

Thomson's formulation of the second law for macroscopic and finite work sources

Armen E. Allahverdyan^{1,2,a)}, Roger Balian^{3,b)} and Theo M. Nieuwenhuizen^{1,c)}

¹⁾ Institute for Theoretical Physics, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

²⁾ Yerevan Physics Institute, Alikhanian Brothers St. 2, Yerevan 375036, Armenia

³⁾ SPhT, CEA-Saclay, 91191 Gif-sur-Yvette cedex, France

^{a)} armena@science.uva.nl

^{b)} balian@spht.saclay cea.fr

^{c)} nieuwenh@science.uva.nl

Received 8 August 2003, Accepted 30 December 2003, Published 18 March 2004

Abstract: Thomson's formulation of the second law states: no work can be extracted from an equilibrium system through a cyclic process. A simple, general proof is presented for the case of macroscopic sources of work. Next the setup is generalized towards situations, where the corresponding work-source is not macroscopic. It is shown that using such a source one can extract energy from an equilibrium system by means of a cyclic process. However, this extraction is accompanied by an entropy increase of the source, in a manner resembling the Clausius inequality.

Keywords: second law; work; entropy.

1 Introduction

The laws of thermodynamics were formulated as basic principles, deduced from studies by Carnot, Joule, Mayer, Clausius and Kelvin. Since one century they have been shown to emerge as consequences of the laws of matter at the microscopic level. Their derivation from first principles requires an identification of concepts such as temperature, work or heat in terms of the elementary constituents, and an elimination of the irrelevant (unobservable) microscopic variables by methods of statistical mechanics. This program has been partially achieved for macroscopic systems in the thermodynamic limit [1, 2]. In contrast, within the domain of finite quantum systems,

the application and even the formulation of the laws of thermodynamics are still the subject of studies [3, 7].

Here we focus on Thomson's (later Lord Kelvin) formulation of the second law: *no work can be extracted from an equilibrium system by means a cyclic process generated by an external work-source*. We have two basic goals. The first is to discuss a rigorous version of this formulation and to present a simple proof of it. This approach uses a minimal number of assumptions and has a larger region of applicability than text-book presentations of Thomson's law, see e.g. [1, 2], where it is typically derived from the law of entropy increase. In contrast, the presented derivation does not invoke the concept of entropy at all and applies equally well to macroscopic and finite equilibrium systems. The crucial fact used in the proof is that the external work-source is macroscopic (in the sense specified below). Our second goal is to determine limits of applicability of Thomson's formulation for sources that are large but not explicitly macroscopic (may be large but finite). It appears that this problem leads to an intimate connection between the formulations of the second law by Thomson and Clausius. It is useful to recall in this context that out of the standard thermodynamical domain — that is, without the assumption that all the involved subsystems are macroscopic — the various formulations of the second law should be studied separately, as there is no *a priori* reason to expect that they are equivalent [7].

The special importance of Thomson's formulation comes from the fact that other formulations of the second law rely more explicitly upon notions such as entropy or temperature which require rather elaborate identification in terms of microscopic variables, whereas this formulation concentrates on the concept of work (energy exchange) which can be considered as a primary, microscopic concept, as we show below.

The present paper is organized as follows. In section 2 we recall the general setup for Thomson's formulation. It is proven in section 3. In section 4 we discuss finite sources of work. Our conclusions are presented in the last section.

2 General setup for Thomson's formulation

In its original form Thomson's formulation states that if a thermodynamic system S , which interacts with a single macroscopic thermal bath B at uniform temperature T and with a source of work W , evolves in a closed cycle, it cannot yield work to that external system W . In other words, such a process cannot extract energy from the heat bath.

We assume that interaction of B with S is sufficiently weak, yielding a negligible contribution to the energy. Over sufficiently long times, energy transfer identified with heat may result from this interaction. The bath can therefore set up a canonical equilibrium at the temperature T for S , without interacting directly with the source of work.

In order to separate from each other the effects of the interactions of S with B and with W , we shall consider in the following a cyclic process that takes place in two steps. A canonical thermal equilibrium of S is assumed to have already been established for times $t < 0$. The interaction process that couples S and W takes place between the times $t = 0$ and $t = \tau$. The delay τ is supposed to be sufficiently short so that no heat is exchanged with the bath during this first step

of the process, which is possible because of the assumed weak coupling between the system S and the bath B. During the second step, for $t > \tau$, the energy of W is left unchanged, while S returns to its original thermal equilibrium state at temperature T through interaction with the bath B. According to Thomson, if the two-step cycle is closed in regard to the system S, a positive amount of energy should have been transferred from W to S between the times $t = 0$ and $t = \tau$, and this amount should be released from S to B for $t > \tau$. This two-step analysis of the process will allow us to focus on the coupled evolution of S and W in the time-interval $0 < t < \tau$, while the subsequent thermal exchange with the bath proceeds in an obvious manner.

More precisely, in order to define the *work* done by the source W, we express the Hamiltonian which generates the evolution in this time-interval as

$$\hat{H} = \hat{H}_S + \hat{H}_I + \hat{H}_W, \quad (1)$$

where $\hat{H}_S$ ($\hat{H}_W$) depends only on the variables of S (W), with $[\hat{H}_S, \hat{H}_W] = 0$, and where $\hat{H}_I$ describes the coupling between S and W. To fix ideas we shall first take $\hat{H}_I$ in the form: $\hat{H}_I = \hat{X}\hat{\xi}$. The observable $\hat{X}$ belongs to the system S, and corresponds to a “force” variable such as the pressure for a gas or the magnetic moment for a magnetic material. The conjugate variable $\hat{\xi}$ refers to the system W, and corresponds to a “position” variable such as the volume for a gas or the magnetic field for a magnetic source.

3 Thomson’s formulation for macroscopic sources

In the thermodynamic limit for the system W (but not necessarily for the system S), the fluctuations in $\hat{\xi}$ are negligible at all relevant times; in relative value they are small as the inverse square root of the number of microscopic degrees of freedom of W contributing to $\hat{\xi}$. The operator $\hat{\xi}$ can therefore be replaced by its expectation value

$$\langle \hat{\xi} \rangle_t = \text{tr} [\hat{\xi} \rho_W(t)] \equiv \xi(t), \quad \rho_W(t) = \text{tr}_S \rho(t), \quad (2)$$

where $\rho(t)$ is the total density matrix of S and W, and where $\rho_W(t) = \text{tr}_S \rho(t)$ and $\rho_S(t) = \text{tr}_W \rho(t)$ are the marginal density operators of W and S, respectively. The evolution of $\xi(t)$ is generated by the Ehrenfest equation:

$$\hbar d\xi(t)/dt = i \langle [\hat{H}, \hat{\xi}] \rangle_t. \quad (3)$$

More precisely, if the fluctuation $\langle \Delta \hat{\xi}^2 \rangle_t$, where $\Delta \hat{\xi} = \hat{\xi} - \langle \hat{\xi} \rangle_t$, is negligible *at all times*, the Schwartz inequality $|\langle \Delta \hat{\xi} \Delta \hat{Z} \rangle_t|^2 \leq \langle \Delta \hat{\xi}^2 \rangle_t \langle \Delta \hat{Z}^2 \rangle_t$, where $\hat{Z}$ is any operator of S, implies that we can make the replacement of the operator $\hat{\xi}$ by its expectation value $\langle \hat{\xi} \rangle_t = \xi(t)$ in its Hamiltonian, which thus takes the form

$$\hat{H} = \bar{H}_S(t) + \hat{H}_W = \hat{H}_S + \hat{X} \xi(t) + \hat{H}_W, \quad (4)$$

where an effective time-dependent Hamiltonian $\bar{H}_S(t) = \hat{H}_S + \hat{X} \xi(t)$ appears for the system S. This form of $\hat{H}$ implies the factorization of the evolution operator and of the density operator

$\rho(t) = \rho_S(t) \otimes \rho_W(t)$ at all times, provided the initial state at $t = 0$ is also factorized. Hence in the Schrödinger picture the density matrix $\rho_S(t) = \bar{U}(t)\rho_S(0)\bar{U}(t)^\dagger$ of the system S evolves according to the unitary transformation $\bar{U} = \mathcal{T} \exp(-i \int_0^t dt' \bar{H}_S(t')/\hbar)$ generated by $\bar{H}_S(t)$.

The equilibrium condition at time $t = 0$ provides the initial density operator of S as

$$\rho_S(0) = e^{-\bar{H}_S(0)/T}/Z, \quad Z = \text{tr} e^{-\bar{H}_S(0)/T}. \quad (5)$$

A cyclic transformation is characterized here by the fact that $\xi(t)$ has returned to its initial value at time $t = \tau$: $\xi(\tau) = \xi(0)$, $\bar{H}_S(\tau) = \bar{H}_S(0)$. Due to the lack of fluctuations in $\hat{\xi}$, this condition is sufficient to ensure that the overall process can be rigorously cyclic at the microscopic level. Indeed, since the state of S is not correlated at the time $t = \tau$ with that of W, it can return later on to its original form (5) at $t = 0$ through interaction with the thermal bath B.

The work given by the macroscopic source W to the system S between the times 0 and τ is identified as the energy yielded by the source W, or equivalently as [2]:

$$W = \text{tr} [\rho_W(0) H_W - \rho_W(\tau) H_W]. \quad (6)$$

Due to (4) this can be also written as $W = \text{tr} [\rho_S(\tau) \bar{H}_S(0) - \rho_S(0) \bar{H}_S(0)]$. With help of (5) one gets, using the fact that $\bar{U}$ is unitary:

$$\frac{W}{T} = \text{tr} [\rho_S(0) \ln \rho_S(0) - \rho_S(\tau) \ln \rho_S(0)] = \text{tr} [\rho_S(\tau) \ln \rho_S(\tau) - \rho_S(\tau) \ln \rho_S(0)] \geq 0. \quad (7)$$

The last expression is nothing but the *relative entropy* of $\rho_S(\tau)$ with respect to $\rho_S(0)$; it is strictly positive if $\rho_S(\tau) \neq \rho_S(0)$ [4].

As shown by (7), the work is positive for a macroscopic source of work W, whether the system S is macroscopic or microscopic. Thus, in the second case this entails an extension of the original Thomson's statement [5, 6], which was verified in specific examples [7]. The irreversibility of the process arises from the lack of symmetry between the times 0 and τ which is introduced by the initial condition (5). Usually, the breaking of time-reversal-invariance is associated with a dissipative process. Here the latter process is the interaction with the heat bath, which will take place after the time τ . In our setup there is no work is done or extracted during this relaxation.

The above treatment can readily be generalized to an interaction with several (not necessarily independent) sources of work, where $\hat{H}_I = \sum_{i=1}^r \hat{X}_i \hat{\xi}_i$. Altogether, if the source W of work is a macroscopic system, the lack of fluctuations in the variables $\hat{\xi}_i$, which couple W to the system S, ensures the second law in Thomson's form provided that S is initially in a canonical equilibrium state.

For non-equilibrium initial states counter examples of the Thomson formulation have been analyzed [7].

4 Thomson's formulation for finite sources.

Let us now turn to the situation where W is no longer a macroscopic object. This is of interest at least for the following two reasons: 1) Real work-source may deviate from the explicit macroscopic limit. One wishes to know to what qualitative implications such deviations may lead. 2)

Non-macroscopic work-sources realize one further step away from the standard thermodynamical domain. It is of relevance to understand what will be now the fate of the second law. In studying these issues we have to deal with the operator nature of the $\hat{\xi}_i$'s and we can no longer leave aside the Hamiltonian $\hat{H}_W$ of the source of work, when studying the dynamics of S.

The first point to note is that work is typically defined for the macroscopic sources only. This corresponds to its intuitive status as “usable energy”. Therefore, for finite sources we shall speak on the decrease

$$\mathcal{W} = \text{tr} [\rho_W(0) H_W - \rho_W(\tau) H_W]. \quad (8)$$

of the *average* energy of the source. In the macroscopic limit for the work-source, this quantity converges to the work as defined by (6). Note that though Eqs. (6, 8) are formally identical, their contexts are different, since these equations refer to macroscopic and finite-size situations, respectively.

We will again assume that in its initial state the system S is in canonical thermal equilibrium, and is not correlated with W:

$$\rho(0) = \rho_S(0) \otimes \rho_W(0), \quad (9)$$

where $\rho_S(0)$ is given by (5) with

$$\bar{H}_S(0) = \hat{H}_S + \sum_{i=1}^r \hat{X}_i \langle \hat{\xi}_i \rangle_0. \quad (10)$$

Initial condition (9) implies that S and W did not interact for $t < 0$, and that S was prepared according to an independent thermalization process as described above. However, the definition of a cyclic process is no longer unambiguous. This condition may indeed be generalized in two different ways, both of which are rather natural [8]. To keep our presentation concise we will here outline only one of the two definitions proposed in [8]. Since the variables $\hat{X}_i$ and $\hat{\xi}_i$ can now be correlated, and since the interaction term $\hat{H}_I$ involves a sum of their products, we search a condition for such a sum of products. We notice that in the considered finite-size case the states of S and W do become correlated for $t > 0$. Since now both $\rho(t)$ and $\hat{H}_I$ irreducibly live in the common Hilbert space of S and W, the most relevant way to define energy exchange between two systems is to require that the average interaction Hamiltonian does not contribute to the total energy budget, i.e., it is zero both initially and finally. For our situation the proper interaction Hamiltonian is $\sum_{i=1}^r \hat{X}_i (\hat{\xi}_i - \langle \hat{\xi}_i \rangle_0)$, since its average is zero at the initial time $t = 0$. Requiring that this average is also equal to zero at some final time τ , we get for a cyclic process with the duration τ :

$$\sum_{i=1}^r \langle \hat{X}_i \hat{\xi}_i \rangle_\tau = \sum_{i=1}^r \langle \hat{X}_i \rangle_\tau \langle \hat{\xi}_i \rangle_0. \quad (11)$$

For a large system W we recover the macroscopic condition for $r = 1$, since the lack of fluctuations in $\hat{\xi}$ implies the lack of correlations with $\hat{X}$ at any time: $\langle \hat{X} \hat{\xi} \rangle_t \simeq \langle \hat{X} \rangle_t \langle \hat{\xi} \rangle_t$, while for a cycle $\langle \hat{\xi} \rangle_\tau = \langle \hat{\xi} \rangle_0$.

5 The short-time limit.

We shall discuss $\mathcal{W}$ for a finite source realizing a cyclic process of a small duration τ , so that we can expand $\mathcal{W}$ in powers of τ . This limit is especially interesting, since non-macroscopic contributions can overcome the main macroscopic term: as we show [8], the macroscopic contribution to $\mathcal{W}$ evaluated by neglecting the fluctuations of $\hat{\xi}_i$'s and given by (7), scales as τ^6 for small τ , whereas in the general case it is much larger, $\mathcal{W} \approx C\tau^2$. In this approach the possibility to extract energy from an equilibrium system shows up as the case $C < 0$.

Expanding the Heisenberg equations of motion one gets for the work [8]:

$$\frac{2\hbar^2 \mathcal{W}}{\tau^2} = \sum_{i,k=1}^r \langle \{ \Delta \hat{\xi}_i, \Delta \hat{\xi}_k \}_+ \rangle \sum_{\alpha < \beta} (p_\alpha - p_\beta)(h_\beta - h_\alpha) \Re \{ \langle \alpha | \hat{X}_i | \beta \rangle \langle \beta | \hat{X}_k | \alpha \rangle \} \quad (12)$$

$$+ i \sum_{i,k=1}^r \langle [\hat{\xi}_k, \hat{\xi}_i] \rangle \sum_{\alpha < \beta} (p_\alpha + p_\beta)(h_\beta - h_\alpha) \Im \{ \langle \alpha | \hat{X}_k | \beta \rangle \langle \beta | \hat{X}_i | \alpha \rangle \}, \quad (13)$$

where $\{\dots, \dots\}_+$ denotes anti-commutator, $\Re$ and $\Im$ stand for the real and imaginary parts, respectively. $\{|\alpha\rangle\}$ is the common eigenbase of $\rho_S(0)$ and $\bar{H}_S(0)$, and p_α , h_α are the corresponding eigenvalues: $p_\alpha = e^{-\beta h_\alpha} / \sum_\alpha e^{-\beta h_\alpha}$. For $r = 1$, where only the term (12) survives, one can show that $\mathcal{W} > 0$ because $\hat{\rho}_S(0)$ is a decreasing function of $\hat{H}_S$. Since the thermodynamic contribution to the work is negligible small, the amount of dissipated energy predicted by the second law, is *enhanced* for finite sources. The contribution (13) reflects an effect of interference between different types of work (channels), since it contains non-diagonal terms over i, k and becomes zero with $\langle [\hat{\xi}_k, \hat{\xi}_i] \rangle$. Now assume that the temperature of the system S is so high (fluctuations are strong) that $p_\alpha \simeq p_\beta$ for all the β and α that are connected by matrix elements of $\hat{X}_k$ and $\hat{X}_i$. Then the contribution of (12) can be disregarded, and provided that $\langle [\hat{\xi}_k, \hat{\xi}_i] \rangle \neq 0$ and $\Im \{ \langle \alpha | \hat{X}_k | \beta \rangle \langle \beta | \hat{X}_i | \alpha \rangle \} \neq 0$, the sign of (13) can be made negative. Thus, one *can* extract energy from an equilibrium system by means of a cyclic process realized by a *finite* (possibly large but not macroscopic) source of work.

The above results also apply to the classical limit, where Poisson brackets take the place of commutators.

Let us now shortly consider the thermalization process that takes place for $t > \tau$. It is straightforward to show that due to the above definition of cyclic processes, the marginal density operator of S is the equilibrium distribution (5), both at the preparation time $t = 0$ and at times $t \gg \tau$, after the interaction with the thermal bath has taken place.

As we already stressed the following important fact distinguishes between macroscopic and finite work sources: in the former case the states of S and W do not correlate, i.e. $\rho(t) = \rho_S(t) \otimes \rho_W(t)$ for all relevant times t , and then the evolutions of both S and W alone is unitary. Due to this, the von Neumann entropies of S and W are conserved separately. In general these do not happen for finite sources, and only the overall entropy of S+W is conserved between the times 0 and τ . Denote by $\hat{U}$ the overall unitary evolution generated by the Hamiltonian (1): $\hat{U} = \exp(-\frac{t}{\hbar} H)$. Let $\rho(\tau) = \hat{U} \rho(0) \hat{U}^\dagger$ be the overall density matrix of S and W at the moment τ where a cyclic

process is finished in the sense described above. Consider the relative entropy

$$S[\rho(\tau) || \rho_S(0) \otimes \rho_W(\tau)] \equiv \text{tr} [\rho(\tau) \ln \rho(\tau) - \rho(\tau) \ln \{ \rho_S(0) \otimes \rho_W(\tau) \}] \geq 0, \quad (14)$$

between the final state $\rho(\tau)$ and an auxiliary state $\rho_S(0) \otimes \rho_S(\tau)$, where $\rho_S(0)$ is given by (9), and where $\rho_W(\tau) = \text{tr}_S \rho(\tau)$ is the marginal density operator of W. One now uses the following facts: *i)* $\text{tr} [\rho(\tau) \ln \rho(\tau)] = \text{tr} [\rho(0) \ln \rho(0)]$ due to the unitarity of $\hat{U}$; *ii)* Eqs. (9, 10) together with the definition (8) of $\mathcal{W}$; *iii)* The cyclicity condition (11). These conditions allow to go from (14) to a relation for the change of energy of W, $\Delta E_W = -\mathcal{W}$,

$$T \Delta S_W \geq \Delta E_W. \quad (15)$$

$\Delta S_W = -\text{tr} [\rho_W(\tau) \ln \rho_W(\tau)] + \text{tr} [\rho_W(0) \ln \rho_W(0)]$ is the entropy increase of W. Recall that T is the temperature of the bath, to which S is coupled.

A macroscopic source exchanges energy without changing its the entropy, so (15) again proves the Thomson formulation $\mathcal{W} \geq 0$.

For a finite source W we may split up ΔE_W in heat received by W and work done on W, $\Delta E_W = \Delta Q_W + \Delta W_W$. When $\Delta E_W > 0$ we expect both terms to be positive, in which case Eq. (15) implies the Clausius inequality $T \Delta S_W \geq \Delta Q_W$. In other words, if for finite sources Thomson's formulation of the second law can be violated ($\mathcal{W} < 0$, $\Delta E_W > 0$), this violation must be accompanied by an entropy increase of the work-source, that typically exceeds the increase predicted by Clausius. It is useful to stress that this conclusion is general and does not depend on details of W and its interaction with S.

In conclusion, we have seen that different formulations of the second law have different limits of applicability for finite systems. In this respect we may recall that another formulation, positivity of the rate of energy dispersion, may be violated using a macroscopic work source, even in processes starting from an equilibrium state [7].

References

- [1] L.D. Landau and E.M. Lifshitz, *Statistical Physics, I*, (Pergamon Press Oxford, 1978).
- [2] R. Balian, *From Microphysics to Macrophysics*, volumes I and II, (Springer, 1992).
- [3] *Quantum Limits to the Second Law*, ed. D.P. Sheehan, (AIP Conf. Proc. 643, 2002)
- [4] F. Schlögl, Phys. Rep. **62**, 268 (1980).
- [5] A. Lenard, J. Stat. Phys **19**, 575, (1978). W. Pusz, L. Woronowicz, Comm. Math. Phys., **58**, 273, (1978). G.N. Bochkov and Yu.E. Kuzovlev, Physica A, **106**, (1981) 443. J. Kurchan, cond-mat/0007360.
- [6] A.E. Allahverdyan and Th.M. Nieuwenhuizen, Physica A **305**, 542 (2002).
- [7] A.E. Allahverdyan and Th.M. Nieuwenhuizen, Phys. Rev. Lett **85**, 1799 (2000); J. Phys. A, **36**, 875 (2003); Phys. Rev. E **64** 056117 (2001). Th.M. Nieuwenhuizen and A.E. Allahverdyan, Phys. Rev. E **66**, 036102 (2002)
- [8] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen, unpublished.