European Journal of
Physics

Reminder of Jaynes' views on two key issues of Gibbsian statistical mechanics

To cite this article: Enda McGlynn 2025 *Eur. J. Phys.* **46** 065103

OPEN ACCESS
IOP Publishing European Journal of Physics
Eur. J. Phys. 46 (2025) 065103 (17pp) https://doi.org/10.1088/1361-6404/ae0dca

Reminder of Jaynes' views on two key issues of Gibbsian statistical mechanics

Enda McGlynn

School of Physical Sciences, Dublin City University, Glasnevin, Dublin 9, Ireland

E-mail: enda.mcglynn@dcu.i.e

Received 1 August 2025, revised 15 September 2025
Accepted for publication 30 September 2025
Published 31 October 2025

Abstract

Gibbs' approach to statistical mechanics incorporates key foundational concepts and is at the core of modern statistical thermodynamics. However, two aspects of Gibbs' approach still cause problems for students and teachers alike. The first is the conservation of phase space density in a thermally isolated system (the Liouville theorem) which means that the entropy of the system, according to Gibbs' definition, is necessarily constant. This then leads to measures such as coarse-graining to enable an approach to statistical equilibrium. The second aspect is the Gibbs' paradox, concerning the extensivity of the entropy of ideal gas mixtures in classical statistical mechanics. However, Edwin Jaynes addressed both of these issues in a logically consistent manner many decades ago, and provided interesting and insightful alternative ways to view them. Surprisingly Jaynes' treatment of both these issues seems to be largely forgotten. In the present work Jaynes' treatment is recalled and linked back to Gibbs' original work. Several important issues merit being highlighted once again for a new generation of teachers and students, including the details of Gibbs' statistical mechanical definition of entropy, the exact definition of entropy in classical thermodynamics, as well as Jaynes' insights on the distinction between physical and thermodynamical systems, the relationship to the anthropomorphic nature of entropy, and the information known about a system.

Keywords: Gibbs, statistical mechanics, entropy, Jaynes, Gibbs' paradox, Liouville's theorem

Introduction

Statistical mechanics developed from the mid 1800s onwards, with Boltzmann and Gibbs generally regarded as the progenitors of key concepts and definitions [1–3]. The approaches of Boltzmann and Gibbs are quite different. Boltzmann focuses on a constant energy system (a micro-canonical ensemble in Gibbs' terms) and considers the number of microstates of a configuration as the fundamental aspect. The equilibrium state is identified with the macrostate that maximizes the number of microstates. In this regard one can then speak about the approach to statistical equilibrium as the system inexorably moving towards this macrostate. In Boltzmann's definition, the entropy is then proportional to the natural logarithm of the number of microstates [4–6]. Gibbs by contrast uses the concept of ensembles, collections of independent replicas of a system. Averages over such ensembles allow relevant macroscopic properties of a system to be extracted [3, 7]. In particular Gibbs uses certain ensembles with defined coefficients of probability in phase space as a function of energy to define the constant energy isolated system (micro-canonical ensemble) and a system with a defined equilibrium temperature (canonical ensemble). Within the context of these ensembles however Gibbs does not single out a particular subgroup of the ensemble as constituting the equilibrium configuration but rather takes the entirety of the ensemble as constituting the equilibrium state. He defines the entropy, S, as proportional to the ensemble average of the natural logarithm of the occupation probability density, P, with the constant of proportionality equal to Boltzmann's constant, k_{B} [8]. This yields the Gibbs entropy:

$$S = -k_{\mathrm{B}} \int \dots \int P . \ln(P) \mathrm{d}p \mathrm{d}q = -k_{\mathrm{B}} \overline{\ln(P)}. \tag{1}$$

Despite the differences in approaches, and rather lively discussion and debate through the years, in the thermodynamic limit and with reasonable assumptions regarding ergodicity in fact both yield the same results for observable properties of system and the reasons for their agreement are quite readily understood [9]. While some fine details still generate disagreement, at the undergraduate level the approaches are entirely consistent.

However, Gibbs' approach has increasingly come to be viewed as a more elegant and logically consistent one, whereby the various ensembles used (micro-canonical, canonical and grand-canonical) allow a wider range of phenomena to be treated within a unified framework [2]. Despite this, two aspects of Gibbs' approach still cause problems for students and teachers alike. Firstly, in an ensemble of systems undergoing evolution according to a Hamiltonian function, the Liouville theorem states that the local density of the ensemble probability distribution around any moving point in the ensemble cannot change. So, while the shape of the distribution can change significantly, the average value of the logarithm of the probability density is a rigorously conserved quantity and hence the Gibbs entropy cannot change. This is clearly at odds with the phenomenon of e.g. an irreversible free adiabatic expansion of a gas, where the classical thermodynamic entropy increases. Gibbs realized the conceptual difficulties that this caused and devoted a large part of chapter XII of his work 'Elementary Principles in Statistical Mechanics' [7] (referred to as EPSM from here onwards) to attempts to explain it and come up with an acceptable conceptual framework. The essence of this was what has come to be known as coarse-graining, whereby phase space is divided into finite sized elements and the microscopic structure of the probability distributions within each element are replaced by their mean values. This provides a work-around solution, but in classical mechanics, where there is no natural minimum-sized unit in phase space (unlike the case for quantum mechanics), it is undeniably an arbitrary process without physical justification. Gibbs also says at the end of chapter XII in EPSM that any indefiniteness in the time sampling of an ensemble will achieve the same effect, but the fundamental problem remains. Even in the earliest reviews of Gibbs' work in EPSM, as well as the reflective commentary on Gibbs's work published in 1936, a number of authors are clearly aware of the unsatisfactory nature of this approach [9–11]. However, the coarse-graining approach continues to appear in many textbooks on the subject [12, 13].

Secondly, Gibbs' approach yields a result for the entropy of ideal gas mixtures which is non-extensive, called the Gibbs paradox, and which is at odds with our physical intuition. Gibbs discussed this result in EPSM and proposes a mathematical solution, which essentially is to divide the result by the factorial of the number of particles, distinguishing between 'generic' and 'specific' phases in this way. And once again this provides a workaround solution, as well as fuel for much subsequent discussion, debate and argument amongst physicists and philosophers [14–17].

The purpose of the present contribution is to highlight the insights of Professor Edwin Jaynes on both these issues, which were published in the open literature between 30 and 60 years ago [18, 19][1]. Jaynes was a towering figure in 20th century physics and information theory, and an expert on Gibbs' work. In these articles Jaynes addresses both the issues of phase space conservation and the Gibbs paradox in an interesting and logically consistent manner which provides an alternative viewpoint, i.e. that arbitrary assumptions and work-around solutions are unnecessary. However, Jaynes' insights seem not to have permeated into the mainstream of working and teaching physicists. Jaynes' logic is clear and persuasive, but can be quite terse and often assumes familiarity of readers with much of Gibbs' work. I try below to present Jaynes' ideas on both these issues and to provide some discussion of the physical underpinnings in the hope that his and Gibbs' works will be re-read by a new generation of physics students and teachers and thereby gain a new following. In particular I hope that this work may be of interest and value to teachers who use Gibb's approach, as well as students in their classes, and even to future textbook writers including this subject within their content.

Jaynes' treatment of the conservation of phase space density and its relation to changing entropy and the second law of thermodynamics. As described above, Gibbs' definition of the entropy of the system as proportional to the ensemble average of the natural logarithm of the occupation probability density means that for an isolated system a clash with the implication of the Liouville theorem is unavoidable when one speaks about a system reaching equilibrium. The latter says that this ensemble average cannot change and hence that the entropy of such a system cannot change, in clear conflict with irreversible processes such as the irreversible free adiabatic expansion of a gas.

Jaynes has treated this in the second half of his 1965 paper entitled 'Gibbs versus Boltzmann Entropies' [19], specifically sections 4 and 5. To an extent he essentially inverts the normal discussion of this topic and shows that the conservation of phase space volume under an adiabatic change is the direct origin of the second law of thermodynamics and the increase of entropy. He does this by defining the 'experimental entropy' (S_{e}) for any probability distribution W_N as the Gibbs entropy (S_{G}) calculated for a canonical distribution with the same average energy as the probability distribution W_N. By well-known arguments about the extremal properties of the canonical distribution shown by Gibbs himself [7] the Gibbs entropy calculated for a canonical distribution is greater than that calculated for any other probability distribution with the same average energy (or indeed the same distribution in energy; theorems I and II of EPSM, respectively). Hence $S_{\mathrm{G}} \leqslant S_{\mathrm{e}}$. For example, if an

---

[1] A PDF copy of [18] can be found at: https://bayes.wustl.edu/etj/articles/gibbs.paradox.pdf. PDF copies of many of Jayne's other papers can be found at: https://bayes.wustl.edu/etj/node1.html.

initial state at $t = 0$ is represented by a canonical distribution with an average energy E and entropy S_G ($=S_{ei}$ for the initial state), and then undergoes an adiabatic change (e.g. free expansion) to a later final state, the value of S_G will not change due to the Liouville theorem, but since the probability distribution in the final state will not be canonical, the new 'experimental entropy' S_{ef} (associated with a new, different canonical distribution in the increased volume) must be greater than or equal to S_{ei}.

Jaynes elaborates on this approach in section 5 of 'Gibbs versus Boltzmann Entropies', which is based on his statement that '…the experimental entropy is a measure of all conceivable ways in which the final macrostate can be realized, and not merely of all ways in which it could be produced in one particular experiment'. He also explicitly dismisses the use of coarse-graining as being essentially artificial and of no physical significance, saying that 'Any really satisfactory demonstration of the second law must therefore be based on a different approach than coarse-graining'.

George Phillies has noted in his textbook that entropy is an ensemble, rather than a mechanical, variable and that 'The value of an ensemble variable can be obtained by examining an entire ensemble, but can not be calculated from a complete microscopic description of a single element of that ensemble.' [20]. This is certainly the case, but Jaynes goes beyond this in his 1965 paper, to essentially say that a thermodynamic state must be associated in statistical mechanics not with any single ensemble, but with the sub-set of all possible ensembles consistent with the system variables of the macrostate, and with the same mean value of ensemble energy. Different ensembles yielding equal and unchanging values for these macroscopic system variables must be considered to represent the same thermodynamic state. From this perspective the association of a macroscopic thermodynamic state with a single ensemble is the entire origin of the clash with the Liouville theorem. Jaynes also says that the only meaningful criterion for determination of whether a system is in equilibrium is by experimentation, i.e. is it stable over sufficient periods of time, with unchanging macroscopic state variables?

It is clear that in order to associate a value for the macroscopic state variable of entropy with an ensemble whose average yields the desired macroscopic state variable behavior one must have a process of some sort. Gibbs, in chapter XII of EPSM, advocates a process of taking the ensemble average of the natural logarithm of the occupation probability density after a prior step which we now call coarse-graining. Essentially, he proposes that the probability distribution of any ensemble will change and spread out to form a permanent (stationary) equilibrium distribution after sufficient time has elapsed. He makes significant efforts to physically justify this view and the coarse-graining approach using a range of arguments, but these arguments (as opposed to the efficacy of the process in terms of production of the correct numerical value for the entropy) are doomed to failure because in classical systems coarse-graining is an arbitrary mechanism without physical justification.

Jaynes advocates for a different process to associate a value for the macroscopic entropy with an ensemble, without such arbitrary mechanisms, which reflects the recognition of the macroscopic thermodynamic state as being represented by a sub-set of ensembles, all with the same average value of the relevant property (the energy), rather than by any single ensemble. He says that one should choose one representative ensemble from the sub-set and base the choice on well-established properties of the chosen ensemble and then use the Gibbs' process of calculation of the ensemble average of the natural logarithm of the representative ensemble probability density. As stated previously, Jaynes' logic is unquestionable, but his presentation is also relatively minimal. In particular, he defines what he calls the 'experimental entropy' for the sub-set of ensembles representing a thermodynamic state by the value calculated for the canonically distributed ensemble within this sub-set. Given the extremal properties of the canonical distribution mentioned above, this distribution is certainly a unique choice and hence mathematically logical. Jaynes justifies the choice more in terms of knowledge or ignorance of the state, which is aligned with his interests in the maximum entropy principle and information theory [21–23]. But for teaching at an undergraduate physics level one can also ask; why is it the most physically plausible and justified choice?

To address this point, we need to look carefully at both the definition and properties of entropy in classical thermodynamics, and also at the arguments by which Gibbs develops what he refers to as 'thermodynamic analogies' in chapters IV and XIII in EPSM, i.e. the motivation for the identification of the ensemble average of the natural logarithm of the occupation probability density with the entropy. Firstly, the definition of entropy as a state function in classical thermodynamics is [24]:

$$S_{\mathrm{B}} - S_{\mathrm{A}} = \int_{\mathrm{A}}^{\mathrm{B}} \frac{\mathrm{d}q}{T} \text{ along a reversible path from A to B.} \tag{2}$$

There are several points worth noting here: (i) that only entropy differences are defined by the second law of thermodynamics (a point also noted by Gibbs [7]), (ii) that this definition is for a closed system (no particles entering or leaving), (iii) that the numerical values of entropy differences are defined only for changes along reversible paths, and (iv) that the entropy of equilibrium states with constant configuration variables (e.g. volume) in classical thermodynamics is a monotonic function of the internal energy U [25]. This last point (iv), in combination with Gibbs' demonstration of the extremal properties of a canonical distribution compared to any other probability distribution with the same average energy, is the physical justification for Jaynes' specification of the sub-set of ensembles he uses, i.e. all those with the same average energy.

In relation to point (iii), and referring to the Clausius inequality, this implies that any heat exchange must be with a reservoir with a well-defined temperature and consequently that at each stage of the reversible path involving heat exchange the system must have a well-defined temperature equal to that of the most recently contacted reservoir. The final heat exchange step along the reversible path (at which point the final entropy is determined) is then always based on interaction with a reservoir with a well-defined temperature. Furthermore, in relation to this, the thermodynamic analogies developed in chapters IV and XIII in EPSM are explicitly based on the assumption that changes in the ensemble bring it from one canonically distributed form to another canonically distributed form and are formulated in the context of infinitesimal changes between such forms. This is the basis of the discussion at the end of chapter IV of EPSM, from equation (EPSM.106) to equation (EPSM.116), and the later discussion in chapter XIV. Nowhere in Gibbs' treatment is there any justification for the identification of the ensemble average of a general (i.e. not canonical) probability distribution, albeit one yielding experimentally identical outcomes, with the thermodynamic entropy. This latter step is the one which leads to the inevitable clash with the Liouville theorem. If we step back from this, the associated problems vanish.

Hence it is clear that the classical thermodynamic definition of entropy and Gibbs' identification of an ensemble average with entropy are centrally rooted in the use of interactions with thermal reservoirs with well-defined temperatures and canonically distributed ensembles, respectively. This, along with the extremal properties of the canonical ensemble, provides the main physical justification for its use as the representative ensemble of a sub-set in Jaynes' definition of the 'experimental entropy'.

How does the physical point concerning interactions with thermal reservoirs with well-defined temperatures link to the mathematical one concerning the use of canonically distributed ensembles? Does the use of a reversible path involving different thermal reservoirs physically link to the use of canonical distributions with different temperatures? Does the classical thermodynamic definition of entropy support the use of a canonically distributed ensemble as the representative ensemble of a sub-set in Jaynes' definition?

Gibbs actually provides all the tools to deal with this issue in EPSM. In chapter XII Gibbs discusses the process of incompressible mixing of two fluids of different colors, as an analogy to the spreading and distortion at fixed local density of the ensemble probability distribution. Most tellingly on page 145 he says 'Yet no fact is more familiar to us than that stirring tends to bring a liquid to a state of uniform mixture, or uniform densities of its components, which is characterized by minimum values of the average squares of these densities. It is quite true that in the physical experiment the result is hastened by the process of diffusion, but the result is evidently not dependent on that process'. But the last part of this comment is simply not the case. True intermixing which will change the local density must involve a process beyond incompressible hydrodynamics, which can never achieve this result. And equivalently, for an ensemble probability density, the phenomenon of spreading out in phase space of a region of the ensemble (and hence reduction of the local probability density) must be related to a non-Hamiltonian evolution to which the Liouville theorem does not apply.

Interaction of systems with thermal reservoirs implies random thermal perturbations not described by the Hamiltonian of the systems, and hence, for the representation of an ensemble of such systems, the Liouville theorem is not obeyed during interaction with a thermal reservoir. Gibbs' method of treating the effects on an ensemble of the interaction of systems with thermal reservoirs is discussed in chapter XIII of EPSM, based on theorem VII of chapter XI. Gibbs treats this interaction by considering an ensemble of systems and an ensemble of reservoirs and then combining the two ensembles, whereby each member of each ensemble is treated as being mechanically coupled to every member of the other, to create a larger ensemble of members with a greater number of degrees of freedom. This larger ensemble is isolated and hence it obeys the Liouville theorem for its probability density (which is a function of the increased number of degrees of freedom). But in the discussion relating to equations (EPSM.455) to (EPSM.463) he clearly shows that the sum of the average values of the probability distributions of the system and reservoir ensembles taken separately must decrease compared to that of the average value of the combined ensemble. And in practice it is only the average properties of the ensembles taken separately which can relate to the experimental observations of the system and/or reservoir. Essentially the correlations between the ensembles of systems and reservoirs are preserved in the higher dimension probability distribution of the larger isolated ensemble which evolves according to the combined Hamiltonian of the system and obeys the Liouville theorem due to these correlations. But the separate probability distributions of the system and reservoir ensembles, found by integrating over (i.e. tracing out) the degrees of freedom of the reservoir and system, respectively, are spread out by these integrations. The net effect is the required diffusive spreading out in phase space of the system ensemble which cannot be achieved in a system governed by evolution solely according to a Hamiltonian but is readily achieved when a system interacts with a reservoir. Thus, the system ensemble can evolve towards a canonical distribution with the required extremal properties. Gibbs' detailed analysis of this key point is reproduced in the appendix (where I use Gibbs' notation to help readers who wish to look at the original work). This provides the linkage between the physical point concerning interactions with thermal reservoirs with well-defined temperatures and the mathematical one concerning the use of canonically distributed ensembles. I now examine the application of this to a specific case of relevance, that of a thermally isolated system which undergoes an irreversible process (with no change in the Gibbs entropy), and which is then put in contact with a thermal reservoir at the same temperature. The analysis clearly shows that the Gibbs entropy of the system must increase, while that of the reservoir is essentially unchanged in the limit of a very large reservoir.

Because the (isolated) combined ensemble evolves according to the Liouville theorem we can assume that $\overline{\eta^0} = \overline{\eta^1}$ in all circumstances, where η is the index of probability, as defined in the appendix ($\eta = \ln(P)$, where P is the probability distribution), and the superscripts '0' and '1' refer to earlier and later times, t_0 and t_1, respectively. We note that Gibbs' analysis in chapter XIII of EPSM includes the additional statement that $\overline{\eta^1} < \overline{\eta^0}$ due to his discussion of the approach of systems towards statistical equilibrium over long time periods, but this condition is not used here and does not affect the present discussion or conclusion in any way. As stated above, we have an initially thermalised and thereafter thermally isolated system which undergoes an irreversible process, e.g. a free adiabatic expansion. The shape of the associated system ensemble probability distribution can change significantly from the initial canonical distribution, but the average value of the logarithm of the probability density is a rigorously conserved quantity and hence the Gibbs entropy cannot change. But that system must have a temperature, and when put in contact with a canonically distributed reservoir of the same temperature, no net heat exchange will occur. Hence the ensemble averages of the system and reservoir energies in this case won't change in such an interaction and neither will the classical thermodynamic entropies of either the system or the reservoir.

But what happens to the ensemble average values of the logarithms of the system and ensemble probabilities when the system is put in contact with this reservoir? There are two points.

Firstly, the average value of the logarithm of the reservoir probability must either remain the same or increase (although we will show below that any increase will be vanishingly small), since a canonical distribution has the minimum value for a particular ensemble average energy, as shown by Gibbs in theorem II, chapter XI of EPSM. Hence the average value of the logarithm of the system probability must decrease in all circumstances, as shown below.

Given that $\overline{\eta^1} - \overline{\eta^1_{\text{system}}} - \overline{\eta^1_{\text{reservoir}}} > 0$ after the interaction between the system and reservoir, and because the combined ensemble evolves according to the Liouville theorem (so $\overline{\eta^0} = \overline{\eta^1}$), we have:

$$\overline{\eta^1} - \overline{\eta^1_{\text{system}}} - \overline{\eta^1_{\text{reservoir}}} > 0 \tag{3}$$

$$\overline{\eta^1_{\text{system}}} < \overline{\eta^1} - \overline{\eta^1_{\text{reservoir}}} \tag{4}$$

$$\overline{\eta^1_{\text{system}}} < \overline{\eta^0} - \overline{\eta^1_{\text{reservoir}}} \tag{5}$$

The reservoir ensemble is assumed to be a thermalised ensemble at the initial time and therefore canonically distributed, so $\overline{\eta^0_{\text{reservoir}}}$ has the smallest possible value consistent with the initial average energy of the reservoir ensemble and $\overline{\eta^1_{\text{reservoir}}} = \overline{\eta^0_{\text{reservoir}}} + \overline{\Delta\eta_{\text{reservoir}}}$, where $\overline{\Delta\eta_{\text{reservoir}}}$ must be greater than zero, since the average energy of the reservoir ensemble has not changed and $\overline{\eta^0_{\text{reservoir}}}$ is the smallest possible value consistent with this average energy of the reservoir ensemble.

Hence:

$$\overline{\eta^1_{\text{system}}} < \overline{\eta^0} - \overline{\eta^0_{\text{reservoir}}} - \overline{\Delta\eta_{\text{reservoir}}} \tag{6}$$

Since, prior to an initial time t_0, the system and reservoir ensembles are not interacting we have that:

$$\overline{\eta^1_{\text{system}}} < \overline{\eta^0_{\text{system}}} - \overline{\Delta\eta_{\text{reservoir}}} \tag{7}$$

Since $\overline{\eta^0_{\text{system}}} - \overline{\Delta\eta_{\text{reservoir}}} < \eta^0_{\text{system}}$ because $\overline{\Delta\eta_{\text{reservoir}}} > 0$, we have:

$$\overline{\eta^1_{\text{system}}} < \overline{\eta^0_{\text{system}}} \tag{8}$$

This shows that the effect on the system of the interaction between system and reservoir is that $\overline{\eta^1_{\text{system}}} < \overline{\eta^0_{\text{system}}}$, i.e. the average value of the system ensemble's probability distribution decreases, which is the required diffusive spreading in phase space mentioned in the main text, which enables an evolution towards a canonical distribution.

Hence contact with a thermal reservoir will cause a diffusive spreading in phase space which must ultimately lead to the extremal canonical distribution.

Secondly, the reservoir ensemble is assumed to be a thermalised ensemble at the initial time and therefore canonically distributed and hence has the extremal properties of the canonical distribution, so $\overline{\eta^0_{\text{reservoir}}}$ is the smallest possible value consistent with the initial average energy of the reservoir ensemble. The reservoir is also assumed to be a system with a far greater number of degrees of freedom than the system and is minimally affected by the interaction with the system (the standard assumption for a reservoir to be much larger than a system and hence minimally affected by interactions with the system).

These two facts imply that the average value of the logarithm of the reservoir probability is largely unaffected by the interaction with the system. We can write for average value of the logarithm of the probability:

$$\overline{\eta} = \int \dots \int \eta e^{\eta} dp dq \tag{9}$$

For the initially canonically distributed reservoir: we have that $\eta = \frac{-E}{k_B T} + \frac{C}{k_B T}$, where C is a constant. Hence, if the change in the index of probability of the reservoir due to the interaction with the system, away from the initially canonical form, is $\Delta\eta$ (which will be a function of q and p), the change in quantity $\overline{\eta}$ for the reservoir is found by differentiating inside the integral with respect to η:

$$\Delta\overline{\eta} = \int \dots \int \Delta\eta e^{\eta} dp dq + \int \dots \int \Delta\eta \eta e^{\eta} dp dq \tag{10}$$

$\Delta\eta$ is going to be a small random deviation due to the interaction of the reservoir with the (much smaller) system. Since the probability P is given by $\exp(\eta)$; $\Delta P = \Delta\eta \exp(\eta)$. Because the integral of ΔP over the entire phase space must vanish (the integral of P over phase space must equal 1), the first integral must vanish identically. The second integral in (10) can be split into two parts if the system is canonically distributed initially:

$$\Delta\overline{\eta} = \int \dots \int \Delta\eta e^{\eta} \left(\frac{-E}{k_B T} \right) dp dq + \int \dots \int \Delta\eta e^{\eta} \left(\frac{C}{k_B T} \right) dp dq \tag{11}$$

The second term in (11) vanishes identically for the same reasons the first term in (10) does, while the first vanishes since the average energy is unchanged in the particular interaction specified here, i.e. that the system and reservoir are at the same temperature and hence there is no net heat exchange between them.

Hence overall $\Delta\overline{\eta} = 0$ for a disturbance of the reservoir away from an initially canonically distributed form[2]. Because the reservoir has a far greater number of degrees of freedom than the system its index of probability is affected very little by the interaction with the system and thus it will remain very close to canonically distributed during the period of interaction (and any variations will be of second order in $\Delta\eta$) and hence $\overline{\eta^0_{\text{reservoir}}} \lesssim \overline{\eta^1_{\text{reservoir}}}$ (i.e. $\overline{\Delta\eta_{\text{reservoir}}} \gtrsim 0$).

This implies that the effect of the interaction of system and reservoir is that $\overline{\eta^1_{\text{system}}} < \overline{\eta^0_{\text{system}}}$, with a minimal effect on the reservoir, i.e. the average value of the system ensemble's probability distribution decreases, which is the required diffusive spreading in phase space mentioned in the main text, and which enables an evolution towards a canonical distribution in the system.

Given that at the macroscopic level, since there has been no energy exchange between system and reservoir and hence the classical thermodynamic entropies of these two objects have not changed, the choice of a canonically distributed ensemble with the appropriate average ensemble energy as the representative ensemble to be used in the definition of the system entropy is supported strongly by this argument.

In summary, a physically plausible and justified choice for the representative ensemble to be used in the definition of the entropy is a canonical distribution with the appropriate average energy and other macroscopic system variables, and the present discussion is intended to remind readers of, and to expand on, Jaynes' discussion and to provide additional physical motivating arguments, some of which are drawn from Gibbs' own work. We also note quite recent work by Tomé & de Oliveira at a mathematically more advanced level incorporating stochastic motion for the case of the microcanonical ensemble which has shown evolution to which the Liouville theorem does not apply, where the entropy increases with time [26]. Our present work links to the work of Tomé & de Oliveira, by identifying the interaction with the reservoir as the physical origin of the stochastic forces introduced by these authors.

Jaynes' treatment of Gibbs' paradox, and the extensivity of the entropy in classical statistical mechanics. The Gibbs' paradox concerns the extensivity of the entropy of ideal gas mixtures in classical statistical mechanics. There is such a volume of literature on the topic that is difficult to give an adequately representative sample [14–16, 18]. However, the essential points are quite simple. Firstly, consider a box with a partition with a controllable valve which divides the box into two equal volumes and imagine closing the valve and filling the two sides with chemically dissimilar gases. If one then opens the valve and allows the gases to intermix, an increase in entropy is expected in this irreversible mixing process, of $2k_\text{B}N.\ln(2)$, where N is the total number of gas molecules, based on classical thermodynamics. The paradox arises from the fact that this predicted increase in entropy persists even when similar (or, in the limit, identical) gases populate the two halves of the container.

---

[2] In fact, even if the reservoir is only close to being canonically distributed, with a small random deviation from canonical form, the same result holds to a high degree of accuracy. If we take the index of probability of the reservoir to be $\eta = \frac{-E}{k_\text{B}T} + \frac{C}{k_\text{B}T} + \delta\eta$, where $\delta\eta$ is a small random deviation (which will be a function of q and p), then the expression for $\Delta\overline{\eta}$ in equation (11) will have an additional term, $\int\ldots\int \Delta\eta\delta\eta \mathrm{e}^{\eta}\mathrm{d}p\mathrm{d}q = \int\ldots\int \Delta P\delta\eta \mathrm{d}p\mathrm{d}q$. If $\delta\eta = \delta\eta_\text{av} + \Delta(\delta\eta)$, where $\delta\eta_\text{av}$ is a constant independent of q and p, and $\Delta(\delta\eta)$ averages to zero, then $\int\ldots\int \Delta P\delta\eta \mathrm{d}p\mathrm{d}q = \int\ldots\int \Delta P\delta\eta_\text{av} \mathrm{d}p\mathrm{d}q + \int\ldots\int \Delta P\Delta(\delta\eta)\mathrm{d}p\mathrm{d}q$. The first term in the preceding equation vanishes for the same reasons discussed above. Since $\Delta(\delta\eta)$ and ΔP are both small random deviations centered on zero and averaging to zero, and are uncorrelated, the integral of their product will average to be very close to zero (and will tend to zero as the reservoir becomes increasingly larger with more degrees of freedom). Hence the additional term $\int\ldots\int \Delta\eta\delta\eta \mathrm{e}^{\eta}\mathrm{d}p\mathrm{d}q$ also vanishes to a high degree of accuracy.

When the valve is opened and the volumes of identical gases allowed to mix and the valve then closed, the classical equations predict an entropy increase which is completely at odds with our common sense view that we have simply done nothing to the state of the system and there is no change in any state function. The paradox is often restated in terms of the lack of extensivity of the entropy function for the gas from classical thermodynamic considerations [27].

The issue of the Gibbs' paradox actually predates Gibbs' work on statistical mechanics and the publication of EPSM in 1902, and appears in classical thermodynamics. In fact, the term seems to have first been used by Wiedeburg in 1894, in relation to Gibbs' earlier work 'On the Equilibrium of Heterogeneous Substances' (EHS; 1875–78) [28, 29]. This issue also arises in earlier classical statistical mechanics and indeed is evident in Boltzmann's work where it is dealt with by a rather casual attitude to additive constants in his work, which is not always clearly noted by later commentators[3].

As mentioned above, there is a huge volume of literature and endless discussion and disagreement on the origin, meaning and interpretation of this paradox. This is greatly complicated by the fact that the issue does not naturally arise in the equations when quantum statistics are applied to ideal gases, and the mathematical origin of this is the identicality of the particles in the quantum treatment. This has led numerous authors to claim the origin of the paradox in classical treatments is in an inadequate consideration of the identicality/indistinguishability of the particles, either fundamentally or in practical terms, and hence that the solution must be to take into account this identicality/indistinguishability in the mathematics. The outcome then is to subtract $\ln(N!)$ or $N.\ln(N)$ at some point in the derivation, generally stating that this 'takes account' of the extra microstates due to interchange of identical/indistinguishable particles which are judged to be physically meaningless [7]. Gibbs uses this approach in chapter XV at the end of EPSM. The topic has an extended reach and has been discussed in terms of the absolute values of the entropy of gases (Sackur–Tetrode equation), including by Ehrenfest and Trkal, and more recently by Swendsen [17, 27, 30, 31].

Given the volume of literature that has been published it is impossible to address all the issues, but it is worth pointing out at the start that the use of subtraction of a factor of $\ln(N!)$ or $N.\ln(N)$ to deal with identicality/indistinguishability must be viewed with great suspicion. If one treats the statistics of identical particles correctly, with no other quantum properties involved, one gets Bose–Einstein statistics (rather than Boltzmann statistics), so the entire premise of using an arbitrary subtraction of a factor of $\ln(N!)$ or $N.\ln(N)$ when deriving classical results is questionable.

---

[3] Swendsen discusses Boltzmann's original entropy definition and the fact that it is extensive, and quotes extensively from Boltzmann's 1877 paper to support a viewpoint that Boltzmann considered this issue carefully in this 1877 paper [4, 5, 17]. However, it is also worth noting that Boltzmann displayed a rather casual attitude to additive constants at times. One quotation worth mentioning but absent from Swendsen's paper, occurs below equation (36) in Boltzmann's paper: 'The variable Ω, which differs from the logarithm of the number of permutations only by an additive constant, is of special importance for this work and we call it the permutability measure. I note, incidentally, that suppression of the additive constants has the advantage that the total permutability measure of two bodies is equal to the sum of the permutability measures of each body'. This comment, mentioned by Boltzmann as incidental, might just as reasonably be interpreted as the extensivity outcome being viewed by Boltzmann essentially as a happy accident. For reference, Boltzmann's original German text is: 'Wir wollen diese GröBe Ω, welche nur durch einen constanten Addenden vom Logarithmus der Permutabilität verschieden ist, und welche für die Folge einen besondere Wichtigkeit hat, als das PermutabilitätsmaB bezeichnen. Ich bemerke übrigens, dass die Unterdrückung dieser Constanten noch den Vortheil hat, daB dadurch bewirkt wird, daB das gesammte PermutabilitätsmaB des Verins zweier Körper gleich der Summe der PermutabilitätsmaBe jedes einzelnen Körpers ist'.

Jaynes' approach to the Gibbs' paradox is considerably more surgical and is presented in chapter 1 of 'Maximum Entropy and Bayesian Methods' (published in 1992 [18]). Rather than grapple with the hydra-headed multitude of symptoms outlined above, he moves quickly to diagnose the origin of the problem and prescribes the correct course of treatment. Jaynes starts his discussion by stating that he found a much clearer treatment of the issue in Gibbs' earlier EHS work than in EPSM and he proceeds from there. He clearly states that there is no paradox because the entropy of mixing in the case of dissimilar gases is fundamentally different to that of identical gases. He focuses on the reverse process, i.e. what external change is required to separate the mixture back to its original thermodynamic state, and specifically is clear that the mathematical approach leading to the $2k_\mathrm{B}N.\ln(2)$ term above is predicated on returning the system to a state where molecules originally in the right-hand volume of the container before mixing return to that side, and those originally in the left-hand volume return there. However, clearly that is not what is meant by returning the system to its original thermodynamic state in the case of identical gases, rather simply that there are equal numbers of gas molecules in both sides, regardless of where they were initially. By contrast for the case of dissimilar gases, to recover the initial thermodynamic state one must return the system to a state where molecules originally in the right-hand volume of the container return to that side, and those originally in the left-hand volume return there. The origin of the problem quickly identified by Jaynes is in the very practical definition of what the term 'original thermodynamic state' actually means, which is vastly different, for very clear reasons, in the cases of identical and dissimilar gases mixing. Jaynes specifies that the thermodynamic state is defined in this case '…by specifying only the chemical composition, total energy, volume, and number of moles of a gas; and nothing else'.

He therefore avoids any metaphysical musing on the identicality or indistinguishability of particles. Of course, the subtraction of a factor of $\ln(N!)$ or $N.\ln(N)$ yields a mathematically suitable outcome, but the attempts to justify this on the basis of removal of physically meaningless additional microstates due to interchange of identical/indistinguishable particles are not well-founded.

Jaynes then moves to discuss the anthropomorphic nature of entropy, using a thought experiment based on imaginary elements 'whifnium' and 'whafnium', which can interact differently with two previously unrecognized types of argon (A1 and A2) which are otherwise identical in all respects. He describes how one's knowledge of this situation affects the thermodynamics, with observable consequences in terms of extractable work processes. Essentially the meaning of 'original state' becomes dependent on an experimenter's knowledge of the system and hence their ability to extract work from the system. This is a crucial point since statements about the anthropomorphic nature of entropy and its subjectivity are often confusing given the seemingly hard-headed practical nature of thermodynamics as the science of work, steam engines and generally practical and objectively real phenomena.

Interestingly Jaynes does not give as much prominence to the distinction between a physical and a thermodynamic system which he emphasized in his 1965 paper mentioned in the previous section, even though he clearly specifies how a thermodynamic state is to be defined. In his 1965 paper he states that '…thermodynamics knows of no such notion as the 'entropy of a physical system'. Thermodynamics does have the concept of the entropy of a thermodynamic system; but a given physical system corresponds to many different thermodynamic systems'. This omission from his later work seems strange to me, since essentially the entire discussion about 'whifnium' and 'whafnium' is about different thermodynamic systems associated with the same physical system. There is no conflict between these views of the same physical because they deal with different thermodynamic systems;

the first deals with a system of identical argon atoms where the experimenter has neither knowledge and/or means and/or desire to distinguish A1 from A2, while the second thermodynamic system deals with a system where the experimenter has such knowledge and/or means and/or desire. Jaynes also makes some insightful related comments towards the end of his 1965 paper on engineers' use of 'steam tables' and how certain thermodynamic degrees of freedom are ignored in the generation and use of these tables [19].

Finally, Jaynes discusses the broader topic, related to my point above on the questionability of attempts to justify subtraction of a factor of $\ln(N!)$ or $N.\ln(N)$ on the basis of removal of physically meaningless additional microstates due to particle identicality or indistinguishability. He restates the basic definition of entropy in classical thermodynamics shown in equation (2) above, emphasizing that it is solely a definition of an entropy change, and is only defined for a closed system (no particles entering or leaving). He then notes that for a closed system the addition of a 'constant' of integration to the integral in equation (2) above with an arbitrary functional dependence on the total particle number N (which is unchanged in a closed system) leaves entropy differences unchanged. He notes that Wolfgang Pauli also noticed this point and that the nature of this additional factor is necessarily an additional condition to be imposed on the system [18, 32]. Furthermore, van Kampen has also written in detail about the same point [33].[4] There is no a priori reason that the entropy of a gas should be extensive, either in classical or quantum statistical mechanics. This property can depend on various factors, including the presence of long-range forces, as Jaynes mentions, even if the particles composing the system are identical. The justification of the subtraction of a factor of $\ln(N!)$ or $N.\ln(N)$ based on the identicality/indistinguishability of particles is clearly at odds with this consideration, as it will be present regardless of the details of long-range forces. Hence the condition/requirement of extensivity must logically be imposed as an additional condition, and, if relevant for a thermodynamic system, can be taken account of for that system by means of a suitable choice of the functional dependence of the 'constant' of integration (e.g. using $f(N) = -\ln(N!)$ or $f(N) = -N.\ln(N)$.) The fact that the use of quantum statistics leads to an extensive entropy is simply a fluke from this perspective. A 'constant' of integration with an arbitrary functional dependence on the total particle number N can also be added to the quantum entropy, which might render it non-extensive, and which might be the physically appropriate description for that system. There is nothing intrinsically associated with quantum mechanics or the identicality/indistinguishability of particles that a priori determines the extensivity of a thermodynamic system.

Jaynes concludes his paper on 'The Gibbs Paradox' with a wider discussion of various topics around Gibbs' likely reaction to Jaynes' views, as well as a discussion of the assumptions Gibbs makes in developing the thermodynamic analogies in chapters IV and XIII in EPSM (which are explicitly based on the assumption that changes in the ensemble bring it from one canonically distributed form to another canonically distributed form and are

---

[4] van Kampen's treatment parallels Jaynes' closely in many respects, including discussion of the arbitrary functional dependence on the total particle number N, the anthropomorphic nature of entropy (and interestingly also mentions steam tables as an example in this context, albeit in a slightly different way to Jaynes' comment), and reaches similar key conclusions about the unimportance of indistinguishability and quantum mechanics. However, his discussion focuses on the different possible processes by which similar and dissimilar gases can be reversibly mixed. Jaynes' approach, following Gibbs' explanation in EHS, initially focuses on the meaning of the original thermodynamic state, and only then moves to discuss the anthropomorphic nature of entropy, external changes/processes required to separate a mixture back to this state, and the arbitrary functional dependence on the total particle number N etc. Notwithstanding the depth of knowledge and understanding evident in van Kampen's work, Jaynes' approach seems to me to be a much better method of presentation for teaching purposes, especially at the undergraduate level.

formulated solely in the context of infinitesimal changes between such forms). He also mentions the association of the absolute value of the thermodynamic entropy with the natural logarithm of the occupation probability density, and the absence of discussion by Gibbs of any *N*-dependent 'constant' of integration factor of the type mentioned above.

In summary, Jaynes proposes that there is no solution required to the Gibbs' paradox, because there is no paradox, rather a lack of clarity over the definition of the meaning of a return to the original thermodynamic state in the cases of similar and dissimilar gases mixing. He continues on to treat the anthropomorphic nature of entropy, and the less well-known issue that extensivity must logically be imposed as an additional condition on any system, and, if relevant for a particular system, can be taken account of for that system by means of a suitable choice of the functional dependence of the arbitrary 'constant' of integration allowed by the classical thermodynamics definition of the entropy changes in a closed system, rather than by insisting on identicality of particles.

Conclusions

As stated in the introduction, the purpose of this contribution is to highlight the insights of Edwin Jaynes on the issues of conservation of phase space density in an isolated system entropy and its relation to changing entropy and the second law of thermodynamics, as well as on the Gibbs' paradox and the extensivity of the entropy in classical statistical mechanics, which were published between 30 and 60 years ago. Jaynes has addressed both these topics in an interesting and logically consistent manner and provided valuable alternative viewpoints. Despite this, Jaynes' insights seem not to have permeated into the mainstream of working and teaching physicists[5].

Additionally, I have tried to provide some broader discussion and physically plausible justifications for both issues in the hope that Jaynes' and Gibbs' works will be (re-)read by a new generation of physics students and teachers and thereby gain a new following.[6] There are of course also other advantages to be gained by students and teachers from looking at Jaynes' earlier work on these issues, including gaining an appreciation of his approach to the definition of quantities in non-equilibrium thermodynamics, as well as an introduction to the broader topics of maximum entropy (MAXENT) methods and information theory for which Jaynes' is famous [21–23]. In fact, Jaynes definition of the 'experimental entropy' (S_e) for any probability distribution W_N as the Gibbs entropy (S_G) calculated for a canonical distribution with the same average energy as the probability distribution W_N is just the MAXENT approach applied to this thermodynamic situation, since the canonical distribution is the extremal one, subject to constant average energy.

Acknowledgments

The author is very grateful to Dr. Paul van Kampen for helpful comments on the manuscript during its preparation and to Prof. George Phillies for earlier discussions about some of the topics contained in it, which greatly clarified the author's thinking. The author would like to

---

[5] The reasons why Gibbs did not make these connections himself, given his profound mathematical and physical insight, are ultimately unclear. However, Jaynes has suggested that Gibbs' failing health towards the end of his life was a key factor and even pinpoints the place in EPSM where he believes the deterioration began to be clearly apparent, i.e. at the beginning of chapter XII [18, 34].

[6] Revisiting the writings of great physicists can turn up many unexpected treasures. Kramers' views on the third law of thermodynamics, as presented by Casimir, are just one more example of this [35, 36].

dedicate this work to the late Mr. Kevin Wade whose broad and life-long interest in, and curiosity about, the natural world, manifest through his wide reading on topics relating to science and natural history, has been an inspiration.

Data availability statement

No new data were created or analyzed in this study.

Appendix. Presentation and extension of Gibbs' analysis of the interaction of a system with a thermal reservoir from chapter XIII of EPSM, based on theorem VII of chapter XI of EPSM

Gibbs indicates that one can consider the combined ensemble as consisting of two parts, which I will call the system ensemble and the reservoir ensemble. The degrees of freedom of the system ensemble are the canonical variables $q_1,\ldots,q_m,\ p_1,\ldots,p_m$, while those of the reservoir ensemble are $q_{m+1},\ldots,q_n,\ p_{m+1},\ldots,p_n$.

We denote the probability density function of the entire combined ensemble as $P(q_1,\ldots,q_n,\ p_1,\ldots,p_n)$ and the index of probability $\eta(q_1,\ldots,q_n,\ p_1,\ldots,p_n)$ is given by $P = \exp(\eta(q_1,\ldots,q_n,\ p_1,\ldots,p_n))$.

The probability for a member of the combined ensemble to be within specified limits is given by:

$$\int\ldots\int \mathrm{e}^{\eta(q_1,\ldots,q_n,p_1,\ldots,p_n)}\mathrm{d}p_1,\ \ldots,\ \mathrm{d}p_n,\ \mathrm{d}q_1,\ \ldots,\ \mathrm{d}q_n, \tag{A1}$$

The probability density function for the system ensemble can then be written as:

$$P_{\text{system}} = \mathrm{e}^{\eta_{system}(q_1,\ldots,q_m,p_1,\ldots,p_m)} = \int\ldots\int \mathrm{e}^{\eta(q_1,\ldots,q_n,p_1,\ldots,p_n)}\mathrm{d}p_{m+1},\ \ldots,\ \mathrm{d}p_n,\ \mathrm{d}q_{m+1},\ \ldots,\ \mathrm{d}q_n. \tag{A2}$$

Where the integral covers all values of the phases of the reservoir (i.e. tracing out these degrees of freedom).

Similarly, the probability density function for the reservoir ensemble can similarly be written as:

$$P_{\text{reservoir}} = \mathrm{e}^{\eta_{\text{reservoir}}(q_{m+1},\ldots,q_n,p_{m+1},\ldots,p_n)} = \int\ldots\int \mathrm{e}^{\eta(q_1,\ldots,q_n,p_1,\ldots,p_n)}\mathrm{d}p_1,\ \ldots,\ \mathrm{d}p_m,\ \mathrm{d}q_1,\ \ldots,\ \mathrm{d}q_m, \tag{A3}$$

where the integral covers all values of the phases of the system (again tracing out these degrees of freedom).

Hence the probability for the system ensemble to be within specified limits is given by:

$$\int\ldots\int \mathrm{e}^{\eta_{\text{system}}}\mathrm{d}p_1 1,\ \ldots,\ \mathrm{d}p_m,\ \mathrm{d}q_1,\ \ldots,\ \mathrm{d}q_m. \tag{A4}$$

With the limits of integration expressed in the coordinates of the system.

And the probability for the reservoir ensemble to be within specified limits is given by:

$$\int\ldots\int \mathrm{e}^{\eta_{\text{reservoir}}}\mathrm{d}p_{m+1},\ldots,\mathrm{d}p_n,\ \mathrm{d}q_{m+1},\ldots,\mathrm{d}q_n. \tag{A5}$$

With the limits of integration expressed in the coordinates of the reservoir.

The normalization conditions are:

$$1 = \int \ldots \int \mathrm{e}^{\eta} \mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}q_1, \ldots, \mathrm{d}q_n, \tag{A6}$$

$$1 = \int \ldots \int \mathrm{e}^{\eta_{\text{system}}} \mathrm{d}p_1, \ldots, \mathrm{d}p_m, \mathrm{d}q_1, \ldots, \mathrm{d}q_m, \tag{A7}$$

$$1 = \int \ldots \int \mathrm{e}^{\eta_{\text{reservoir}}} \mathrm{d}p_{m+1}, \ldots, \mathrm{d}p_n, \mathrm{d}q_{m+1}, \ldots, \mathrm{d}q_n. \tag{A8}$$

Where the limits in each case cover all values of the phases associated with the relevant variables.

Gibbs defines the entropy of an ensemble, S, with an arbitrary index of probability $\eta(q, p)$ as:

$$S = -k_{\mathrm{B}} \int \ldots \int \eta \mathrm{e}^{\eta} \mathrm{d}p \mathrm{d}q = -k_{\mathrm{B}} \overline{\eta}. \tag{A9}$$

Hence the entropy of the combined ensemble, the system and the reservoir are:

$$S_{\mathrm{CE}} = -k_{\mathrm{B}} \int \ldots \int \eta \mathrm{e}^{\eta} \mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}q_1, \ldots, \mathrm{d}q_n = -k_{\mathrm{B}} \overline{\eta}, \tag{A10}$$

$$S_{\text{system}} = -k_{\mathrm{B}} \int \ldots \int \eta_{\text{system}} \mathrm{e}^{\eta_{\text{system}}} \mathrm{d}p_1, \ldots, \mathrm{d}p_m, \mathrm{d}q_1, \ldots, \mathrm{d}q_m = -k_{\mathrm{B}} \overline{\eta_{\text{system}}}, \tag{A11}$$

$$S_{\text{reservoir}} = -k_{\mathrm{B}} \int \ldots \int \eta_{\text{reservoir}} \mathrm{e}^{\eta_{\text{reservoir}}} \mathrm{d}p_{m+1}, \ldots, \mathrm{d}p_n, \mathrm{d}q_{m+1}, \ldots, \mathrm{d}q_n = -k_{\mathrm{B}} \overline{\eta_{\text{reservoir}}}. \tag{A12}$$

Gibbs wishes to show that the sum of the average values of the probability distributions of the system and reservoir ensembles taken separately must decrease compared to that of the average value of the combined ensemble, i.e. $S_{\text{system}} + S_{\text{reservoir}} \geqslant S_{\mathrm{CE}}$. If this is true then the sum of equations (A10)–(A12) should obey:

$$\int \ldots \int (\eta - \eta_{\text{system}} - \eta_{\text{reservoir}}) \mathrm{e}^{\eta} \mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}q_1, \ldots, \mathrm{d}q_n \geqslant 0. \tag{A13}$$

Multiplying equations (A7) and (A8) and adding to equations (A13) and then subtracting equations (A6) (which in total is an addition of a quantity identically equal to zero) gives:

$$\int \ldots \int \{(\eta - \eta_{\text{system}} - \eta_{\text{reservoir}}) \mathrm{e}^{\eta} + \mathrm{e}^{\eta_{\text{system}} + \eta_{\text{reservoir}}} - \mathrm{e}^{\eta}\} \mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}q_1, \ldots, \mathrm{d}q_n \geqslant 0 \tag{A14}$$

Denoting $u = \eta - \eta_{\text{system}} - \eta_{\text{reservoir}}$ we have:

$$\int \ldots \int \{u \mathrm{e}^{u} + 1 - \mathrm{e}^{u}\} \mathrm{e}^{\eta_{\text{system}} + \eta_{\text{reservoir}}} \mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}q_1, \ldots, \mathrm{d}q_n \geqslant 0 \tag{A15}$$

The condition $S_{\text{system}} + S_{\text{reservoir}} \geqslant S_{\mathrm{CE}}$ will be satisfied if $u\mathrm{e}^{u} + 1 - \mathrm{e}^{u} \geqslant 0$. Gibbs shows just below (EPSM.423) that this condition is true, and the equality sign only holds if $u = 0$ for all elements of the combined ensemble, i.e. if the distributions of the phases in of the system and reservoir are completely independent, that the system and reservoir are not interacting. If they interact then the inequality holds.

If we imagine that prior to an initial time t_0 that the system and reservoir ensembles are not interacting, then $\overline{\eta^0} - \overline{\eta^0_{\text{system}}} - \overline{\eta^0_{\text{reservoir}}} = 0$. If they are put into thermal contact at t_0 they will then interact as a combined ensemble and at a later time t_1 we will have that $\overline{\eta^1} - \overline{\eta^1_{\text{system}}} - \overline{\eta^1_{\text{reservoir}}} > 0$.

This analysis underpins the general effects of interactions between a system and reservoir, including heat transfer. However, in the context of the present work I present a slightly more focussed example with no net heat exchange, as described in the main text above.

References

[1] Brush S G 1983 *Statistical Physics and the Atomic Theory of Matter; from Boyle and Newton to Landau and Onsager* (Princeton University Press)

[2] Wightman A S 1989 On the prescience of J. Willard Gibbs *Proceedings of the Gibbs Symposium* ed D G Caldi and G D Mostow (American Mathematical Society & American Institute of Physics)

[3] Mehra J 1998 Josiah Willard Gibbs and the foundations of statistical mechanics *Found. Phys.* **28** 1785

[4] Sharp K and Matschinsky F 2015 Translation of Ludwig Boltzmann's paper 'on the relationship between the second fundamental theorem of the mechanical theory of heat and probability calculations regarding the conditions for thermal equilibrium' Sitzungberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch–Naturwissen Classe. Abt. II, LXXVI 1877, pp 373-435 (Wien. Ber. 1877, 76:373-435). Reprinted in Wiss. Abhandlungen, Vol. II, reprint 42, p. 164-223, Barth, Leipzig, 1909 *Entropy* **17** 1971–2009

[5] Boltzmann L 1877 Über die Beziehung zwischen dem zweiten Hauptsatze des mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht, Sitzungsberichte Der Mathematisch–Naturwissenschaftlichen Classe Der Kaiserlichen Akademie Der Wissenschaften **76** 373

[6] Ehrenfest P and Ehrenfest T 1990 *The Conceptual Foundations of the Statistical Approach in Mechanics* (Dover Publications)

[7] Gibbs J W 1960 *Elementary Principles in Statistical Mechanics* (Dover Publications)

[8] Lebowitz J L 1993 Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy *Physica A* **194** 1

[9] Epstein P S 1936 Critical appreciation of Gibbs' statistical mechanics *A Commentary on the Scientific Writings of J. Willard Gibbs_Volume II-Theoretical Physics* ed A Haas (Yale University Press)

[10] Haas A 1936 The chief results of Gibbs' statistical mechanics *A Commentary on the Scientific Writings of J. Willard Gibbs_Volume II-Theoretical Physics* ed A Haas (Yale University Press)

[11] Popp B D *Contemporary Reaction to Gibbs's Statistical Mechanics* https://arxiv.org/pdf/2412.11425

[12] Sethna J P 2006 *Statistical Mechanics; Entropy, Order Parameters, and Complexity* (Oxford University Press)

[13] Thorne K S and Blandford R D 2017 *Modern Classical Physics; Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics* (Princeton University Press)

[14] Saunders S 2018 The Gibbs paradox *Entropy* **20** 552

[15] Dieks D 2018 The Gibbs paradox and particle individuality *Entropy* **20** 466

[16] Ben-Naim A 2007 On the so-called Gibbs paradox, and on the real paradox *Entropy* **9** 132–6

[17] Swendsen R H 2006 Statistical mechanics of colloids and Boltzmann's definition of the entropy *Am. J. Phys.* **74** 187

[18] Jaynes E T 1992 The Gibbs paradox *Maximum Entropy and Bayesian Methods—Seattle, 1991* ed C R Smith *et al* (Springer Science+Business Media, B.V.)

[19] Jaynes E T 1965 Gibbs versus Boltzmann entropies *Am. J. Phys.* **33** 391

[20] Phillies G D J 2000 *Elementary Lectures in Statistical Mechanics* (Springer)

[21] Jaynes E T 1978 Where do we stand on maximum entropy *The Maximum Entropy Formalism—A Conference Held at the Massachusetts Institute of Technology on May 2–4, 1978* ed R D Levine and M Tribus (The MIT Press)

[22] Jaynes E T 1985 Where do we go from here *Maximum-Entropy and Bayesian Methods in Inverse Problems* ed C R Smith and W T Grandy Jr (D. Reidel Publishing Company)

[23] Jaynes E T 2003 *Probability Theory; The Logic of Science* (Cambridge University Press)

[24] Pippard A B 1966 *Elements of Classical Thermodynamics for Advanced Students of Physics* (Cambridge Univ. Press)

[25] Callen H B 1985 *Thermodynamics and an Introduction to Thermostatistics* 2nd edn (Wiley)

[26] Tomé T and De Oliveira M J 2022 Stochastic motion in phase space on a surface of constant energy *Phys. Rev. E* **106** 034129

[27] Swendsen R H 2002 Statistical mechanics of classical systems with distinguishable particles *J. Stat. Phys.* **107** 1143

[28] Wiedeburg O 1894 Das Gibbs'sche paradoxon *Ann. Phys.* **289** 684

[29] Gibbs J W 1876 On the equilibrium of heterogeneous substances *Trans. Connecticut Acad. Arts Sci.* **3** 108

[30] Ehrenfest P and Trkal V 1921 Deduction of the dissociation-equilibrium from the theory of quanta and a calculation of the chemical constant based on this *Proc. R. Netherlands Acad. Arts Sci.* **23** 162

[31] Ehrenfest P and Trkal V 1921 Ableitung des dissoziationsgleichgewichtes aus der quantentheorie und darauf beruhende berechnung der chemischen konstanten *Ann. Phys.* **370** 609

[32] Pauli W 1973 *Pauli Lectures on Physics; Thermodynamics and the Kinetic Theory of Gases* vol 3 (The MIT Press)

[33] van Kampen N G 1984 The Gibbs paradox *Essays in Theoretical Physics; In Honour of Dirk Ter Haar* ed W E Parry (Pergamon)

[34] Jaynes E T 1967 Foundations of probability theory and statistical mechanics *Delaware Seminar in the Foundations of Physics* (Springer)

[35] Casimir H B G 1978 Nernst's theorem XXX *Fundamental Problems in Statistical Mechanics IV : Proc. Int. Summer School in Statistical Mechanics IV* ed E G D Cohen *et al* XXX

[36] Casimir H B G 1963 Über die statistische Begründung des Nernstschen Wärmetheorems *Z. Phys.* **171** 246