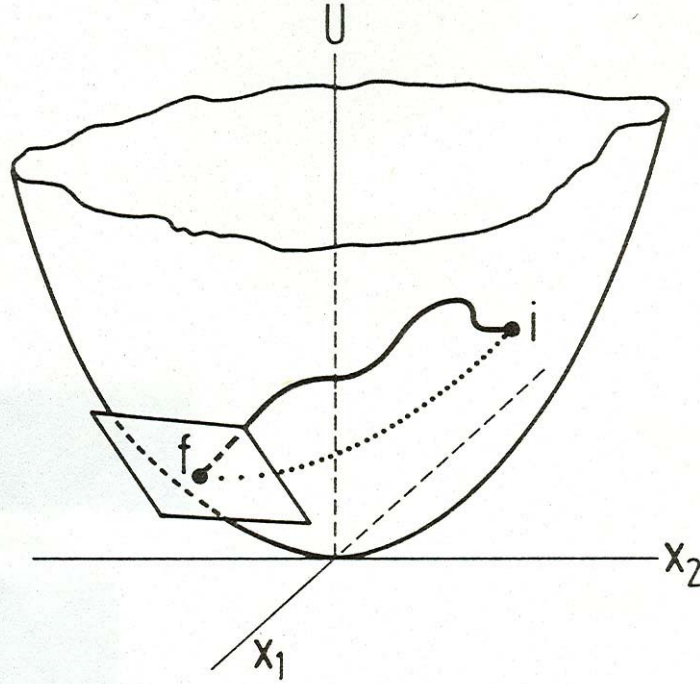


Figure 2. The state space of a thermodynamic system represented by the internal energy U and the other extensive variables X_1 and X_2 . The paraboloid illustrates the equilibrium states of the system, while a tangent plane represents an environment with its constant intensive variables. A system initially in state i can move toward equilibrium with the environment at the tangent point f along an infinite number of internal equilibrium paths. The length of the shortest of these (dotted) is determined by the Weinhold metric, but the actually traversed path (solid) may be longer. The availability dissipated in the process is bounded by the square of this distance times an internal relaxation time and divided by the duration of the process.

where C is the heat capacity of the system. For comparison, the bound from traditional thermodynamics is only

$$-\Delta A \geq 0. \quad (2.16)$$

An analogous expression exists for the total entropy production during the process:

$$\Delta S^u \geq \frac{L^2 \varepsilon}{\tau}. \quad (2.17)$$

The length L is then calculated relative to the entropy metric

$$\mathbf{M}_S = - \left\{ \frac{\partial^2 S}{\partial X_i \partial X_j} \right\} \quad (2.18)$$

which (when expressed in identical coordinates!) is related to \mathbf{M}_U by (Salamon et al. 1984)

$$\mathbf{M}_U = -T_0 \mathbf{M}_S, \quad (2.19)$$

where T_0 as usual is the environment temperature. In statistical mechanics, where entropy takes the form

$$S(\{p_i\}) = - \sum_i p_i \ln p_i, \quad (2.20)$$

the metric \mathbf{M}_S is particularly simple, being the diagonal matrix (Feldmann et al. 1985)

$$\mathbf{M}_S = - \left\{ \frac{1}{p_i} \right\}. \quad (2.21)$$

The same procedure of calculating metric bounds for dynamic systems has been applied to coding of messages (Flick et al. 1987) and to economics (Salamon et al. 1987).

More recently (Andresen and Gordon 1994) we have relaxed a number of the assumptions in the original work, primarily those restricting the system to be close to equilibrium at all times and the average form of the relaxation time ε . The more general bound replacing eq. (2.17) then becomes

$$\Delta S^u \geq \frac{1}{\Xi} \left(\int_{\xi_i}^{\xi_f} \frac{1}{T \sqrt{C}} \left| \frac{dU}{d\xi} \right| \sqrt{1 + \frac{\theta}{CT} \frac{dU}{d\xi} + \dots} d\xi \right)^2 \quad (2.22)$$

with $\Xi = \xi_f - \xi_i$ being the total duration of the process in natural dimensionless time units,

$$d\xi = dt/\varepsilon(T), \quad (2.23)$$

and where we have defined

$$\theta(T) = 1 + \frac{T}{2C} \frac{\partial C}{\partial T}. \quad (2.24)$$

The equality (lower bound) in eq. (2.22) is achieved when the integrand is a constant, i.e. when

$$\frac{dS^u}{d\xi} = \text{constant}. \quad (2.25)$$

Consequently, constant rate of entropy production, when expressed in terms of natural time, is the path or operating strategy which produces the least overall entropy.

One can express the optimal path in a form similar to eq. (2.15):

$$\frac{dT}{dt} \sqrt{1 + \frac{\theta(T) \varepsilon (dT/dt)}{T}} + \dots = \text{constant} \times \frac{T}{\varepsilon \sqrt{C}}. \quad (2.26)$$

The constant thermodynamic speed algorithm, eq. (2.15), is thus the leading term of the general solution in an expansion about equilibrium behavior.

3. Optimal path

3.1 Optimal path calculations

A knowledge of the maximum work that can be extracted during a given process, e.g. calculated by one of the procedures described in the previous section, may not by itself be sufficient. One may also want to know *how* this maximum work can be achieved, i.e. the time path of the thermodynamic variables of the system. The primary tool for obtaining this path is optimal control theory.

We have a system described by n state variables $\mathbf{x} = (x_1, x_2, \dots, x_n)$, e.g. volume, pressure, temperature etc. For this system we want to maximize a certain function, the objective function, which we for full generality write as an integral over the full cycle of the process,

$$\int_0^\tau A(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, \dots) dt . \quad (3.1)$$

Note that it may depend not only on the state variables but also on any of their derivatives. In this expression we have written the path parameter as time, but really it can be any parameter which is convenient for the problem. To achieve the extremum we have at our disposal k controls $\mathbf{c} = (c_1, c_2, \dots, c_k)$ which can be varied instantly and at no cost. Usually each control is limited to a certain range, $[c_{\min}, c_{\max}]$ or the physical situation may require it to be positive (e.g. only positive reservoir temperatures). During the process certain constraints must be obeyed like energy or mass conservation and the evolution equations of the state variables,

$$\mathbf{B}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, \dots) = 0 . \quad (3.2)$$

Averaged constraints, i.e. true only on average over the cycle, are also possible and are represented by integrals like eq. (3.1).

Solving such an extremum problem with constraints is usually very difficult. It becomes a little easier if we can remold the problem to be without constraints. This can be achieved by defining a Hamiltonian

$$H = A(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, \dots) + \mathbf{p} \cdot \mathbf{B}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}, \dots) \quad (3.3)$$

which incorporates the constraints through a new set of functions $\mathbf{p} = (p_1, p_2, \dots, p_n)$ conjugate to the state variables. Possible solutions, just like in classical mechanics, then satisfy Hamilton's equations,

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \quad (3.4a)$$

$$\frac{dp_i}{dt} = - \frac{\partial H}{\partial x_i} , \quad (3.4b)$$

one pair of equations for each pair of independent state and conjugate variables. Note that eqs. (3.4a) are just a restatement of the constraints eq. (3.2).

Eqs. (3.4) naturally contain the (so far unknown) controls \mathbf{c} . Those may often be found by application of the very powerful Pontryagin maximum principle which states that if \mathbf{x}^* , \mathbf{c}^* , and \mathbf{p}^* are the fully optimal state functions, controls, and conjugate functions, then

$$H(\mathbf{x}^*, \mathbf{c}^*, \mathbf{p}^*) \geq H(\mathbf{x}^*, \mathbf{c}, \mathbf{p}^*) , \quad (3.5)$$

i.e. if the state and conjugate functions are still optimal and only the controls are suboptimal, the Hamiltonian can be no larger than the Hamiltonian of the fully optimal trajectory. With this principle in hand one can often argue that the optimal controls c^* must assume one of their extreme values whenever certain combinations of the variables are either positive or negative. For a more detailed description of this procedure see e.g. Rubin (1979).

With these values of the controls solutions of eqs. (3.4), which are called branches or interior solutions, are still only possible extremal solution, because boundary solutions may be even better. This is equivalent to finding extremal values of a simple function of a single variable $f(x)$ where we need to set the derivative $df/dx = 0$ as well as evaluate boundary values.

The next step in the optimization is to piece these possible branches together to a full solution. The requirements are that the Hamiltonian is constant and that all state and conjugate functions are continuous also at switchings from one branch to another, although the controls are free to have discontinuities. This piecing together of branches is usually not unique. It may have one, several or no solutions, and the only way to find out which is the optimum is to try them all.

To summarize let me point out that in order to set up the optimal control problem one must specify:

- the controls, i.e. the variables that can be manipulated by the operator (they may be a volume, rate, voltage, heat conductance, etc.);
- limits on the controls and on the state variables, if any (in order to avoid unphysical situations such as negative temperatures and infinite speeds);
- the equations that govern the time evolution of the system (they will usually be differential equations describing heat transfer rates, chemical reaction rates, friction, and other loss mechanisms);
- the constraints that are imposed on the system (e.g. conserved quantities, the quantities held constant, or any requirements on reversibility. The constraints may either be differential, instantaneous i.e. algebraic, or integral i.e. not obeyed at each point but over the entire interval);
- the desired quantity to be maximized, called the objective function (usually expressed as an integral); and finally
- whether the duration of the process is fixed or part of the optimization.

Typical manipulation usually leads to a set of coupled, non-linear differential equations for which a qualitative analysis and a numerical solution are the only hope. Thus answering the more demanding question about the optimal time path rather than the standard question about maximum performance requires a considerably larger computational effort. On the other hand, once the time path is calculated, all other thermodynamic quantities may be calculated from it, much like the wave function is the basis of all information in quantum mechanics. More extensive explanations of optimal control calculations are widely available (see e.g. Boltyanskii 1971; Tolle 1975; Leondes 1964).

3.2 Criteria of performance

Efficiency, the earliest criterion of performance for engines, measured how much water could be pumped out of a mine by burning a ton of coal. Other familiar criteria include effectiveness, change of thermodynamic potential, and loss of availability, all of which are measures of work. Potentials for heat can also be defined (see Sect. 2.1) but are less common, and we have used the minimization of entropy production in a separate study (Salamon et al. 1980b).

The Curzon-Ahlborn analysis and many of our own analyses use a quite different criterion, that of *power*. This quantity is of course zero for any reversible system, and maximizing power forces us to deal with systems operating at finite rates. Other criteria of performance are the rate of entropy production and the rate of loss of availability. Entropy production was a function introduced in the

earliest thinking about irreversible thermodynamics (Onsager 1931; Prigogine 1962, 1967), but more from the differential, local, instantaneous viewpoint than from the global, integral view of entire optimized processes. Under some circumstances, optimizing one of these quantities is equivalent to optimizing another. For example, minimizing the entropy production is equivalent to minimizing the loss of availability, at least in those cases in which the irreversibilities can be represented as spontaneous heat flows (Salamon et al. 1980b).

Salamon and Nitzan (1981) have optimized the Curzon-Ahlborn engine for a number of these objective functions. Assuming the working fluid to be in contact with the hot reservoir for the period τ_1 and the cold reservoir for the period τ_2 , the optimal time distributions are shown in Fig. 3. The diagonal $\tau = \tau_1 + \tau_2$ indicates fixed total cycle time, and only processes above the curve labeled 'zero power' produce positive power. It is quite obvious that different criteria of merit dictate different operating conditions for the process. Even when not knowing the precise objective function but only that it belongs to a specified class, one can sometimes say a good deal about the possible optimal behavior of the system. If one considers the Curzon-Ahlborn engine to be a model of a power plant which buys heat (coal) at the unit price α and sells work (electricity) at the unit price β , its net revenue is $\Pi = \beta w - \alpha q_1$. All solutions to the problem of maximizing this revenue are bounded on one side by the solutions to the maximum power problem (when α is insignificant compared to β) and on the other side by the solutions corresponding to minimum loss of availability (when coal and electricity are priced according to their availability contents). While this is a very simple example, this approach has far reaching possibilities for describing biological, ecological, and economic systems.

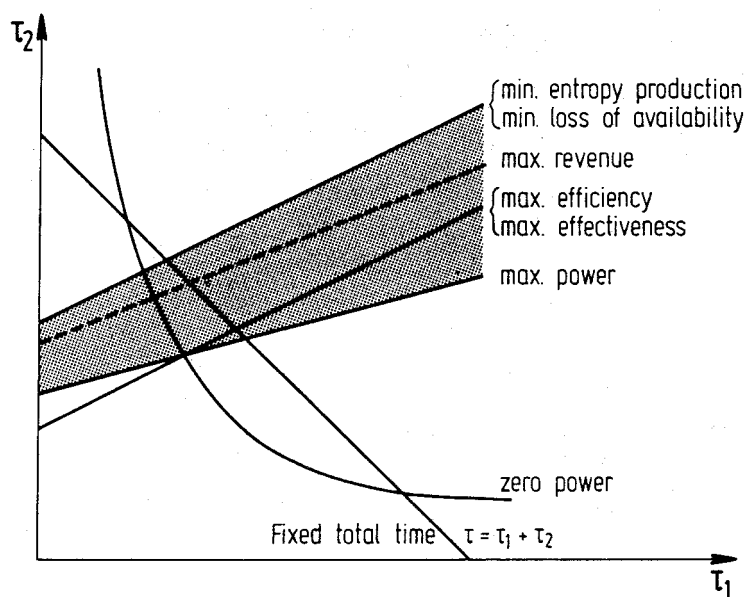


Figure 3. If an endoreversible engine (Fig. 1) spends time τ_1 in contact with the hot reservoir and τ_2 in contact with the cold reservoir, the optimal proportioning between τ_1 and τ_2 depends on what one chooses to optimize, as indicated. The locus of maximum revenue for a power producing system falls in the shaded area for any choice of prices, as described in the text. Only contact times above the hyperbola marked 'zero power' actually correspond to positive power production.

3.3 Paths for endoreversible engines

Maximum power. Rubín (1979) was the first to apply optimal control theory to a finite-time system. He calculated the optimal path for the endoreversible engine of Fig. 1 which maximizes its power output. Without any restrictions as to type of branches, but only limiting the reservoir temperatures, heat conductances, and rate of change of volume, he found the optimal cycle to consist

of six branches: two isotherms and four maximum-power branches, but no adiabats. The optimal controls (reservoir temperature, heat conductance, and relative rate of volume change) are shown in Fig. 4 and the optimal state variables (temperature and pressure) in Fig. 5. Note that only the extreme available temperatures, T_H and T_L , are required, as is the maximum conductance to the reservoirs. The adiabats of a Carnot cycle are replaced by the maximum-power branches along which the volume changes at its maximum permissible rate while the working fluid is still in contact with one or the other heat reservoir. Thus the optimal cycle is never isolated from both its reservoirs. If the limitation on the rate of volume change is lifted, these new branches proceed instantaneously and in effect become adiabats. This optimal cycle is the finite-time equivalent of the Carnot cycle in reversible thermodynamics and serves the same purpose, viz. as an idealized reference processes for evaluation of real cycles. Curzon and Ahlborn's (1975) assumption of an interior Carnot cycle in their system (see Sect. 1.2) is now justified by this optimization.

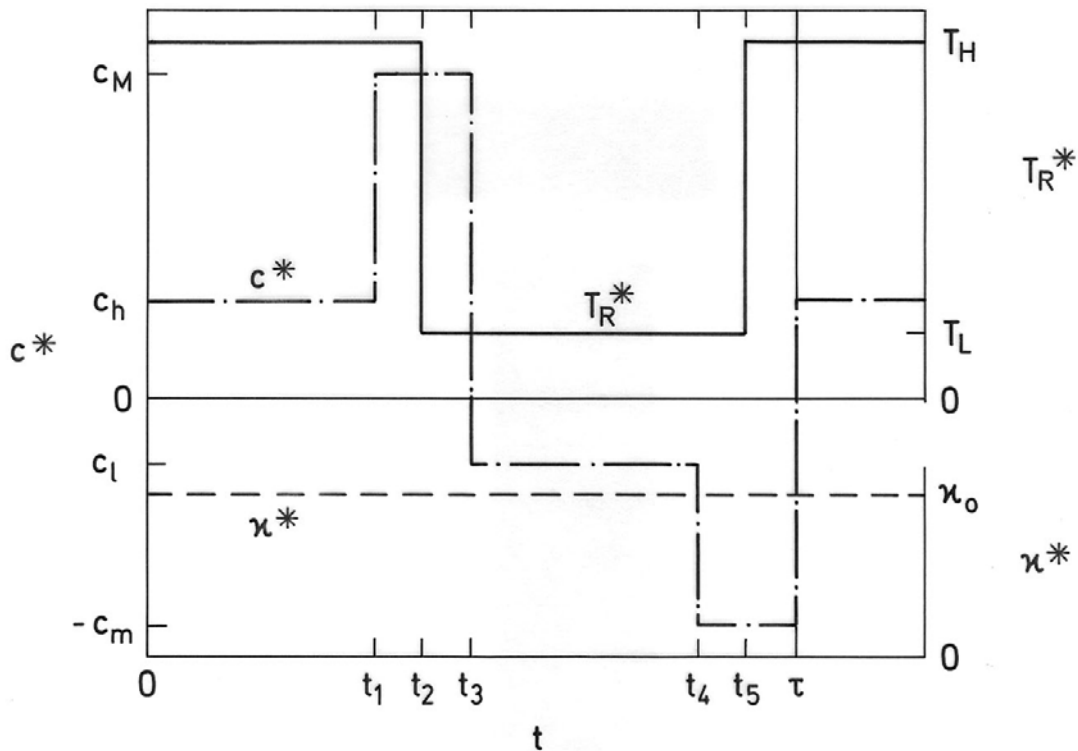


Figure 4. The optimal controls over a period τ for an unrestricted optimization of the endoreversible engine (Fig. 1). $c^*(t) = \dot{V}/V$ is the relative rate of change of volume, and $\kappa^*(t)$ is the heat conductance to the reservoir with temperature $T_R^*(t)$. Note that, except for the rate of change of volume on the isothermal branches $[0; t_1]$ and $[t_3; t_4]$, only the extreme values of the controls are required.

Minimum entropy production. About the same time we analyzed the performance of endoreversible engines in terms of their entropy production (Salamon et al. 1980b). Minimizing that is equivalent to minimizing the loss of availability, but not to maximizing efficiency or power, as most other studies have done. The system exchanges generalized fluxes with the surroundings at a rate which depends on the generalized forces in a completely arbitrary way, but not on time or any time derivative of the forces. Then, regardless of whether the internal or the external temperature can be controlled by the operator, minimum total entropy production always implies constant rate of entropy production on each branch of the cycle. If the fluxes are linear in the forces, this constant entropy production rate must be the same on all branches of the cycle. If the class of systems is further

restricted to admit only Fourier heat flow through a conductance κ to the reservoir, one finds that the total entropy produced in a cycle with period τ is bounded by

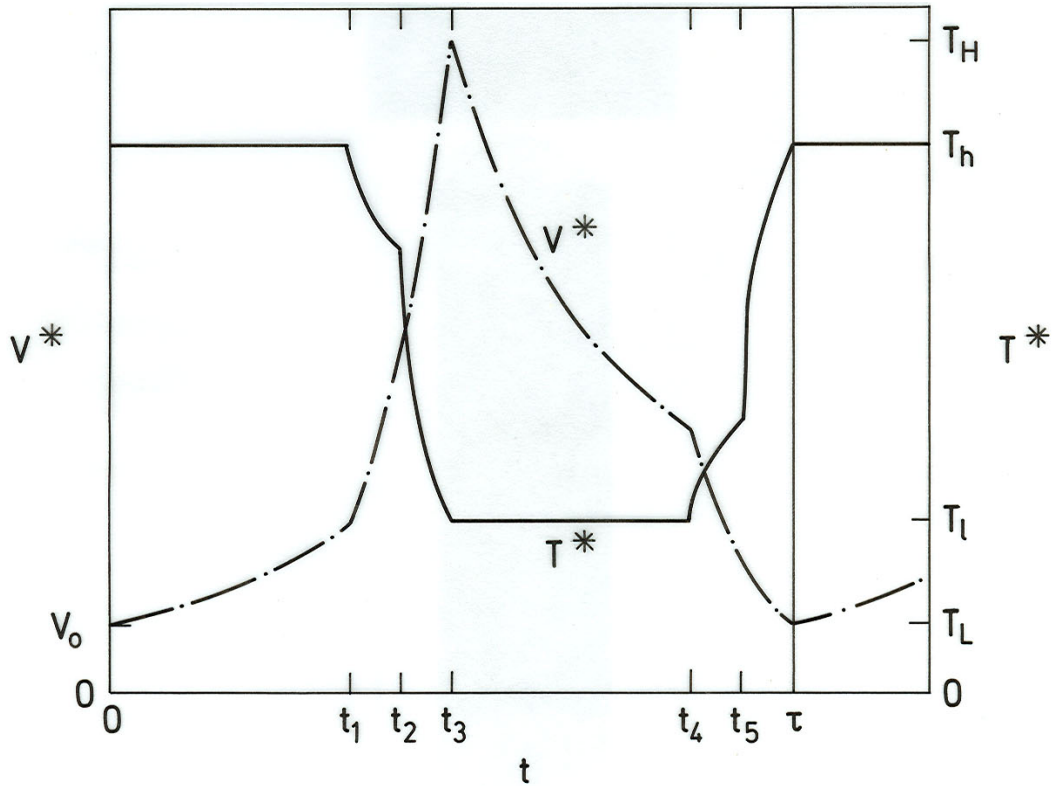


Figure 5. The optimal trajectories associated with the controls in Fig. 4. $V^*(t)$ and $T^*(t)$ are the volume and temperature of the working fluid, respectively.

$$\Delta S \leq \left(\sum_i |\sigma_i| \right)^2 / \kappa \tau, \quad (3.6)$$

where σ_i is the entropy change of the working fluid on branch i . Combine this expression with the equality form of the Second Law of thermodynamics, eq. (1.3), and we see that the work W produced by such an engine is bounded by

$$W \leq W_{\text{rev}} - T_0 \left(\sum_i |\sigma_i| \right)^2 / \kappa \tau, \quad (3.7)$$

where T_0 is the temperature of the environment with respect to which availability is defined. This expression is a particularly clear exposition of one kind of cost of finite-time operation.

The main result, that the ‘best process’ is associated with a constant rate of entropy production, is reminiscent of Prigogine’s (1967) theorem of irreversible thermodynamics which states that the entropy production rate is minimum at near-equilibrium steady states. Both results are derived from variational principles. However, the approach of Prigogine (‘irreversible thermodynamics’) involves instantaneous quantities while our formalism (‘finite-time thermodynamics’) investigates the extrema of integrals over time.

Staging. Endoreversible engines have the same staging property as Carnot engines: If one puts two or more engines of the same kind in sequence, then the whole system behaves as a single engine of the that kind (Rubin and Andresen 1982). In our unconstrained optimization the interface between the two stages, which for the present model is the intermediate temperature and the relative timing of the two engines, is arbitrary and can be used to satisfy other, non-thermodynamic constraints. A possible constraint which could be added to make the problem more realistic concerns the volume swept by the two engines, e.g. the total volume or the sum of the compression ratios. However, if we want to go into such detail, we must also specify the amount of working fluid in each engine and its equation of state, and those new parameters precisely balance the added constraints, so the interface is still arbitrary. This conclusion implies that one of the stages could shrink to zero without it being detected from the outside. There is thus no *thermodynamic* reason for the usual practice of dividing turbines, compressors, refrigerators, and the like into stages.

This staging property of endoreversible engines makes it a unique building block for constructing and analyzing more complex finite-time thermodynamic systems — much like the Carnot engine is in reversible thermodynamics.

3.4 Design of distillation column by equal thermodynamic distance

In Sect. 2.4 we saw that the concept of thermodynamic length not only provided a lower bound on dissipation (eqs. (2.14), (2.17), and (2.22)), it also predicted which path would achieve that bound, namely operation at constant thermodynamic speed. We will use that line of thought to improve the performance of a conventional distillation column.

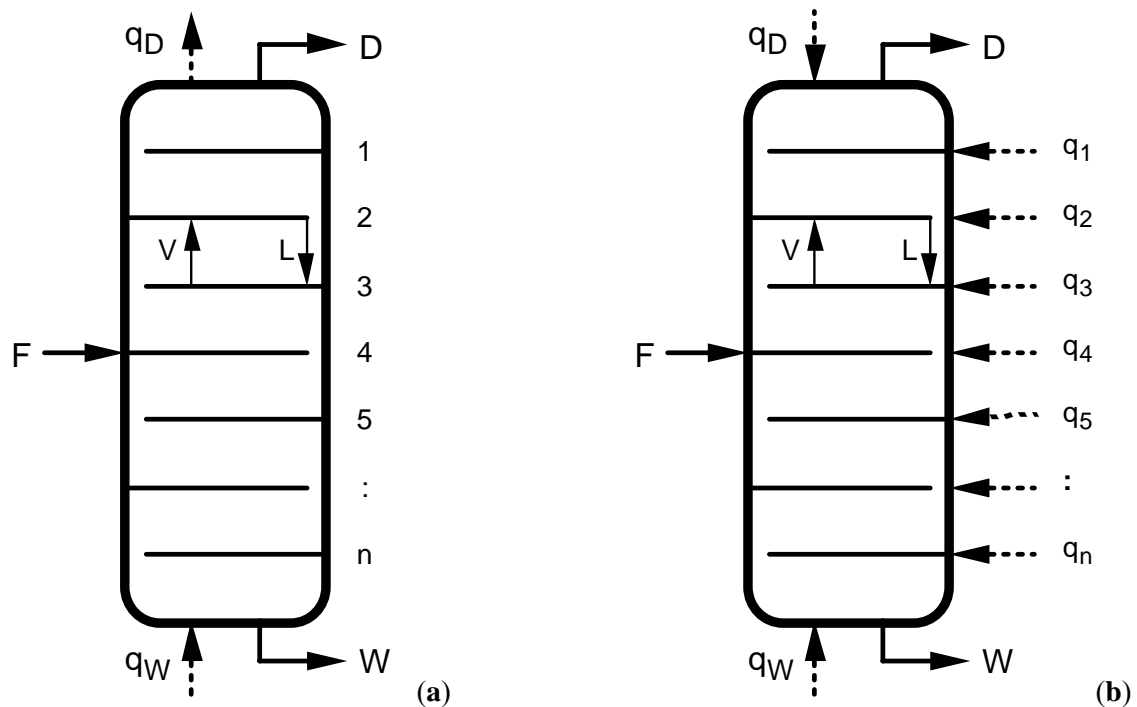


Figure 6. (a) Sketch of a conventional distillation column with feed, distillate, and waste (bottoms) rates F , D , W , heating and cooling rates q_W and q_D , and tray numbers n . (b) Sketch of an equal-thermodynamic-distance distillation column with heating or cooling on all n trays.

In a conventional distillation column, Fig. 6a, heat is added and withdrawn only in the reboiler and the condenser, respectively. In such a column the internal distribution of temperature T and mole fractions x of light and heavy components is fixed exclusively by the laws of energy and mass conservation. In most columns that leads to an S-shaped curve of mole fractions versus plate number with most of the variation occurring near the end points of the column, connected with a flat stretch around the feed point (see Fig. 7). A qualitatively similar picture emerges for the temperature. This profile implies that the major part of the entropy production in the distillation process occurs near the ends of the column and is thus not uniformly distributed. Even more importantly, if the number of plates is increased, essentially only the middle flat section is extended while the segments of rapid variation are unchanged. This means that dissipation does not approach zero as the number of plates goes to infinity.

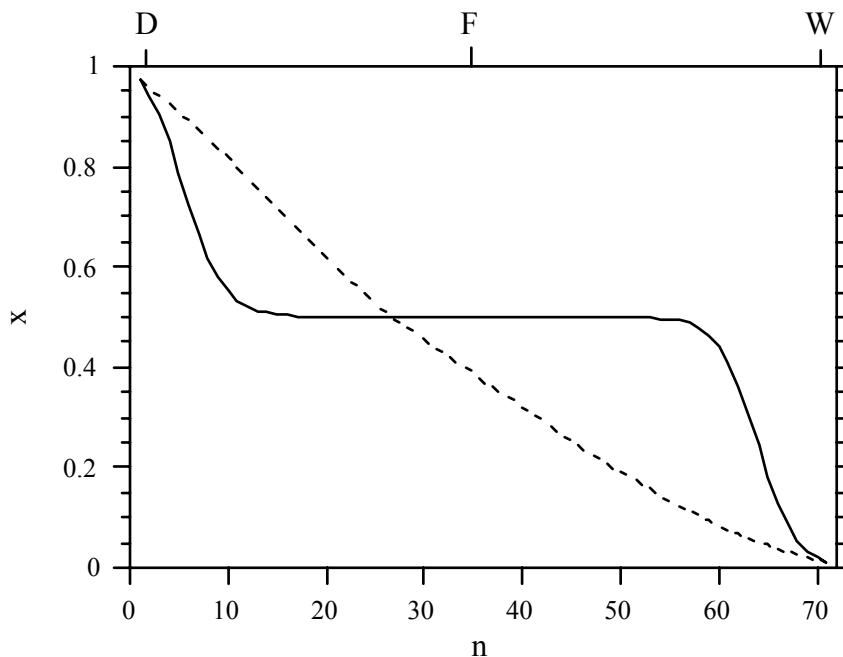


Figure 7. The liquid composition profile x as a function of tray number, counted from the condenser, for conventional (solid) and equal-thermodynamic-distance (dashed) separation of an ideal benzene-toluene system.

The principle of constant thermodynamic speed for continuous systems, or equivalently equal thermodynamic distance for discrete systems, aims at minimizing total dissipation in the column by distributing it evenly among the plates whatever their number (Andresen and Salamon 2000). In Nulton et al. (1985) a general quasistatic step process was optimized, i.e. a process composed of N discrete steps where the system equilibrates fully after each step. The standard description of a distillation column is exactly such a process where it is assumed that gas and liquid come to equilibrium at a particular temperature on each plate. Entropy is produced when the up- and down-moving flows encounter liquid on the next plate at slightly different temperature and composition.

The result of the optimization (Nulton et al. 1985), for processes not too far from equilibrium, was that the minimum of the entropy production

$$\Delta S^u \geq L^2/2N \quad (3.8)$$

is achieved when the thermodynamic distance between each pair of plates is kept constant, i.e. they are equidistant in that measure. L is the total thermodynamic distance from one end of the column to the other. For any other distribution of the total length along the column dissipation is larger.

Using the metric eq. (2.18), the thermodynamic distance D_S^k from plate $k-1$ to plate k is equal to

$$D_S^k = \sqrt{\Delta \mathbf{X}^k \mathbf{M}_S^k \Delta \mathbf{X}^k} \quad (3.9)$$

where $\Delta \mathbf{X}^k = \mathbf{X}^k - \mathbf{X}^{k-1}$ is the difference in extensities from plate $k-1$ to plate k . The total thermodynamic length of the column is thus

$$L_S = \sum_{k=1}^N D_S^k. \quad (3.10)$$

Two consequences of this general result are immediate: dissipation (here entropy production) must be equally distributed along the column; and the total dissipation approaches zero as N , the number of plates, goes to infinity (eq. (3.8)), i.e. the separation becomes reversible.

In binary distillation 8 extensive quantities are involved on each plate, besides the objective function entropy, namely enthalpy, volume, mole number of light component, and mole number of heavy component, each one for both gas and liquid, leading to an 8×8 metric matrix. Fortunately a number of relations allow one to reduce the dimensionality of the problem dramatically. First of all, the two components are usually considered non-interacting, at once making \mathbf{M}_S block-diagonal. Next, constant pressure in the column, energy and mass conservation, and the gas-liquid equilibrium equation reduce the problem to just one free variable which may conveniently be taken to be temperature. After a lot of algebra the plate-to-plate distance finally becomes simply

$$D_S^k = \frac{\sqrt{C_r^k}}{T^k} \Delta T, \quad (3.11)$$

where C_r is an effective heat capacity involving all the constraints mentioned above in a complicated expression.

The computational procedure is to integrate eq. (3.11) from the distillate temperature T_D to the reboiler temperature T_B (both of course given by the required product purities) to obtain the total thermodynamic column length L_S . The distance from one plate to the next must then be fixed at $D_S = L_S/N$ for optimal performance by adjusting the plate temperatures appropriately according to eq. (3.11).

Obviously such freedom of adjustment does not exist in a conventional adiabatic column. Rather, it is necessary to allow individual heat exchange with each plate to maintain it at the desired temperature, see Fig. 6b. This heat addition/removal is of course part of the energy balance used above. The result of the whole calculation is either a graph like Fig. 7 specifying the temperature of each plate in the column or a graph of the amount of heat added/removed at each plate. It should be noted that the total amount of heat used to perform a certain equal-thermodynamic-distance separation is only marginally different from that required by a conventional column, but a large part of it is used over a much smaller temperature difference than T_B to T_D leading to a correspondingly smaller entropy production.

3.5 Simulated annealing.

The global optimization procedure simulated annealing (Kirkpatrick et al. 1983) is intimately connected with a real thermodynamic process. Specifically, if the correspondence with statistical

mechanics implied in the simulated annealing procedure involving phase space and state energies is valid, then further results from thermodynamics will probably also be possible to carry over to simulated annealing.

So far all suggested simulated annealing temperature paths (annealing schedules) have been of the a priori type and thus have not adjusted to the actual behavior of the ‘system’ as the annealing progresses. Examples of such paths are

$$T(t) = a \exp(-t/b) \quad (3.12)$$

$$T(t) = a/(b+t) \quad (3.13)$$

$$T(t) = a/\ln(b+t). \quad (3.14)$$

The real annealing of physical systems often has rough parts where the surrounding temperature must be decreased slowly due to phase transitions or regions of large heat capacity or slow internal relaxation. The same behavior is seen in the abstract systems, so annealing schedules which take such variation into account are preferable in order to keep computation time at a minimum for a given accuracy of the final result (Ruppeiner 1988). Since asking a question (= one evaluation of the energy function) in information theoretic terms is equivalent to producing one bit of entropy, the temperature path $T(t)$ which produces minimum entropy, as calculated with thermodynamic length [c.f. eq. (2.17)], at once suggests itself as the optimal simulated annealing schedule:

$$\frac{dT}{dt} = -\frac{vT}{\varepsilon\sqrt{C}} \quad (3.15)$$

or equivalently

$$\frac{\langle E \rangle - E_{\text{eq}}(T)}{\sigma} = v. \quad (3.16)$$

In these expressions v is the (constant) thermodynamic speed, C and ε are the heat capacity and internal relaxation time of the system, respectively, $\langle E \rangle$ and σ the corresponding mean energy and standard deviation of its natural fluctuations, and finally $E_{\text{eq}}(T)$ is the internal energy the system would have if it were in equilibrium with its surroundings at temperature T . The physical interpretation of eq. (3.16) is that the environment should at all times be kept v standard deviations ahead of the system. Similarly eq. (3.15) indicates that the annealing should slow down where internal relaxation is slow and where large amounts of ‘heat’ have to be transferred out of the system. In case C and ε do not vary with temperature, eq. (3.15) integrates to the standard schedule eq. (3.12).

The extra temperature dependent variables of the constant thermodynamic speed schedule of course require additional computational effort. Since systems often change considerably in a few steps, ergodicity is not fulfilled, so the use of time averages to obtain $\langle E \rangle$, σ , C , and ε is usually not satisfactory. Instead we (Andresen et al. 1988) suggest running an ensemble of systems in parallel, i.e. with the same annealing schedule, in the true spirit of the analogy to statistical mechanics. Then these variables can be obtained anytime as true ensemble averages based on the system degeneracies $p_i = p(E_i)$:

$$Z(T) = \sum_i p_i \exp(-E_i / T) \quad (3.17)$$

$$E(T) = T^2 \frac{d \ln Z}{dt} \quad (3.18)$$

$$C(T) = \frac{dE}{dT} = \frac{\langle (\Delta E)^2 \rangle}{T^2} \quad (3.19)$$

$$\varepsilon(T) = \frac{-1}{\ln \lambda_2} \approx \frac{T^2 C(T)}{\sum_i p_i \sum_{j>i} (E_j - E_i)^2 P_{ji} \exp(-E_i/T)}, \quad (3.20)$$

where λ_2 is the second largest eigenvalue of the thermalized version of the transition probability matrix \mathbf{P} among all the energy levels ($\lambda_1 = 1$ corresponds to equilibrium).

But where does one get the degeneracies p_i from? Actually (Andresen et al. 1988), information to calculate the temperature-independent (or infinite-temperature, if you prefer) transition probability matrix \mathbf{P} can be accumulated during the annealing run by simply adding up in a matrix \mathbf{Q} the number of attempted moves (not just the accepted ones) from level i to j as the calculation progresses. Normalization of \mathbf{Q} yields a good estimate of \mathbf{P} ,

$$P_{ji} = Q_{ji} / \sum_k Q_{ki}. \quad (3.21)$$

The degeneracies \mathbf{p} are then the eigenvector of \mathbf{P} corresponding to the eigenvalue 1.

This use of ensemble annealing is particularly well suited for implementation on present day parallel computers. A further analysis of its performance has been carried out by Ruppeiner et al. (1990), and the trade-off between ensemble size and duration of annealing for fixed total computation cost has been addressed by Pedersen et al. (1990).

4. Summary

Finite-time thermodynamics was ‘invented’ in 1975 by R. S. Berry, P. Salamon, and myself as a consequence of the first world oil crisis. It simply dawned on us that all the existing criteria of merit were based on reversible processes and therefore were totally unrealistic for most real processes. That made an evaluation of the potential for improvement of a given process quite difficult.

Finite-time thermodynamics is developed from a macroscopic point of view with heat conductances, friction coefficients, overall reaction rates, etc. rather than based on a microscopic knowledge of the processes involved. Consequently most of the ideas of traditional thermodynamics have been assimilated, e.g. the notions of thermodynamic potential (Sect. 2.1) and availability (Sect. 2.2). At the same time we have seen new concepts emerge, e.g. the non-equivalence of well-honored criteria of merit (Sect. 3.2), the importance of power as the objective (Sect. 1.2 and 3.2), the generality of the endoreversible engine (Sect. 3.3), and in particular thermodynamic length (Sect. 2.4). Several of these abstract concepts have been successfully applied to practical optimizations (see e.g. Hoffmann et al. 1985, Mozurkewich and Berry 1981, 1982, Hoffmann et al. 1997, 2003). In line with the global philosophy of finite-time thermodynamics Gordon and Zarmi (1989) have calculated the global wind pattern.

More recently the notions and results of finite-time thermodynamics, at times in connection with statistical mechanics and information theory, have been used to perfect the global optimization method simulated annealing (Sect. 3.5). However, the surface is only scratched, there is still plenty of room for inspiration from such well-known concepts as state entropy, free energy, and transition state theory.

5. Acknowledgments

Much of the work reported in this review has been done in collaboration with colleagues near and far. In particular I want to express my gratitude to profs. R. Stephen Berry, Peter Salamon, Jim Nulton, and Karl Heinz Hoffmann.

References

- Andresen, B. and Gordon, J. M. 1994. On constant thermodynamic speed for minimizing entropy production in thermodynamic processes and simulated annealing, *Phys. Rev. E* **50**, 4346.
- Andresen, B. and Salamon, P. 2000. Optimal distillation using thermodynamic geometry. In Sieniutycz, S. and De Vos, A. (eds.) *Thermodynamics of energy conversion and transport* (Springer, New York), p. 319.
- Andresen, B., Hoffmann, K. H., Mosegaard, K., Nulton, J., Pedersen, J. M. and Salamon, P. 1988. On lumped models for thermodynamic properties of simulated annealing problems, *J. Phys. (France)* **49**, 1485.
- Andresen, B., Rubin, M. H. and Berry, R. S. 1983. Availability for finite-time processes: General theory and a model, *J. Chem. Phys.* **87**, 2704.
- Boltyanskii, V. G. 1971. *Mathematical methods of optimal control* (Holt, Reinhart and Winston, New York).
- Callen, H. B. 1985. *Thermodynamics* (Wiley, New York).
- Curzon, F. L. and Ahlborn, B. 1975. Efficiency of a Carnot engine at maximum power output, *Am. J. Phys.* **43**, 22.
- Feldmann, T., Andresen, B., Qi, A. and Salamon, P. 1985. Thermodynamic lengths and intrinsic time scales in molecular relaxation. *J. Chem. Phys.* **83**, 5849.
- Flick, J. D., Salamon, P. and Andresen, B. 1987. Metric bounds on losses in adaptive coding. *Info. Sci.* **42**, 239.
- Gaggioli, R. A. (ed.) 1980. *Thermodynamics: Second law analysis*, ACS Symposium series **122** (American Chemical Society, Washington, D. C.).
- Geusic, G., Schulz-Dubois, E. O. and Scovil, H. E. D. 1967. Quantum equivalent of the Carnot cycle, *Phys. Rev.* **156**, 343.
- Gordon, J. M. 1989. Maximum power point characteristics of heat engines as a general thermodynamic problem, *Am. J. Phys.* **57**, 1136.
- Gordon, J. M. and Ng, K. C. 1994. Thermodynamic modeling of reciprocating chillers, *J. Appl. Phys.* **57**, 2769.
- Gordon, J. M. and Ng, K. C. 2000. *Cool thermodynamics: The engineering and physics of predictive, diagnostic and optimization methods for cooling systems* (Cambridge International, Cambridge).
- Gordon, J. M. and Zarmi, Y. 1989. Wind energy as a solar-driven heat engine: A thermodynamic approach, *Am. J. Phys.* **57**, 995.
- Hermann, R. 1973. *Geometry, physics, and systems* (Marcel Dekker, New York).
- Hoffmann, K. H., Burzler, J. and Schubert, S. 1997. Endoreversible systems, *J. Non-Equil. Therm.* **22**, 4.

- Hoffmann, K. H., Burzler, J., Fischer, A., Schaller, M. and Schubert, S. 2003. Optimal process paths for endoreversible systems, *J. Non-Equil. Therm.* **28**, 233.
- Hoffmann, K. H., Watowitch, S. and Berry, R. S. 1985. Optimal paths for thermodynamic systems: The ideal diesel cycle, *J. Appl. Phys* **58**, 2125.
- Keenan, J. H. 1941. *Thermodynamics* (MIT Press, Cambridge, Massachusetts).
- Kirkpatrick, S., Gelatt, C. D. Jr. and Vecchi, M. P. 1983 Optimization by simulated annealing, *Science* **220**, 671.
- Leondes, C. T. (ed.). 1964 and onward. *Advances in control systems* (Academic Press, New York) [especially articles in the first volumes].
- Levine, R. D. and Tribus, M. 1979. *The maximum entropy formalism* (MIT Press, Cambridge, Massachusetts).
- Mozurkewich, M. and Berry, R. S. 1981. Finite-time thermodynamics: Engine performance improved by optimized piston motion, *Proc. Natl. Acad. Sci. USA* **78**, 1986.
- Mozurkewich, M. and Berry, R. S. 1982. Optimal paths for thermodynamic systems: The ideal Otto cycle, *J. Appl. Phys* **53**, 34.
- Nulton, J. and Salamon, P. 1988. Statistical mechanics and combinatorial optimization, *Phys. Rev. A* **37**, 1351.
- Nulton, J., Salamon, P., Andresen, B. and Anmin, Q. 1985. Quasistatic processes as step equilibrations, *J. Chem. Phys.* **83**, 334.
- Onsager, L. 1931a. Reciprocal relations in irreversible processes. I, *Phys. Rev.* **37**, 405.
- Onsager, L. 1931b. Reciprocal relations in irreversible processes. II, *Phys. Rev.* **38**, 2265.
- Pedersen, J. M., Mosegaard, K., Jacobsen, M. O. and Salamon, P. 1990. Optimal ensemble size for parallel implementations of simulated annealing. *Appl. Math. Lett.* **3**, 53.
- Prigogine, I. 1962. *Non-equilibrium statistical mechanics* (Wiley, New York).
- Prigogine, I. 1967. *Introduction to thermodynamics of irreversible processes* (Wiley, New York), pp. 76 ff.
- Rubin, M. H. 1979. Optimal configuration of a class of irreversible heat engines. II, *Phys. Rev. A* **19**, 1277.
- Rubin, M. H. and Andresen, B. 1982. Optimal staging of endoreversible heat engines, *J. Appl. Phys.* **53**, 1.
- Ruppeiner, G. 1988. Implementation of an adaptive, constant thermodynamic speed simulated annealing schedule. *Nucl. Phys. B (Proc. Suppl.)* **5A**, 116.
- Ruppeiner, G., Pedersen, J. M. and Salamon, P. 1991. Ensemble approach to simulated annealing. *J. Phys. I.* **1**, 455.
- Salamon, P. and Berry, R. S. 1983. Thermodynamic length and dissipated availability, *Phys. Rev. Lett.* **51**, 1127.
- Salamon, P. and Nitzan, A. 1981. Finite time optimizations of a Newton's law Carnot cycle, *J. Chem. Phys.* **74**, 3546.
- Salamon, P., Andresen, B. and Berry, R. S. 1977. Thermodynamics in finite time. II. Potentials for finite-time processes, *Phys. Rev. A* **15**, 2094.
- Salamon, P., Andresen, B., Gait, P. D. and Berry, R. S. 1980a. The significance of Weinhold's length, *J. Chem. Phys.* **73** 1001; erratum *ibid.* **73**, 5407E.

- Salamon, P., Komlos, J., Andresen, B. and Nulton J. 1987. A geometric view of welfare gains with non-instantaneous adjustment. *Math. Soc. Sci.* **13**, 153.
- Salamon, P., Nitzan, A., Andresen, B. and Berry, R. S. 1980b. Minimum entropy production and the optimization of heat engines, *Phys. Rev. A* **21**, 2115.
- Salamon, P., Nulton, J. and Ihrig, E. 1984. On the relation between entropy and energy versions of thermodynamic length. *J. Chem. Phys.* **80**, 436.
- Tisza, L. 1966. *Generalized thermodynamics* (MIT Press, Cambridge, Massachusetts).
- Tolle, H. 1975. *Optimization methods* (Springer, New York).
- Tolman, R. C. and Fine, P. C. 1948. On the irreversible production of entropy, *Rev. Mod. Phys.* **20**, 51.
- Weinhold, F. 1975a. Metric geometry of equilibrium thermodynamics, *J. Chem. Phys.* **63**, 2479.
- Weinhold, F. 1975b. Metric geometry of equilibrium thermodynamics. II. Scaling, homogeneity, and generalized Gibbs-Duhem relations, *J. Chem. Phys.* **63**, 2484.
- Weinhold, F. 1975c. Metric geometry of equilibrium thermodynamics. III. Elementary formal structure of a vector-algebraic representation of equilibrium thermodynamics, *J. Chem. Phys.* **63**, 2488.
- Weinhold, F. 1975d. Metric geometry of equilibrium thermodynamics. IV. Vector-algebraic evaluation of thermodynamic derivatives, *J. Chem. Phys.* **63**, 2496.
- Weinhold, F. 1978. Geometrical aspects of equilibrium thermodynamics, in *Theoretical chemistry, advances and perspectives*, Vol. 3, edited by D. Henderson and H. Eyring (Academic Press, New York).
- Wheatley, J., Hofler, T., Swift, G. W. and Migliori, A. 1983. Experiments with an intrinsically irreversible acoustic heat engine, *Phys. Rev. Lett.* **50**, 499.

Problems

1. Curzon-Ahlborn engine – Consider the endoreversible engine of Fig. 1. Calculate the maximum power w_{\max} that can be extracted from this engine and its thermal efficiency η_w at this maximum power, i.e. repeat Curzon and Ahlborn's discovery. For simplicity you may assume that the conductances are equal, $\kappa_h = \kappa_l = \kappa$. Although the physical variables which are free to be varied in order to achieve this maximum are the inner temperatures T_h and T_l , calculations become simpler if you replace one of them by either the ratio T_h/T_l or by the inner Carnot efficiency $1 - T_l/T_h$ as free variable.

2. Gibbs potentials – Calculate the work and heat potentials P_W and P_Q for an isothermal isobaric process.

3. Ammonia work potential – The ideal gases H_2 , N_2 , and NH_3 are enclosed in a variable volume V at isothermal temperature T . The initial amounts are 3 moles of hydrogen, 1 mole of nitrogen, and no ammonia. The compounds are considered to be in chemical equilibrium at all times.

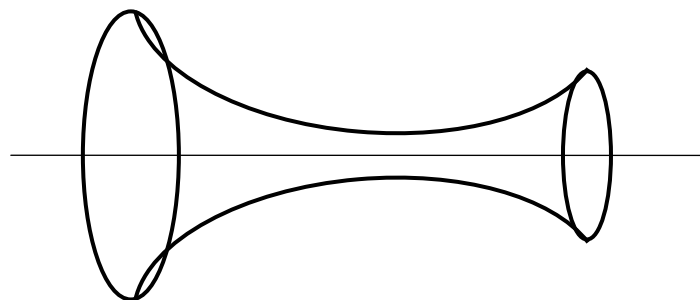
- Find the thermodynamic work potential P_W for this reacting gas. Show step by step which constraints are involved and how they lead to the potential. It will prove convenient to work with the reaction coordinate $\xi = \text{amount of } NH_3$ rather than the individual concentrations.
- (ambitious) Find the thermodynamic work potential P_W for the reacting gas under adiabatic instead of isothermal conditions. For this calculation you need the heat of reaction which may be assumed constant. Likewise, the heat capacity of the gas mixture can be taken to be independent of the extent of reaction and of temperature.

4. Optimal shape of hanging rope – A rope of length L and mass density (mass per length) ρ is suspended between points $A = (x_1, y_1)$ and $B = (x_2, y_2)$ under the influence of gravity. It is assumed that L is larger than the distance between A and B .

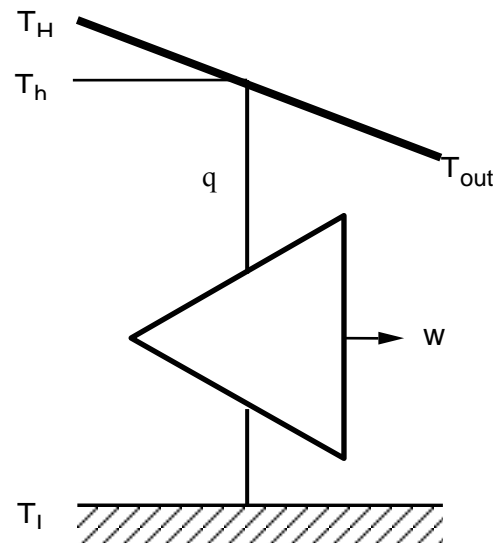
- Calculate the shape of the rope, $y(x)$.
- Is this shape different on Earth and on the Moon? Why?

5. Minimum surface area – Two disks of radii R_1 and R_2 are placed a distance L apart on the same rotational axis. A "rubber stocking" connects the two disks creating a volume of revolution resembling somewhat the shape of an hourglass.

- Calculate the shape of the minimum surface area of this volume for all R_1 , R_2 , and L .
- Compare this 3-dimensional result with the 2-dimensional result of the hanging rope.



6. Criteria of merit – Consider the Curzon-Ahlborn-like engine below. It is connected to its low temperature reservoir without heat resistance. The high temperature source is a heat exchanger with fluid entering at the fixed temperature T_H and leaving at T_{out} , the latter depending on how much heat is removed to the engine, i.e. $T_{out} = T_H - q/c_p$, where q is the heat rate entering the heat engine and c_p is the heat capacity of the stream. For simplicity assume that all heat enters the engine at $T_h = (T_H + T_{out})/2$. Thus the interior engine operates between T_h and T_L .



- Calculate the maximum power w which may be produced, the corresponding T_{out} , and the rate of entropy production s in this setup in the following two cases: (i) Heat leaving the heat exchanger at T_{out} may be used for something else and thus carries value; or (ii) this heat has no further value and is discarded into the environment at T_L .
- Calculate the minimum entropy production s in cases (i) and (ii) above as well as the corresponding power production w and T_{out} .
- Explain why maximum power production and minimum entropy production are not always the two opposites on the “rate of operation axis”.