

Thermodynamics for Processes in Finite Time

BJARNE ANDRESEN

Physics Laboratory II, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

R. STEPHEN BERRY*

Department of Chemistry and the James Franck Institute, The University of Chicago, Chicago, Illinois 60637

MARY JO ONDRECHEN

Department of Chemistry, Northeastern University, Boston, Massachusetts 02115

PETER SALAMON

Department of Mathematical Sciences, San Diego State University, San Diego, California 92182

Received September 1, 1983

Thermodynamics serves several functions for chemistry and its sister sciences and technologies. The first of these historically, and still one of the most important, is the provision of limits on the performance of processes and devices. The very origins of thermodynamics developed from the creative ways people addressed the problem of how best to pump water out of mines.¹ The provision of limits of performance is the main concern of the work we describe here. Research during the past 9 years has led to new ways of finding natural bounds on performance under the constraint that the system operate at a nonzero rate, thereby giving more realistic bounds than those derived from reversible processes. Some approaches tell us only the limits; others tell us also the process that would yield those limits. Frequently, we can infer how the lower bound on the energy we pay for operating at a nonzero rate depends on that rate of operation. There even exist processes for which no zero-rate reversible counterpart exists.²

The difference between the performance of a reversible process and a real-time process may be considerable, so great that relying on the bounds based on reversible models can be quite misleading. For example, a theoretical heat engine operating between 300 K and 2700 K with friction, finite heat conductance between the engine and its reservoirs, and a heat leak,³ all chosen to correspond approximately to an automobile engine, has an ideal Carnot efficiency of $(2700-300)/2700 = 0.89$. The efficiency of the engine when operating to generate maximum power is only 0.33.

Another example is the frequent use of roasting with subsequent reduction in ore processing, both processes

which usually are carried out with large expenditures of energy, even in cases when the metal has lower free energy content than the ore, and the refining therefore in principle could have *supplied* work. In short, the constraints attending operating in finite time reduce the limits on performance so much that it is desirable to find general methods to calculate more realistic limits that reflect those constraints.

Extension to Finite-Time Thermodynamics

The first natural pathway to those more realistic criteria is the extension of the concepts of thermodynamic potential⁴ and availability⁵ to incorporate the constraints on time or rate with the other constraints—such as constant temperature, pressure, or volume—in the construction of the potential (see later). Such extensions of thermodynamic potentials, traditionally functions only of the state variables of the system itself, to include time or rate, are justified by some existence theorems that need not be reproduced here.⁴ In many cases one wants more than just the maximum work that can be produced in a certain time. One may need to know the detailed time path that yields this maximum work. The tool to obtain this is optimal control theory. Applying this method is usually much more involved than finding a potential and must often be done numerically, but in return the optimal time path contains all the information about the optimized system, including the maximum work it can provide—much as the wave function is the source of all information in quantum mechanics. The optimization of the Otto cycle illustrates this procedure.⁶ In contrast to the usual *parameter* optimizations done in engi-

Bjarne Andresen was born in Copenhagen in 1949. He studied at Wesleyan University and at Cal Tech prior to receiving his Ph.D. from the University of Copenhagen in 1974. He is now a Lecturer in Physics at the H. C. Ørsted Institute, University of Copenhagen.

R. Stephen Berry was born in Denver, Colorado, in 1931. He received his A.B. (1952) and his Ph.D. (1956) from Harvard University. He was an instructor at Harvard and the University of Michigan and then moved to Yale University as an Assistant Professor. In 1964 he moved to the University of Chicago and became a Professor there in 1967, and he remains there.

Mary Jo Ondrechen was born in Philadelphia in 1953. She received her B.A. at Reed College in 1974 and her Ph.D. at Northwestern University in 1978. She is now an Assistant Professor at Northeastern University.

Peter Salamon received his B.A. in 1971 from Lindenwood College, his M.S. from Drexel University in 1972, and his Ph.D. from the University of Chicago in 1978. In 1981 he became Associate Professor at San Diego State University and remains there now.

(1) D. S. L. Cardwell, "From Watt to Clausius", Cornell University Press, Ithaca, NY, 1971.

(2) J. Wheatley, T. Hoffer, G. W. Swift, and A. Migliori, *Phys. Rev. Lett.*, **50**, 499 (1983); M. Mozurkewich and R. S. Berry, *J. Appl. Phys.*, **54**, 3651 (1983).

(3) B. Andresen, P. Salamon, and R. S. Berry, *J. Chem. Phys.*, **66**, 1571 (1977).

(4) P. Salamon, B. Andresen, and R. S. Berry, *Phys. Rev. A*, **15**, 2094 (1977).

(5) B. Andresen, M. H. Rubin, and R. S. Berry, *J. Phys. Chem.*, **87**, 2704 (1983).

(6) M. Mozurkewich and R. S. Berry, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 1986 (1981); *J. Appl. Phys.*, **53**, 34 (1982).

neering design, in this calculation, the parameters of period, heat transfer, fuel assumption, etc., were fixed at conventional values, and only the time path of the piston was varied from its conventional sinusoidal form to optimize the power. Optimizing the piston motion alone could increase the power output of the engine by 8–15%.

The maximum work a system can do in a given process is one index of performance; it is not the only one we could choose. Some others are as follows: the theoretical, or Carnot, efficiency $\eta_c = 1 - (T_L/T_H) = (\text{ideal work, } W)/(\text{heat to produce } W)$; the thermal efficiency $\eta = (\text{work actually delivered, } W_e)/(\text{heat to produce } W_e)$; the effectiveness $\epsilon = W_e/(\text{change in availability})$; the power; the dissipated availability (availability lost and not transformed to work); the entropy produced; and the net revenue. Some criteria of performance, such as power, are not meaningful for reversible processes. Others, such as entropy generated, can be separated into loss terms that have no relation to the rate of the process and other loss terms that depend specifically on time or rate. Entropy of mixing is of the first kind, and ohmic heating and other friction-based heat generation processes are clearly of the second kind. A system optimized by one criterion need not be optimal with respect to others. A later section illustrates differences that follow from different choices of criterion.

One goal of finite-time thermodynamics is the identification of generic models that can play roles of great generality, in the way the Carnot engine is the archetypical heat engine in reversible thermodynamics. Assuredly, one uses many other idealized models such as the air-standard cycles (Otto, Diesel, Stirling, Brayton, etc.) to describe classes of real engines approximately but more accurately than with the Carnot engine, even when retaining the conditions of reversibility. With the introduction of time constraints and the corresponding time-dependent loss processes, the possibilities might seem to open without limit. It appears now, however, that one may need only a few idealized paradigms for finite-time behavior.

One generic class consists of mechanical processes driven by heat, for which the irreversibilities occur entirely at the boundaries between system and surroundings. These systems are called "endoreversible", and two subclasses have been identified until now, those dominated by constraints on heat transfer between the system and its reservoirs and those dominated by friction and heat leaks. Systems with both kinds of losses divide³ into sets dominated by one kind of irreversibility or the other.

Mechanical Processes Driven by Heat

We will illustrate some of the consequences of finite-time operation by examining several generic processes. First, consider systems consisting of a working fluid which may be connected to a range of heat reservoirs, operating either in a cycle (e.g., a reciprocating steam engine) or in continuous flow (e.g., a steam turbine). The most general approach to such heat engines is provided by the black box model known as the tricycle formalism,³ which only involves energy balance and the second law, and keeps track of entropy flow and generation, without requiring any specific knowledge of how the process works in detail. An early model of a laser used this idea⁷ but without developing its gen-

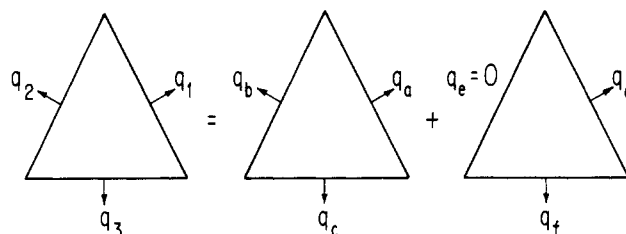


Figure 1. Tricycle decomposition of a general heat process into its reversible and irreversible parts. Each side of the tricycle represents heat flow q_i into a reservoir with temperature T_i . The input from the hot reservoir is q_2 ; q_b is set equal to q_2 to make all losses accountable to the work (q_1) and waste heat (q_3) flows. The first tricycle on the right, being reversible, has zero entropy production per cycle.

described as flows between a working system and three reservoirs. In this way simple heat engines can be represented by a triangle, standing for the system, with flows through each side of the triangle representing flows between the system and one of the (heat or work) reservoirs. This pictorial representation of a cyclic process is the "tricycle". More complex systems may require more than one triangle for their representation.

The crux of the tricycle method is the decomposition of the real process into a reversible part and an irreversible part, as illustrated in Figure 1. Some processes, such as separation, are best decomposed into three parts: a reversible part, a part containing the inherently irreversible contributions that occur even with infinitely slow operation, and a third part that contains the time-dependent irreversibilities. By choosing generic expressions for the most important losses—friction, finite heat conductance, and heat leakage—in terms of system parameters and the cycle time, it becomes possible to optimize the performance of the engine.

The important finding³ is that the optimal operation of the general heat engine shows a bifurcation, as a function of the ratio of friction to heat resistance, into two classes: those which are friction dominated and those which are dominated by resistance to heat flow.

Curzon and Ahlborn⁸ considered a model of an engine which is a particularly simple yet rich elaboration of the Carnot engine. The Curzon–Ahlborn engine has a cycle of two adiabats and two isotherms and suffers losses due only to thermal resistance between the reservoirs at temperatures T_H and T_L and the working fluid. We shall refer to such an engine as a CA engine. Its Carnot efficiency is

$$\eta_c = 1 - T_L/T_H \quad (1)$$

obtained for infinitely slow operation. Curzon and Ahlborn⁸ asked not about maximum efficiency but instead about the efficiency *when the engine produces maximum power* and found it to be

$$\eta_w = 1 - (T_L/T_H)^{1/2} \quad (2)$$

independent of the thermal conductance κ . This latter efficiency, Curzon and Ahlborn found, corresponds considerably closer to the actual performance of power plants than does η_c !

In general, the CA engine can be optimized according to a variety of criteria, subject to the constraints of the reservoir temperatures and a fixed time τ for the period

(7) J. Geusic, E. O. Schulz-duBois, and H. E. D. Scovil, *Phys. Rev.*, **156**, 343 (1967).

(8) F. L. Curzon and B. Ahlborn, *Am. J. Phys.*, **43**, 22 (1975).

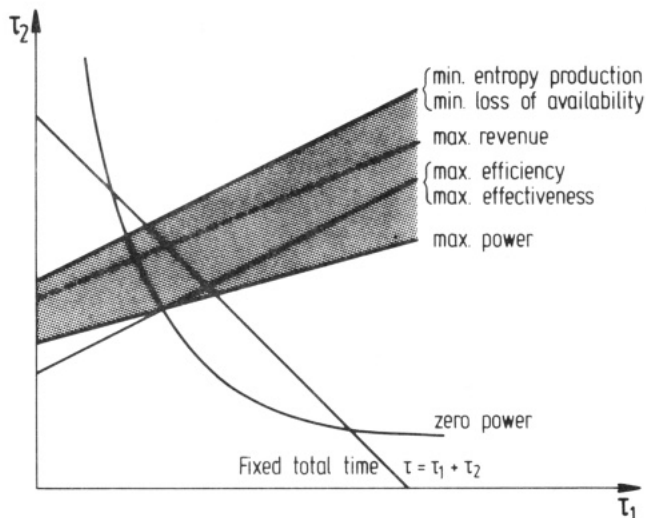


Figure 2. Distribution of the times an endoreversible engine is in contact with its hot reservoir (τ_1) and its cold reservoir (τ_2) when different criteria of performance are optimized. The location of solutions corresponding to maximum revenue (see text) is situated in the shaded region.

of the cycle. Control to attain the optimum behavior is achieved by varying the portion of τ allotted to each branch of the cycle. Adiabatic branches are ordinarily taken to be instantaneous in analyses of this kind, since no dissipation is associated with them. Thus the optimization can be described by the durations of the high- and low-temperature isotherms τ_1 and τ_2 where $\tau_1 + \tau_2 = \tau$; a graph showing the optima⁹ is presented in Figure 2. The loci of solutions for the CA engine, determined by maximizing power, maximizing efficiency, minimizing entropy production, and minimizing loss of availability are straight lines. Note that for CA engines minimizing entropy production is equivalent to minimizing loss of availability.¹⁰

We can identify the rate of revenue R for the process as a linear combination of power and rate of loss of availability, $R = \alpha P + \beta \dot{A}$, because the power P is a saleable good and the availability A represents input resources, so \dot{A} is negative; α and β represent the prices of power and availability.⁹ Then, for any pair of positive prices α and β optimal R defines a line between the line of maximized power and the line of minimized loss of availability. Hence all economically optimal solutions for any pair of prices α , β lie inside the shaded region of Figure 2. If resources are free, R is equivalent to P , and if power is given away, R is equivalent to \dot{A} . This is an example of how one can bound the optimal solution with limits determined by nature, without knowing the exact criterion of performance.

A more general class of heat engines whose reservoirs have variable temperatures can be operated to minimize the total entropy production.¹⁰ When this optimization is applied, the rate of entropy production is a constant, given by

$$\Delta S_u^{\min} / \tau = (\sum_i |\sigma_i|)^2 / \kappa \tau^2 \quad (3)$$

where σ_i is the entropy change of the working fluid along branch i . This result holds also for an endorev-

ersible engine in which the working fluid receives generalized fluxes from reservoirs, provided such fluxes are functions, however nonlinear, of the generalized forces at the boundary but are not functions of time or of any time derivatives of the forces. The maximum power of such engines is given by¹¹

$$P_{\max} = \kappa \text{Variance} (T_{\text{res}}(t))^{1/2} \quad (4)$$

where Variance refers to the time average of the square deviation of $(T_{\text{res}}(t))^{1/2}$ from its mean value.

These examples have either assumed a specific cycle (Carnot in the case of CA engines) or focused on the temperature variations of the working fluid, sidestepping the volume variations necessary to achieve them and their possible restrictions. Rubin¹² made the first complete optimal control calculations of the general endoreversible engine, the only restrictions being limits on piston velocity, heat conductance between the working fluid and the reservoir, and the range of available reservoir temperatures. Fairén and Ross¹³ have studied constraints due to inertial effects. A qualitative summary of these findings is that the working fluid should optimally accept and reject heat isothermally, at temperatures which minimize the losses across the thermal resistance, and "jump" from one isotherm to the other as quickly as the constraints allow. If there are no limits on piston velocity, this means instantaneously, i.e., adiabatically, thus recovering the interior Carnot cycle assumed in the Curzon-Ahlborn analysis.⁸ In addition all 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 that kind.¹⁴ This makes the endoreversible engine a unique building block for analyzing larger finite-time thermodynamic systems.

Much progress has also been made on the form of the optimal trajectories for specific working fluids. Such analyses typically include considerations of friction,^{3,6,15,16} inertia,¹³ and other losses and/or constraints of operation and must eventually resort to numerical methods.

Processes Driven by Chemical Reactions

Many familiar power-producing systems are driven by heat generated by an exothermic chemical reaction. Such engines have two important features. Firstly, the heat carrier is generally finite in size and, secondly, there is usually a limit on the rate of production of the heat used to drive the engine.

The source of heat for a chemically driven engine is the reaction product mixture. Because this has finite heat capacity, although it is initially at temperature T_H , heat from the mixture cannot be converted into work with the Carnot efficiency η_c , since the temperature of the high-temperature reservoir decreases as heat is transferred from the mixture. For a reversible engine working between a heat source with finite, constant heat

(11) P. Salamon, Y. B. Band, and O. Kafri, *J. Appl. Phys.*, **53**, 197 (1982).

(12) M. H. Rubin, *Phys. Rev. A*, **19**, 1272 (1979); *ibid.*, **19**, 1277 (1979); *ibid.*, **22**, 1741 (1980).

(13) V. Fairén and J. Ross, *J. Chem. Phys.*, **75**, 5485 (1981).

(14) M. H. Rubin and B. Andresen, *J. Appl. Phys.*, **53**, 1 (1982).

(15) Y. B. Band, O. Kafri, and P. Salamon, *J. Appl. Phys.*, **53**, 8 (1982).

(16) V. Fairén and J. Ross, *J. Chem. Phys.*, **75**, 5490 (1981).

(9) P. Salamon and A. Nitzan, *J. Chem. Phys.*, **74**, 3546 (1981).

(10) P. Salamon, A. Nitzan, B. Andresen, and R. S. Berry, *Phys. Rev. A*, **21**, 2115 (1980).

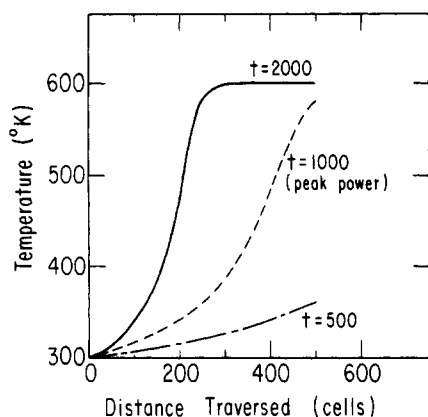


Figure 3. Optimal temperature profile along the reactor tube for an exothermic reaction with a temperature-dependent rate constant obeying an Arrhenius law. The dashed curve with $t = 1000$ corresponds to maximum power output.

capacity initially at T_H and an infinitely large cold reservoir at T_L , the reversible efficiency is given by¹⁷

$$\eta_f = 1 - \frac{T_L}{T_H - T_L} \ln(T_H/T_L) < 1 - \frac{T_L}{T_H} \quad (5)$$

for the case where the source is cooled all the way down to T_L . Further discussion of the production of work from finite heat sources is given in ref 17 and 18.

We have thus far discussed systems whose time dependence enters through friction, heat conductance, or heat loss. In other systems, performance is limited by the rate at which heat is generated. Consider a continuous flow reactor that supplies heat to an engine—for example, a combustion-heated, steam-driven electric generator. For a finite reaction tube and a nonzero flow rate, the reaction will not, in general, go to completion, and the temperature of the emerging product mixture will depend upon the extent of the reaction. On the other hand the reactor must be run at a nonzero flow rate in order to produce power which clearly entails some sacrifice of fuel efficiency.

Conditions have been established for the achievement of maximum power from such a system.¹⁹ For reactions following first-order, first-order reversible, second-order, and second-order bilinear kinetics, the maximum power is attained at a finite, positive flow rate; maximum fuel efficiency and minimum entropy production are achieved in the uninteresting limit of no flow. As one might expect, power production approaches zero at very slow flow rates (because heat is delivered to the engine very slowly) and at very high flow rates (because the reaction hasn't time to go to completion). A temperature profile for the reactor (temperature as a function of distance along the reactor tube) in Figure 3 shows how the temperature increases slowly at first and then increases quite suddenly because of the positive feedback due to the temperature dependence of the rate constant. At the flow rate corresponding to maximum power, the sudden rise in temperature—the

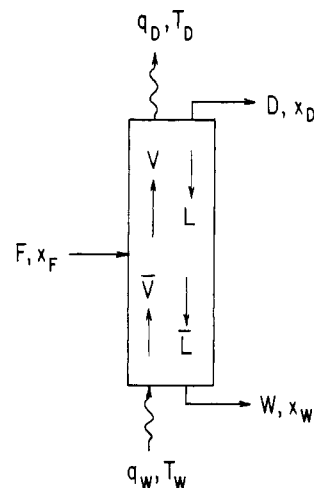


Figure 4. Diagram of a distillation column with feed entering at F , heavy component leaving at B , and light component leaving at D . Heat is supplied at temperature T_B and withdrawn at T_D . In the figure W for waste is used instead of B for bottoms.

“combustion zone”—occurs just before the end of the reactor tube. The maximum power obtainable from such a process is a very sensitive and strictly decreasing function of the activation energy of the reaction which drives the system. Furthermore, the flow rate yielding maximum power is itself a decreasing function of the activation energy.

Systems Doing Chemical Work

The theory of separation, and in particular isotope separation by distillation, got a tremendous boost during and after World War II for obvious reasons. The development was summed up in the 1950's by its main contributors, Benedict and Pigford²⁰ and Cohen and Murphy.²¹ Relatively few new ideas have appeared since then.

In conventional binary distillation feed of composition (mole fraction of light component) x_F is separated into distillate of composition x_D and bottoms of composition x_B by passing heat Q through the distillation column from temperature T_B at the bottom to T_D at the top (see Figure 4). There is a trade-off, largely economic, between the number of plates in the column, i.e., capital, and heat load required to perform a certain separation, i.e., operating cost. Even when heat requirements are pushed to their minimum, the effectiveness of separating a mixture of similar compounds into pure products by distillation, expressed as the reversible work of separation divided by the work equivalent of the heat used in the distillation process (heat used multiplied by $\eta_c(T_B, T_D)$), is approximately only

$$\epsilon = -[x_F \ln x_F + (1 - x_F) \ln (1 - x_F)] < 1 \quad (6)$$

This function has a maximum of 70% for an equimolar mixture, $x_F = 0.5$. The remaining 30% of the heat is lost due to the inherent thermodynamics of the process and cannot be avoided without altering the distillation process itself, e.g., by adding extra boilers and condensers along the column.

(20) M. Benedict and T. F. Pigford, “Nuclear Chemical Engineering”, McGraw-Hill, New York, 1957.

(21) K. Cohen and G. M. Murphy, “The Theory of Isotope Separation as Applied to the Large-Scale Production of U^{235} ”, McGraw-Hill, New York, 1951.

(17) M. J. Ondrechen, B. Andresen, M. Mozurkewich, and R. S. Berry, *Am. J. Phys.*, **49**, 681 (1981); F. d'Isep and L. Sertorio, *Nuovo Cimento B*, **67**, 41 (1982).

(18) M. J. Ondrechen, M. H. Rubin, and Y. B. Band, *J. Chem. Phys.*, **78**, 4721 (1983).

(19) M. J. Ondrechen, R. S. Berry, and B. Andresen, *J. Chem. Phys.*, **72**, 5118 (1980); M. J. Ondrechen, B. Andresen, and R. S. Berry, *J. Chem. Phys.*, **73**, 5838 (1980).

Some work²² has been spent in the past decade pinpointing the optimal position of such extra boilers and condensers for complete separation of very similar components. However, only very recently has the more realistic case of dissimilar components and only partial separation been treated,²³ with the goal of minimizing the entropy production. It is found that the possible saving of entropy production per unit of feed by addition of intermediate heat exchangers is an almost constant function of feed composition roughly in the interval $0.25 < x_F < 0.75$. A very interesting result is that this function has a (gentle) minimum at $x_F = 0.5$ for $x_D > 0.89$ which turns into a (gentle) maximum for $x_D < 0.89$.

Related to our own work on competing processes⁵ is that of Robert T. Ross and co-workers²⁴ on the efficiency of solar energy converters. They analyze a model in which solar energy is absorbed by a quantum system under less than reversible conditions and where some of the absorbed energy is lost by retransmission and by nonradiative recombination, much as in photosynthesis. The analysis yields the optimal spectral absorbance of the device for a range of decay rates and the associated conversion efficiency. The most spectacular result is that the efficiency can be made to exceed the ideal thermal efficiency by rapid equilibration among the electronically excited states which reduces reradiation and actually leaves the chemical potential of the excited electronic band below that of the ground band. A similar result for a lasing system has been discussed by Ben-Shaul and Levine.²⁵

The effects of an oscillatory reaction path have been studied by John Ross and co-workers, both for heat engines^{16,26} and for chemical systems.²⁷ In both cases resonances are found where either dissipation is a minimum or power production is a maximum, frequently even exceeding the performance of the optimal monotonic reaction path. This means that in some cases batch operation of a process may prove advantageous compared with steady-state operation.^{27,28}

Potential Structures

On a more abstract level the availability (or exergy) of any chemical system has been generalized for finite-time processes.⁵ First let us recall that the availability, A , of a system is the maximum amount of work that can be extracted reversibly while it comes to equilibrium with its surroundings. With subscript zero denoting environment quantities, the traditional availability is

$$A = U - T_0 S + P_0 V - \sum_i \mu_{0i} N_i \quad (7)$$

The decrease in availability of a system in going from initial state i to final state f is then the reversible

(maximum) work that can be extracted in the process regardless of path

$$W_{\text{rev}} = -\Delta A \quad (8)$$

These two qualities are very desirable to preserve in the finite-time generalization. This is accomplished by optimizing the Tolman-Fine expression for the second law²⁹

$$W_{\text{max}} = \int \vec{F} \cdot d\vec{z} - \int T_0 dS_u = \int dW_{\text{rev}} - \int T_0 dS_u \quad (9)$$

where \vec{F} is the (vector) set of generalized forces, $d\vec{z}$ is the (vector) set of generalized displacements corresponding to the components of \vec{F} , T_0 is the temperature of the surroundings into which all heat eventually goes, and S_u is the total entropy of the system plus surroundings. The integrals are carried out along the path taken by the system. Then for an arbitrary process with no mass transfer

$$W = -\Delta A - T_0 \int_{t_i}^{t_f} \dot{S}_u dt \quad (10)$$

which leads us to define the finite-time availability \mathcal{A} of a system undergoing any of a certain class of processes Σ_0 in time τ to be the maximum work extractable by any process within the class Σ_0 :

$$\mathcal{A} \equiv W_{\text{max}} = \max \left[\int \vec{F} \cdot d\vec{z} - \int T_0 dS_t \right] \quad (11)$$

where the maximum is understood to be taken subject to the constraints of the problem, including time or rate constraints. We thus define the generalized (finite-time) availability \mathcal{A} by changing eq 9 from a statement into a recipe for computing that generalized availability from the dynamical equations for work, the constraints, and the equations governing the entropy production of the system.

Another finite-time extension of thermodynamics with immediate chemical applications is that of thermodynamic potentials.⁴ We are used to calculating the reversible work output of a process from accompanying changes in, e.g., the Gibbs or Helmholtz free energies for constant T , P and constant T , V processes, respectively. These potentials are obtained from the incomplete differential $dW = P dV$ by adding integrating terms $g dy$, e.g., $V dP$ for constant pressure and $S dT$ for constant temperature. However, this so-called Legendre transform is valid for *any* constraint $dy = 0$, including constraints on time, complicated expressions being constant, or differential expressions vanishing, e.g., arising from equations of motion expressed as differential equations. This complete generality opens up the possibility of including the dynamics and nontrivial constraints of a chemical process in its generalized thermodynamic potential and thus being able to calculate its energetics without a complete kinetic analysis.

Abstract geometry can be used to set a lower limit on the amount of availability which must be lost in any process leading from a specified initial state i to a final equilibrium state f via states of local thermodynamic equilibrium.³⁰ (This excludes turbulent processes during which the local thermodynamic variables are not

(22) Zs. Fonyo and P. Foldes, *Acta Chim. Acad. Sci. Hung.*, **81**, 103 (1974); F. Kayihan, *AIChE Symp. Ser.*, **76**, 1 (1980).

(23) O. C. Mullins and R. S. Berry, *J. Phys. Chem.*, **88**, 723 (1984).

(24) R. T. Ross and J. M. Collins, *J. Appl. Phys.*, **51**, 4504 (1980); R.

T. Ross and A. J. Nozik, *J. Appl. Phys.*, **53**, 3813 (1982).

(25) A. Ben-Shaul and R. D. Levine, *J. Non-Equil. Thermodyn.*, **4**, 363 (1979).

(26) P. H. Richter and J. Ross, *J. Chem. Phys.*, **69**, 5521 (1978).

(27) Y. Termonia and J. Ross, *J. Chem. Phys.*, **74**, 2339 (1981); *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 3563 (1981); P. H. Richter, P. Rehmus, and J. Ross, *Prog. Theor. Phys.*, **66**, 385 (1981).

(28) J. S. Su and A. J. Engel, *AIChE Symp. Ser.*, **76** (192), 6 (1980).

(29) R. C. Tolman and P. C. Fine, *Rev. Mod. Phys.*, **20**, 51 (1948).

(30) P. Salamon and R. S. Berry, *Phys. Rev. Lett.*, **51**, 1127 (1983).

well-defined.) The second derivatives $\partial^2 U / (\partial X_i \partial X_j)$ of the internal energy U as a function of all the extensive variables X_i have been shown by Weinhold³¹ to have the properties of a metric, which can be used to construct a distance $D(f, i)$ (shortest length) between states i and f .³² The dissipated availability is then

$$\Delta A \geq D^2(\epsilon/\tau) \quad (12)$$

where ϵ is a mean relaxation time of the system, and τ is the duration of the process. For endoreversible systems the bound is strengthened to

$$\Delta A \geq L^2(\epsilon/\tau) \quad (13)$$

where L is the length of the actual path traversed from i to f . These expressions give a direct measure of the cost of finite process time, in terms of internal relaxation times and equilibrium properties.

Concluding Remarks

In this *Account* we have pointed out the inadequacy of reversible thermodynamics to describe processes which proceed at nonvanishing rates and have pointed

(31) F. Weinhold, *J. Chem. Phys.*, **63**, 2479 (1975).

(32) P. Salamon, B. Andresen, P. D. Gait, and R. S. Berry, *J. Chem. Phys.*, **73**, 1001, 5407E (1980).

out a number of new methods to treat this situation. The primary goal has been to obtain bounds of performance which are more realistic than the reversible ones. Some of the finite-time procedures are generalizations of traditional quantities, like potentials and availability; others are entirely new, like the thermodynamic length.

Since the central ideas of reversible thermodynamics are retained in finite-time thermodynamics, we are continuing our attempts to generalize traditional concepts to include time. Especially important are connections to statistical mechanics and irreversible thermodynamics, e.g., investigating the finite-time content of Keizer's Σ -function³³ and processes far from equilibrium. The most exciting results emerging at the moment are applications of the thermodynamic length which seems to be able to simplify calculations on such diverse systems as lasers, separation by diffusion, and signal encoding.

The research on this topic carried out by the authors at the University of Chicago was supported initially by the National Science Foundation and then by the Department of Energy and by the Exxon Educational Foundation.

(33) J. Keizer, *Acc. Chem. Res.*, **12**, 243 (1979).

Electromagnetic Model for Surface-Enhanced Raman Scattering (SERS) on Metal Colloids

MILTON KERKER

Department of Chemistry, Clarkson University, Potsdam, New York 13676

Received January 31, 1984 (Revised Manuscript Received May 21, 1984)

The realization that Raman signals from pyridine adsorbed at *roughened* silver electrodes¹ are enhanced 10^5 – 10^6 over signals from dissolved pyridine^{2,3} unleashed a torrent of investigations that has hardly abated.^{4,5} Suitable substrates for adsorbed molecular species, in addition to the roughened electrodes on which this remarkable phenomenon was first observed, include colloidal metal particles, vacuum-deposited metal island films, matrix-isolated metal clusters, roughened surfaces of single crystals under ultrahigh vacuum, tunnel junction structures, smooth metal surfaces in the attenuated total reflection arrangement, metal-capped polymer posts, and holographic gratings. The initial observations of SERS for adsorbates on gold, silver, and copper have been extended to include aluminum, cadmium, lithium, nickel, palladium, platinum, and sodium.

Speculations to account for this surface-enhanced Raman scattering (SERS) immediately fell into two

major categories. A purely physical mechanism was proposed in which the molecules were presumed to respond to gigantic electromagnetic fields generated locally by collective oscillations of the free electrons in small metal structures. In addition, so-called "chemical" mechanisms envisioned charge transfer between metal and adsorbate or else formation of a molecule-metal atom complex with consequent molecular resonances. Such specific molecular interactions may certainly play a role since different molecules on the same surface or different Raman bands of the same molecule may exhibit different SERS effects. Indeed, contributions to SERS from each of these two kinds of mechanisms, electromagnetic and chemical, are not mutually exclusive, yet the predominant current view is that the major contribution is electromagnetic and is due to the local field enhancement associated with resonant excitation

(1) M. J. Fleischmann, P. J. Hendra, and A. J. McQuillan, *Chem. Phys. Lett.*, **26**, 163 (1974).

(2) D. J. Jeanmaire and R. P. Van Duyne, *J. Electroanal. Chem.*, **84**, 1 (1977).

(3) M. G. Albrecht and J. A. Creighton, *J. Am. Chem. Soc.*, **99**, 5215 (1977).

(4) R. K. Chang and T. E. Furtak, "Surface Enhanced Raman Scattering", Plenum Press, New York, 1982.

(5) A. Otto, in "Light Scattering in Solids", M. Carbone and G. Guntherodt, Eds., Springer, New York, 1983, Vol. IV.

Milton Kerker is Thomas S. Clarkson Professor and Dean of Science at Clarkson University. Born in Utica, NY, in 1920, he received the AB, MA, and Ph.D. degrees from Columbia University after which he joined Clarkson's Chemistry Department in 1949. He has served since 1965 as co-editor of the *Journal of Colloid and Interface Science*. In addition to light scattering, a subject on which he has written a treatise, he is interested in the mechanics of aerosols and the history of science.