Article

Constraints of Compound Systems: Prerequisites for Thermodynamic Modeling Based on Shannon Entropy

Martin Pfleger *, **Thomas Wallek and Andreas Pfennig**

Institute of Chemical Engineering and Environmental Technology, Graz University of Technology, NAWI Graz, Inffeldgasse 25/C/I, 8010 Graz, Austria

* Author to whom correspondence should be addressed; E-Mail: martin.georg.pfleger@gmail.com, Tel.: +43-676-4058140.

Abstract: Thermodynamic modeling of extensive systems usually implicitly assumes the additivity of entropy. Furthermore, if this modeling is based on the concept of Shannon entropy, additivity of the latter function must also be guaranteed. In this case, the constituents of a thermodynamic system are treated as subsystems of a compound system, and the Shannon entropy of the compound system must be subjected to constrained maximization. The scope of this paper is to clarify prerequisites for applying the concept of Shannon entropy and the maximum entropy principle to thermodynamic modeling of extensive systems. This is accomplished by investigating how the constraints of the compound system have to depend on mean values of the subsystems in order to ensure additivity. Two examples illustrate the basic ideas behind this approach, comprising the ideal gas model and condensed phase lattice systems as limiting cases of fluid phases. The paper is the first step towards developing a new approach for modeling interacting systems using the concept of Shannon entropy.

Keywords: Shannon entropy; constrained extremalization; compound systems; additivity; thermodynamic modeling; discrete states; discrete modeling

1. Introduction

In his basic work, Shannon [1] defines a function H which measures the amount of information of a system which can reside in either of m possible states by means of the probabilities p_i of the states:

$$H(p_1, \ldots, p_m) = -K \sum_{i=1}^{m} p_i \log p_i \tag{1}$$

Both, the constant K and the basis of the logarithm are arbitrary as they just account for a scaling of H. The set of all p_i can be written as probability distribution $\overline{p}$:

$$\overline{p} = \{p_1, \ldots, p_m\}\,,$$

with the normalization condition

$$\sum_{i=1}^{m} p_i = 1 \tag{2}$$

In the following we set $K = 1$ and choose the natural logarithm. When building the sum over all states, the limits of the summation $(1 \ldots m)$ can be omitted and H can formally be written as function of the probability distribution:

$$H(\overline{p}) = -\sum_i p_i \ln p_i \tag{3}$$

Previous papers [1–8] worked out that this measure has all the properties of thermodynamic entropy as introduced by statistical physics. Throughout this paper we call $H(\overline{p})$ as defined in Equation (3) the *Shannon entropy* of the system under consideration, in order to distinguish it from thermodynamic entropy.

The range of $H(\overline{p})$ is given by $0 \le H(\overline{p}) \le \ln m$. The zero value for a distribution is where one of the p_i equals 1 and, because of the normalization condition (2), all other p_i are zero. The maximum value is given for uniformly distributed probabilities [6]:

$$p_i = \frac{1}{m}\,, i = 1 \ldots m \tag{4a}$$

$$\Rightarrow H(\overline{p}) = \ln m \tag{4b}$$

1.1. Compound Systems

We consider a compound system composed of N subsystems, each characterized by its individual probability distribution:

$$\{\overline{p_1}\} = \{p_{1,1}, \ldots, p_{1,m_1}\}$$

$$\vdots$$

$$\{\overline{p_i}\} = \{p_{i,1}, \ldots, p_{i,m_i}\}$$

$$\vdots$$

$$\{\overline{p_N}\} = \{p_{N,1}, \ldots, p_{N,m_N}\}$$

The state of the compound system is defined by the states of the subsystems. We therefore write the probability distribution of the compound system as

$$\{\overline{p_c}\} = \{p_{1,1,\dots,1}, \dots, p_{i_1,\dots,i_N}, \dots, p_{m_1,\dots,m_N}\}, \tag{5}$$

where $p_{i_1,\dots,i_N}$ is the probability of the compound state where subsystem 1 is in the state i_1, subsystem 2 is in the state i_2, and so on. It is not necessary for the subsystems to have identical probability distributions.

Generally the probability of the compound state $A \cdot B$, comprising the states A and B of two subsystems is given by

$$p(A \cdot B) = p(A/B) \cdot p(B)$$

where $p(A/B)$ is the probability of subsystem 1 to be in state A, given that subsystem 2 is in state B. If the subsystems are statistically independent, *i.e.*, $p(A/B) = p(A)$, then it follows (cf. [9,10]):

$$p(A \cdot B) = p(A) \cdot p(B) \tag{6}$$

If all N subsystems comprising the considered compound system are statistically independent, straightforward application of Equation (6) to the probability distribution (5) gives:

$$\{\overline{p}\} = p_{i_1,\dots,i_N} = p_{i_1} \cdot \dots \cdot p_{i_N}$$

With this probability distribution the Shannon entropy of the compound system is:

$$H = -\sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_N=1}^{m_N} \left(p_{i_1,i_2,\dots,i_N} \ln p_{i_1,i_2,\dots,i_N} \right)$$
$$H = H_1 + H_2 + \dots + H_N \tag{7}$$

Hence the Shannon entropy of independent subsystems is additive. In the special case of N equal and statistically independent subsystems, *i.e.*,

$\{\overline{p_1}\} = \dots = \{\overline{p_N}\} \equiv \{\overline{p_s}\}$, the homogeneity of the Shannon entropy of the compound system follows:

$$H_{\mathrm{c}} = N \, H_{\mathrm{s}} \tag{8}$$

Throughout this paper the index s is used for single systems and c for compound systems.

1.2. The Bridge to Thermodynamic Entropy

Considering a compound system composed of N equal and statistically independent subsystems, all subsystems are characterized by the probability distribution $\{\overline{p_s}\}$, and homogeneity, Equation (8), is guaranteed. For a large number of subsystems, *i.e.*, $N >> 1$, the probabilities p_i can be expressed as relative occupation numbers N_i, designating the number of subsystems residing in the state i:

$$p_i \approx \frac{N_i}{N}, \tag{9}$$

with

$$N = \sum_{i=1}^{m} N_i$$

the total number of subsystems within the compound system, playing the role of the normalization condition (2). Expressing the Shannon entropy of a single subsystem with the occupation numbers in Equation (9) gives

$$H_{\mathrm{s}} = -\sum \frac{N_i}{N} \ln \frac{N_i}{N} \ ,$$

and because of homogeneity, Equation (8), the Shannon entropy of the compound system is:

$$H_{\mathrm{c}} = -\sum_i N_i \ln N_i + N \ln N \ . \tag{10}$$

The right hand side of Equation (10) has the same form as the logarithm of the thermodynamic probability W known from classical statistical mechanics [11]:

$$\ln W(N_1, N_2, \ldots, N_j, \ldots) = -\sum_j N_j \ln N_j + N \ln N \tag{11}$$

The variables N_j in Equation (11) give the number of particles in cell j of the μ-space, and have the very same meaning as the occupation numbers N_i in Equation (10), *i.e.*, the numbers of particles residing in the (mechanical) state i. The set $(N_1, N_2, \ldots, N_j, \ldots)$ is called the occupation of the μ-space, and the thermodynamic probability W is the number of microstates realizing the given occupation. Hence, the left hand sides of Equations (10) and (11) stand for the same measure and can be combined to:

$$H = \ln W.$$

Table 1 compares the concepts behind the two measures. Because of

$$S = k_{\mathrm{B}} \ln W$$

where S is the thermodynamic entropy of the system we get the result:

$$S = k_{\mathrm{B}} H$$

This equation reveals the equivalence between Shannon entropy and thermodynamic entropy related by the Boltzmann constant.

Table 1. Comparison between Shannon entropy H and thermodynamic probability W.

Shannon entropy H	thermodynamic probability W
probability distribution: $\{\overline{p}\} = \{p_1, p_2, \ldots, p_i, \ldots\}$	occupation: $(N_1, N_2, \ldots, N_j, \ldots)$
$H = -N \sum_i p_i \ln p_i$ $\sum_i p_i = 1$	$W = \frac{N!}{N_1! N_2! \ldots N_j! \ldots}$ $\sum_j N_j = N$
assumption $N >> 1 \Rightarrow p_i \approx \frac{N_i}{N}$ equal subsystems, statistically independent	 $N, N_1, \ldots, N_j >> 1 \Rightarrow \ln N! \approx N \ln N$ Stirling's formula applied to N and to all N_j
$H = -\sum_i N_i \ln N_i + N \ln N$	$\ln W = -\sum_j N_j \ln N_j + N \ln N$

1.3. Additivity of Shannon Entropy: Its Significance for Thermodynamic Modeling

Homogeneity of a compound system, Equation (8), is the starting point for thermodynamic modeling based on the states of its constituents. On the one hand, as shown in section 1.2, it builds the bridge between Shannon entropy and the classical thermodynamic entropy. On the other hand, the crucial property of additivity enables the calculation of thermodynamic entropy simply by calculating the Shannon entropy of a single subsystem. Subsequently, the compound system's entropy can be expressed by the sum of entropies of the constituting subsystems.

Applied to a gas or fluid this means to derive the Shannon entropy of one atom or molecule based on their respective states. When speaking of molecular or discrete states we do not necessarily consider the quantum-mechanical states of atoms or molecules; the mechanical, continuous states of atoms or molecules are also possible candidates. But when using the discrete formulation of Shannon entropy, Equation (1), a discretization of the continuous states is helpful.

When applying lattice models for describing condensed phase systems, the goal is reduced to derivation of the Shannon entropy of a single lattice site. However, when deriving the Shannon entropy of a compound system by utilizing homogeneity, Equation (8), we made the following preassumptions:

The first is the assumption of statistically independent subsystems, which may be plausible for many thermodynamic systems as long as the subsystems (the particles) are 'not too strongly correlated in some *nonlocal* sense' [10].

We did not emphasize the second assumption, because it seems to be self-evident: We used the probability distributions as given system variables, as if they were properties of the subsystems, which - for statistically independent systems - stay constant. But as known from classical thermodynamics, the

entropy of a system depends on system variables like internal energy, temperature, pressure and so on. In addition, entropy is not predefined directly by these variables, but underlies a maximization principle, stating that a system in thermodynamic equilibrium resides in a state where entropy is a maximum, with respect to the constraints given by the system variables. This means that we cannot deal a priori with *given* probability distributions, but we have to *determine* the very probability distribution which maximizes entropy with respect to the constraints. Hence, if we ask for validity of additivity, Equation (8), we have to investigate the maximized Shannon entropies of a compound system and its constituting single systems separately, as shown in the following section.

2. Probability Distributions with Maximum Entropy

2.1. Maximization of Unconstrained Systems

As can be seen from Equation (4b), the maximum value of the Shannon entropy of unconstrained systems, *i.e.*, uniformly distributed states, depends solely on the number of their possible states. For a single system with m possible states we get $H_s = \ln m$. In the case of a compound system consisting of N single systems, each with m possible states, the number of possible states is m^N, resulting in a Shannon entropy of

$$H_c = N \ln m ,$$

i.e., N times the Shannon entropy of the single system (cf. Equation (4b)). So for unconstrained systems homogeneity, Equation (8), is guaranteed.

2.2. Constrained Maximization of a Single System

Now a single system with m possible states is considered, each of them characterized by the value f_i of a random variable F. Let one constraint be given by the mean value $\langle f \rangle$ of the random variable F:

$$\langle f \rangle = \sum_{i=1}^{m} f_i \, p_i \tag{12}$$

A lot of probability distributions $\overline{p}$ may result in the same mean value $\langle f \rangle$. Among these probability distributions we are looking for the one yielding the maximum value H_s for the Shannon entropy. The maximizing probability distribution and the resulting Shannon entropy H_s will depend on the exact choice for $\langle f \rangle$, so that both can be expressed as functions of this constraint:

$$\overline{p} = \overline{p}(\langle f \rangle)$$
$$H_s = H_s(\langle f \rangle)$$

The maximizing probability distribution considering constraint (12) and the normalization condition (2) can be found by applying Lagrange's method of constrained extremalization [12]. This method introduces the Lagrange function

$$\mathcal{L} = H_s - \lambda_0 \left(1 - \sum_{i=1}^{m} p_i \right) - \lambda_1 \left(\langle f \rangle - \sum_{i=1}^{m} f_i \, p_i \right) \overset{!}{=} max., \tag{13}$$

where λ_0 and λ_1 are the Lagrangian multipliers. $\mathcal{L}$ has to be maximized by equating the derivations with respect to all p_i to zero:

$$\frac{\partial}{\partial p_i} \mathcal{L} = 0 \ \forall i$$

With H_s from Equation (3) and performing all derivations we get:

$$p_i = e^{(-1+\lambda_0 + f_i \lambda_1)} \ \ \forall i \tag{14}$$

Inserting Equation (14) into the normalization condition (2) yields:

$$\sum_{i=1}^{m} p_i = 1 = e^{(-1+\lambda_0)} \sum_{i=1}^{m} e^{f_i \lambda_1} \tag{15}$$

Inserting Equation (14) into constraint (12) results in:

$$\sum_{i=1}^{m} f_i \, p_i = \langle f \rangle = e^{(-1+\lambda_0)} \sum_{i=1}^{m} f_i \, e^{f_i \lambda_1} \tag{16}$$

Using the abbreviation $x_s \equiv e^{\lambda_1}$, combination of Equations (15) and (16) yields:

$$\langle f \rangle = \frac{\sum\limits_{i=1}^{m} f_i \, x_s^{f_i}}{\sum\limits_{i=1}^{m} x_s^{f_i}} \tag{17}$$

This equation can be solved numerically for x_s, which can now be used to express the maximizing probability distribution. For that purpose, Equations (14) and (15) can be rewritten as:

$$p_i = e^{(-1+\lambda_0)} \cdot x_s^{f_i}$$

$$1 = e^{(-1+\lambda_0)} \cdot \sum_{i=1}^{m} x_s^{f_i}$$

Combining these equations yields the maximizing probability distribution

$$p_i = \frac{x_s^{f_i}}{\sum\limits_{j=1}^{m} x_s^{f_j}} \ . \tag{18}$$

Inserting (18) into definition (3) results in:

$$H_s = - \sum_{i=1}^{m} \frac{x_s^{f_i}}{\sum\limits_{j=1}^{m} x_s^{f_j}} \ln \frac{x_s^{f_i}}{\sum\limits_{j=1}^{m} x_s^{f_j}}$$

The denominator of the first factor does not depend on the index i of the outer sum and can be put in front of the sum. The logarithm of the fraction is now written as sum of two terms:

$$H_s = - \frac{1}{\sum\limits_{j=1}^{m} x_s^{f_j}} \sum_{i=1}^{m} x_s^{f_i} \left(\ln x_s^{f_i} - \ln \sum_{j=1}^{m} x_s^{f_j} \right)$$

Further rearrangement finally yields:

$$H_{\mathrm{s}} = -\frac{\sum\limits_{i=1}^{m} x_{\mathrm{s}}^{f_i} \ln x_{\mathrm{s}}^{f_i}}{\sum\limits_{j=1}^{m} x_{\mathrm{s}}^{f_j}} + \ln \sum\limits_{i=1}^{m} x_{\mathrm{s}}^{f_i} \tag{19}$$

This value represents the maximum value of the Shannon entropy among all probability distributions compatible with the normalization condition and the constraint, Equation (12).

2.3. Constrained Maximization of a Cmpound System

Given a compound system composed of N single systems as mentioned in the last section, each of which can reside in either of m possible states and with each state i being assigned a value f_i of a discrete random variable F, the number of possible states is m^N. With q_k the probabilities of the compound system, where k denotes the state, when the single system 1 is in state i_1, single system 2 is in state i_2 and so on: $k \mathrel{\widehat{=}} i_1, \ldots, i_N$, the probabilities can be rewritten as:

$$q_k = q_{i_1,\ldots,i_N} \tag{20}$$

with

$$
\begin{aligned}
k &= 1 \ldots m^N \\
i_1 &= 1 \ldots m \\
&\vdots \\
i_N &= 1 \ldots m
\end{aligned}
$$

Let G be a random variable related to the states of the compound system, and g_k be the according value of G in the state k. g_k is chosen in such a way, that it represents the sum of the random variables F of the single systems:

$$g_k = g_{i_1,\ldots,i_N} = f_{i_1} + \ldots + f_{i_N} \tag{21}$$

with f_{i_1} the value of F of particle 1 in state i_1 and so on. The mean value $\langle g \rangle$ of G, which will act as constraint for the compound system, is (cf. Equation (12)):

$$\langle g \rangle = \sum_{k=1}^{m^N} g_k \, q_k \tag{22}$$

$$= \sum_{i_1=1}^{m} \ldots \sum_{i_N=1}^{m} g_{i_1,\ldots,i_N} \, q_{i_1,\ldots,i_N} \tag{23}$$

$$= \langle f_1 \rangle + \ldots + \langle f_N \rangle \tag{24}$$

The mean value of the compound system is the sum of the mean values of the single systems. If the single systems are equal, their mean value is the same:

$$\langle f_1 \rangle = \langle f_2 \rangle = \ldots = \langle f_N \rangle \equiv \langle f \rangle \,, \tag{25}$$

resulting in

$$\langle g \rangle = N \langle f \rangle . \tag{26}$$

Now we are looking for the probability distribution $\overline{q}$ which fulfills the normalization condition, guarantees the mean value $\langle g \rangle$ given by the constraint, and yields the maximum value H_c for the Shannon entropy. The result will depend on the constraint $\langle g \rangle$:

$$\overline{q} = \overline{q}(\langle g \rangle) \tag{27}$$

$$H_c = -\sum_{k=1}^{m^N} q_k \ln q_k = H_c(\langle g \rangle)$$

Again applying Lagrange's method, the solutions (17), (18) and (19) of the single system can be reused by replacing $f \to g$, $x_s \to x_c$, $H_s \to H_c$ and considering that the number of states is now m^N. With x_c being the solution of (cf. Equation (17))

$$\langle g \rangle = \frac{\sum_{i=1}^{m^N} g_i \, x_c^{g_i}}{\sum_{j=1}^{m^N} x_c^{g_j}} \tag{28}$$

the probabilities of the compound system are (cf. Equation (18))

$$q_k = \frac{x_c^{g_k}}{\sum_{j=1}^{m^N} x_c^{g_j}} \ , \ k = 1 \dots m^N \ , \tag{29}$$

and the Shannon entropy of the compound system is (cf. Equation (19)):

$$H_c = -\frac{\sum_{i=1}^{m^N} x_c^{g_i} \ln x_c^{g_i}}{\sum_{j=1}^{m^N} x_c^{g_j}} + \ln \sum_{i=1}^{m^N} x_c^{g_i} \tag{30}$$

Inserting Equation (21) into Equation (28) results in:

$$\langle g \rangle = N \frac{\sum_{k=1}^{m} f_k x_c^{f_k}}{\sum_{k=1}^{m} x_c^{f_k}} \tag{31}$$

as explicitly derived in the supplementary material. Taking into account Equation (26) we get:

$$\langle f \rangle = \frac{\sum_{k=1}^{m} f_k x_c^{f_k}}{\sum_{k=1}^{m} x_c^{f_k}} \tag{32}$$

By comparing Equation (32) with Equation (17) one can see that the solutions x_s and x_c fulfill the same equations; they are equal and their subscripts can therefore be omitted:

$$x_s = x_c = x \tag{33}$$

Now we again use Equation (21) and insert it into Equation (30). The result is:

$$H_{\text{c}} = N \left(-\frac{\sum\limits_{i=1}^{m} x_{\text{c}}^{f_i} \ln x_{\text{c}}^{f_i}}{\sum\limits_{i=1}^{m} x_{\text{c}}^{f_i}} + \ln \sum\limits_{i=1}^{m} x_{\text{c}}^{f_i} \right) \tag{34}$$

with intermediate steps given in the supplementary material. The same expression can be obtained when Equations (20) and (21) are inserted into the probabilities, Equation (29), resulting in

$$q_{i_1,i_2,\ldots,i_N} = \frac{x_{\text{c}}^{f_{i_1}} \cdot x_{\text{c}}^{f_{i_2}} \cdot \ldots \cdot x_{\text{c}}^{f_{i_N}}}{\left(\sum\limits_{j=1}^{m} x_{\text{c}}^{f_j} \right)^N} , \tag{35}$$

and using this for calculating Shannon entropy by means of Equation (3). This alternative derivation is also included in the supplementary material.

Because of the equivalence of x_{s} and x_{c}, Equation (33), we can rewrite Equations (19) and (34):

$$H_{\text{s}} = -\frac{\sum\limits_{i=1}^{m} x^{f_i} \ln x^{f_i}}{\sum\limits_{j=1}^{m} x^{f_j}} + \ln \sum\limits_{i=1}^{m} x^{f_i}$$

$$H_{\text{c}} = N \left(-\frac{\sum\limits_{i=1}^{m} x^{f_i} \ln x^{f_i}}{\sum\limits_{i=1}^{m} x^{f_i}} + \ln \sum\limits_{i=1}^{m} x^{f_i} \right)$$

Comparing both these equations yields the result:

$$H_{\text{c}} = N H_{\text{s}} \tag{36}$$

Equation (36) illustrates that homogeneity is also fulfilled for compound systems underlying an extremalization principle with respect to one constraint. The crucial assumption we made is that the constraint of the subsystems and the constraint of the compound system obey equation (21).

2.4. Systems Underlying Several Constraints

To be more general we consider systems under several constraints, again beginning with a single system. Let α be the number of random variables $F_1, F_2, \ldots F_\alpha$ as constraints, and m the number of possible states. The value of F_1 in the state i is $f_{1,i}$, the according value of F_2 is $f_{2,i}$ etc. The constraints are given by the mean values $\langle f_1 \rangle, \langle f_2 \rangle, \ldots, \langle f_\alpha \rangle$ of the random variables, with

$$\langle f_1 \rangle = \sum_i p_i \, f_{1,i}$$

$$\langle f_2 \rangle = \sum_i p_i \, f_{2,i}$$

$$\vdots$$

$$\langle f_\alpha \rangle = \sum_i p_i \, f_{\alpha,i} \, .$$

Straightforward application of Lagrange's method of constrained extremalization results [1.0] in (cf. Equation (18)):

$$p_i = \frac{A_i}{\sum\limits_j A_j} \, , \tag{37}$$

with the abbreviation

$$A_i = X_1^{f_{1,i}} \cdot X_2^{f_{2,i}} \cdot \ldots \cdot X_\alpha^{f_{\alpha,i}} \tag{38}$$

and the $X_1, X_2, \ldots, X_\alpha$ being the solutions of the following system of equations (cf. Equation (17)):

$$\frac{\sum\limits_i f_{1,i} \, A_i}{\sum\limits_j A_j} = \langle f_1 \rangle$$

$$\vdots \tag{39}$$

$$\frac{\sum\limits_i f_{\alpha,i} \, A_i}{\sum\limits_j A_j} = \langle f_\alpha \rangle$$

Calculating the Shannon entropy with the probabilities given by Equation (37) yields (cf. Equation (19)):

$$H_{\mathrm{s}} = -\frac{\sum\limits_i A_i \ln A_i}{\sum\limits_j A_j} + \ln \sum\limits_i A_i \tag{40}$$

We consider a compound system composed of N of these single systems, with α random variables $G_1, G_2, \ldots, G_\alpha$, which are associated to the random variables of the single systems in the same way indicated by Equations (22), (23), (24), (25) and (26). The value of G_1 in the state k is $g_{1,k}$, the according value of G_2 is $g_{2,k}$ and so on, and they are related to the random variables of the single system by (cf. Equation (21)):

$$g_{1,k} = f_{1,i_1} + f_{1,i_2} + \ldots + f_{1,i_N} \tag{41}$$

The constraints are given as the mean values $\langle g_1 \rangle, \langle g_2 \rangle, \ldots, \langle g_\alpha \rangle$ of the random variables, with

$$\langle g_1 \rangle = \sum_k q_k \, g_{1,k} = N \langle f_1 \rangle$$

$$\langle g_2 \rangle = \sum_k q_k \, g_{2,k} = N \langle f_2 \rangle$$

$$\vdots \tag{42}$$

$$\langle g_\alpha \rangle = \sum_k q_k \, g_{\alpha,k} = N \langle f_\alpha \rangle \, .$$

Straightforward application of Lagrange's method of constrained extremalization results in (cf. Equations (18) and (37)):

$$q_k = \frac{B_k}{\sum\limits_j B_j} \, , \tag{43}$$

with the abbreviation

$$B_k = Y_1^{g_{1,k}} \cdot Y_2^{g_{2,k}} \cdot \ldots \cdot Y_\alpha^{g_{\alpha,k}} \tag{44}$$

and with the $Y_1, Y_2, \ldots, Y_\alpha$ being the solutions of the following system of equations (cf. Equations (17) and (39)):

$$\frac{\sum_k g_{1,k}\, B_k}{\sum_j B_j} = \langle g_1 \rangle$$

$$\vdots \tag{45}$$

$$\frac{\sum_k g_{\alpha,k}\, B_k}{\sum_j B_j} = \langle g_\alpha \rangle$$

Calculating the Shannon entropy with the probabilities given by Equation (43) yields (cf. Equations (19) and (40)):

$$H_{\mathrm{c}} = -\frac{\sum_k B_k \ln B_k}{\sum_j B_j} + \ln \sum_k B_k \tag{46}$$

According to Equation (42) we replace the right hand side of the first equation of system (45) with $N\langle f_1 \rangle$,

$$N\langle f_1 \rangle = \frac{\sum_k g_{1,k}\, B_k}{\sum_j B_j} \,, \tag{47}$$

and all other equations of system (45) with the according expressions. In the definition of the B_k we replace the exponents $g_{1,k}, g_{2,k}, \ldots, g_{\alpha,j}$ according to Equation (41) and insert the expression into Equation (47). The evaluation yields:

$$\langle f_1 \rangle = N \frac{\sum_i f_{1,i} \hat{A}_i}{\sum_j \hat{A}_j}$$

$$\vdots \tag{48}$$

$$\langle f_\alpha \rangle = N \frac{\sum_i f_{\alpha,i} \hat{A}_i}{\sum_j \hat{A}_j} \,,$$

with $\hat{A}_i$ defined similarly to A_i for the single system, Equation (38), but now with the factors $Y_1, Y_2, \ldots, Y_\alpha$:

$$\hat{A}_i = Y_1^{f_{1,i}} \cdot Y_2^{f_{2,i}} \cdot \ldots \cdot Y_\alpha^{f_{\alpha,i}}$$

The system of equations for $X_1, \ldots, X_\alpha$ for the single system, Equation (39), is the same as for the $Y_1, \ldots, Y_\alpha$ for the compound system, Equation (48), resulting in:

$$Y_1 = X_1$$
$$Y_2 = X_2$$
$$\vdots$$
$$Y_\alpha = X_\alpha$$

Now, replacing all Y-factors in the definition of the B_k, (44), with the according X-factors results in:

$$\sum_k B_k \ln B_k = N \left(\sum_i A_i \ln A_i\right) \left(\sum_i A_i\right)^{N-1}$$

$$\sum_k B_k = \left(\sum_i A_i\right)^N$$

$$\ln \sum_k B_k = N \ln \sum_i A_i$$

Inserting these expressions into the Shannon entropy of the compound system, Equation (46), we get the result:

$$H_\text{c} = N \left(-\frac{\sum_i A_i \ln A_i}{\sum_j A_j} + \ln \sum_i A_i\right)$$

and comparing with Equation (40) we have

$$H_\text{c} = N \, H_\text{s}$$

It is therefore proven that the Shannon entropy of compound systems underlying several constraints is homogeneous if the constraints behave according to Equations (41) and (42).

Equations (21) and (26) for systems underlying one constraint as well as Equations (41) and (42) for systems underlying several constraints reveal that homogeneity of Shannon entropy is guaranteed, if the constraints behave homogeneously, *i.*e., in a linear dependence of the number of subsystems:

$$\langle g(N)\rangle = N \langle f\rangle \quad \Rightarrow \quad \langle g(const.\cdot N)\rangle = const.\cdot N \langle f\rangle = const.\cdot \langle g(N)\rangle \tag{49}$$

We can therefore conclude that the assumption of both independent subsystems and of a compound system underlying a maximization principle with respect to additive constraints lead to the same important result: the homogeneity of the Shannon entropy of the compound system, hence the homogeneity of the modeled thermodynamic entropy. Both assumptions also act as prerequisites for homogeneity and can be regarded as complementary views of the same property of a compound system; neither of those aspects is preferred to the other.

3. Application to Thermodynamic Modeling of Fluid Phases

Amendatory to the previous sections where considerations were established for single and compound systems in general, in the following the application to thermodynamic modeling of fluid phases shall be discussed. These models describe the systems under consideration from the viewpoint of their possible states. For this purpose we reflect on the limiting cases of fluid phases, the ideal gas model and the condensed phase lattice system.

3.1. Ideal Gas

The ideal gas can be considered to be a compound system, consisting of a huge number of ideal and equal particles with no interactions among them. These particles are treated as the subsystems of the compound system. The mechanical state of a particle in the sense of classical mechanics is given by its position and velocity vectors, so the state is described by 6 random variables. The kinetic state of the ideal particle does not depend on its position and vice versa, so the position and velocity vectors are independent random variables. Consequently, as discussed in section 2.4, this leads to two statistically independent probability distributions, and the Shannon entropy splits in a kinetic and a positional term:

$$H = H_{\mathrm{kin}} + H_{\mathrm{pot}}$$

Therefore, maximizing H can be split by maximizing H_{kin} and H_{pot} separately. We restrict our considerations to the maximization of the kinetic term, which still comprises three random variables: one for the velocity and two for the direction of the movement. Assuming isotropy for the ideal gas means that the distribution of the kinetic energy does not depend on the direction of the movement. Hence we can again argue that the velocity is independent from the two other random variables, and we restrict our considerations to the kinetic states of the particles, defined only by the mean of their velocity, equivalent to their kinetic energy e. The kinetic energy is in fact a continuous variable. But in order to use the discrete formulation of Shannon entropy, Equation (1), we can discretize it by introducing an arbitrarily small energy quantum $\Delta\epsilon$:

$$e_i = i\,\Delta\epsilon\ .$$

The discretized energy e_i is meant only as a mathematical artifice. In the limit $\Delta\epsilon \to 0$ all possible continuous states can be represented. The kinetic, discrete state k of the ideal gas is then defined by the kinetic states of the particles: $k \,\widehat{=}\, i_1, i_2, \ldots, i_N$. Obviously, the kinetic energy of the whole system in the state k is the sum of the kinetic energies of the single particles:

$$E_k = E_{i_1,\ldots,i_N} = e_{i_1} + \ldots + e_{i_N}\ ,$$

cf. with Equations (21) and (41), and the additivity of the kinetic energy acting as constraint follows immediatley as

$$\langle E \rangle = N\langle e \rangle\ , \tag{50}$$

in accordance with Equations (26) and (42), and therefore additivity of the Shannon entropy of the kinetic term is guaranteed, cf. with Equation (36):

$$H_{\mathrm{kin,c}} = N H_{\mathrm{kin,s}}$$

This result is the basis for the discrete modeling of the ideal gas, to be presented in a subsequent paper.

3.2. Condensed Phase Lattice Systems

Lattice systems are mostly applied for strongly interacting condensed phases where molecular distances correspond to the liquid or solid state. Many engineering models used in process simulators

for chemical-engineering purposes such as activity coefficient models or equations of state are originally based on lattice models [13][14][15]. One of the reasons for this is that such models can easily be verified by Monte-Carlo simulations, alleviating model development and verification. Therefore, in the following the peculiarities of lattice systems shall be discussed from the viewpoint of Shannon entropy.

A lattice system provides fixed sites, each of which is occupied by one molecule in the simplest case, each of which interacts with its closest neighbors. In the following we apply the maximization presented in section 2 to an exemplary, one-dimensional lattice, comprising molecules of two types.

In terms of section 2, in the following the whole lattice is considered as the compound system, composed of sites which represent the subsystems.

3.2.1. The concept of subsystems applied to a lattice system

The simplest way to define a subsystem in terms of section 1.1 for a lattice is to use a single lattice site isolated from its adjacent neighbors. In a linear lattice system with two components such an isolated site has $2^1 = 2$ possible discrete states, as illustrated in Figure 1(a).

Figure 1. All possible discrete states of a single lattice site as *subsystem* of a linear lattice *system* when observed (a) isolated from its z nearest neighbors and (b) associated with its neighbors. In a linear lattice, considering only the nearest neighbors, $z = 2$.

(a)   (b)
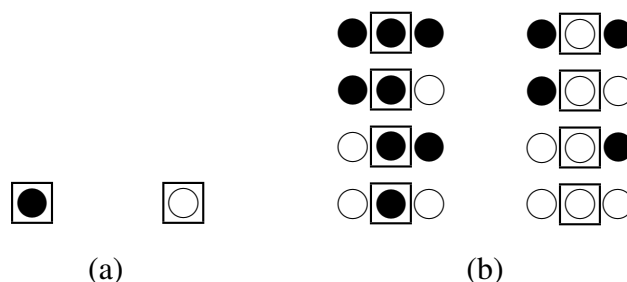
However, a concept of an isolated lattice which does not take its nearest, interacting neighbors into account does not allow for the formulation of constraints including interaction energies between sites, even though such constraints are essential for model development. For this reason, we introduce the concept of a single lattice site *associated* with its z nearest neighbors as subsystem, z representing the coordination number. The subsystem comprises $(z + 1)$ sites, as illustrated in Figure 1(b). In this concept, the discrete state of a lattice site is determined not only by its own molecule type but also by the type and arrangement of the z nearest neighbors contributing to energetic interactions. Here it is assumed that molecular interactions are confined to the nearest neighbors of the central molecule. If molecules beyond the direct neighbors also contribute to interactions, the concept of associated sites can be extended accordingly.

3.2.2. The unconstrained system

Shannon entropy of a subsystem: Without consideration of any constraints, it follows from equation (4a) that both of the two possible states shown in Figure 1(a) have the same probability, $p_{\bullet} = p_{\bigcirc} = \frac{1}{2}$,

corresponding to a system with an equal number of black and white sites. Equation (4b) for an isolated site with 2 possible states yields

$$H_{\text{s, iso}} = \ln 2 \tag{51}$$

For an associated site shown in Figure 1(b), consisting of $(z + 1)$ sites, again according to Equation (4a), all possible states are equally probable. The number of possible states is now 2^{z+1}, yielding the maximum Shannon entropy of

$$H_{\text{s, ass}} = \ln 2^{z+1} = (z + 1) \ln 2 , \tag{52}$$

which is the $(z + 1)$-fold of the Shannon entropy of the isolated site given by (51).

In a next step, the possibility of expressing the Shannon entropy of a compound lattice system by the Shannon entropy of its constituting subsystems shall be examined.

Shannon entropy of a compound system: When distributing two types of molecules over a lattice comprising N sites, the number of possible states is given by 2^N, yielding the maximum Shannon entropy of

$$H_{\text{c}} = \ln 2^N = N \ln 2 = N H_{\text{s, iso}} , \tag{53}$$

which is the N-fold of the maximum Shannon entropy of an isolated site, cf. Equation (51). When considering associated sites as subsystems, where a subsystem consists of $(z + 1)$ sites, the number of subsystems is

$$\hat{N} = \frac{N}{z + 1}$$

Now the number of states of the compound system is $2^{z+1^{\hat{N}}}$, and the according Shannon entropy yields

$$H_{\text{c}} = \ln 2^{z+1^{\hat{N}}} = \hat{N} \ln 2^{z+1} = \hat{N} H_{\text{s, ass}} , \tag{54}$$

which is the $\hat{N}$-fold of the maximum Shannon entropy of an associated site, cf. Equation (52).

Equations (52)-(54) reveal the homogeneity of the Shannon entropy of unconstrained lattice systems, cf. Equation (8):

$$H_{\text{c}} = N H_{\text{s}} ,$$

where H_{c} now denotes the Shannon entropy of the whole lattice as compound system, H_{s} is the Shannon entropy of the considered subsystem, *i.e.*, isolated or associated site, and N is the number of subsystems.

3.2.3. System considering constraints

There are basically three types of constraints to be considered in a lattice system: Energy, composition and the equivalence of contact pairs between molecules of different types.

Energy: As mentioned at the beginning of section 3.2, the intended purpose of lattice systems is the consideration of interaction energies between adjacent lattice sites. Therefore, the concept of a single lattice site associated with its nearest neighbors was introduced in section 3.2.1. to be used as subsystem. Based on this concept, constraints considering interaction energies can be formulated generically in the form

$$\langle u \rangle = \sum_{i=1}^{m} u_i \, p_i^{\text{ass}} \tag{55}$$

which is in line with (12), u_i designating the energy assigned to the central molecule of an associated site, p_i^{ass} its probability of residing in state i and $\langle u \rangle$ the mean value of energy. Figure 2 illustrates this nomenclature.

Figure 2. Examples for states and related energies of an associated lattice site as subsystem. ε denotes the interaction energy between two sites, where each site is assigned the half of it.

To formulate the energy assigned to a compound lattice consisting of N associated sites as subsystems, we use the index k to denote the state of the compound system which is determined by the states of its subsystems: $k \widehat{=} i_1, \ldots, i_N$. With this nomenclature, the energy of a compound system in state k is U_k, where

$$U_k = U_{i_1, \ldots, i_N} = u_{i_1} + \ldots + u_{i_N},$$

which is analogous to (21). Recalling equations (22) to (25) for the mean values of energies, it follows that

$$\langle U \rangle = \langle u_1 \rangle + \ldots + \langle u_N \rangle$$

Because all single subsystems are of the same kind and no single subsystem is preferred to another, their mean value is the same, resulting in

$$\langle u_1 \rangle = \langle u_2 \rangle = \ldots = \langle u_N \rangle \equiv \langle u \rangle$$
$$\langle U \rangle = N \langle u \rangle, \tag{56}$$

which is analogous to (26). Using (56) as the only constraint aside from the normalization of probabilities, maximization of the Shannon entropy analog to (13) requires solution of the Lagrange function

$$\mathcal{L} = H_{\text{c}} - \lambda_1 \langle U \rangle - \lambda_2 \sum_i \{p_i\} \overset{!}{=} max. \tag{57}$$

Application of the maximization principle to (57) in line with (27) - (34) finally results in

$$H_{\text{c}} = N \, H_{\text{s, ass}} \tag{58}$$

as Shannon entropy of the compound lattice system, analogously to Equation (36). Equation (58) reveals that the Shannon entropy of a constrained lattice system can also be expressed through the Shannon entropy of subsystems, whereupon subsystems are single lattice sites associated with their respective nearest neighbors. As shown in section 2.4, several functions with the generic form of equation (55) can also be considered as constraints in the maximization principle.

Composition: In the simplest case of a binary system, there are molecules of two types constituting the lattice. The constraint for the compound system is simply $N_1 = N \, x_1$, the total number of 1-molecules.

x_1 can be interpreted in two ways: as relative fraction of 1-molecules in the system, or as probability to find a 1-molecule at a given site. Hence, if we consider compound systems of identical composition, N_1 behaves according to Equation (49), fulfilling the prerequisites for constraints that enable homogeneity of Shannon entropy. The same holds for x_2 which is related to x_1 by the normalization condition. This can easily be extended to systems comprising an arbitrary number of components.

Equivalence of contact pairs: Contact pairs designate the number of contacts between molecules of different types in the system, e.g. N_{2-1} (read '2 around 1') the number of all 2-molecules around all molecules of type 1, and N_{1-2} the number of all 1-molecules around all molecules of type 2. As the number of contacts between molecules of different types must not depend on the viewpoint, the equivalence $N_{2-1} = N_{1-2}$ has to be fulfilled in lattice systems generally, independent of the respective size. The according constraint for the maximization prinziple is

$$N_{2-1} - N_{1-2} = 0$$

This equating to zero is a homogeneous function in terms of Equation (49).

In summary, all three types of constraints are homogeneous in terms of Equation (49), ensuring that after maximization the Shannon entropy of a compound system is also homogeneous. The complete Lagrange function finally comprises the mean values of energy, composition and equivalence of contact pairs as constraints to be considered in a lattice model. Practical application will be shown in a subsequent paper.

4. Conclusions

The scope of this paper was to clarify prerequisites for applying the concept of Shannon entropy and maximum entropy principle to thermodynamic modeling of extensive systems. The main criterion for applicability of this kind of modeling is the additivity of the Shannon entropy. It was shown that this additivity is guaranteed, provided that the additivity of the constraints is given. If a thermodynamic model comprises additive constraints, this prerequisite is fulfilled, and the method is explicitly applicable to systems of interacting components, *i*.e., real fluids. This was shown for two limiting cases of fluid phases, the ideal gas model and condensed phase lattice systems.

The main benefit of thermodynamic modeling based on Shannon entropy is that it makes the equilibrium distribution of discrete states available. This establishes new possibilities for thermodynamic and mass transport models as it allows consideration of a more detailed picture of physical behavior of matter on a molecular basis, beyond the scope of traditional modeling methods. This will be exploited in subsequent papers.

Acknowledgements

Author Contributions

M.P. contributed sections 2 and 3.1, T.W. contributed section 3.2, sections 1 and 4 were joint contributions of M.P. and T.W. The paper was read and complemented by A.P.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Shannon, C.E. A Mathematical Theory of Communication. *Bell Sys. Techn. Journ.* **1948**, *27*, 379–423,623–656.
2. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
3. Jaynes, E.T. Information Theory and Statistical Mechanics. II. *Phys. Rev.* **1957**, *108*, 171–189.
4. Jaynes, E.T. Gibbs vs Boltzmann Entropies. *Am. Journ. Phys.* **1965**, *33*, 391–398.
5. Wehrl, A. General Properties of Entropy. *Rev. Mod. Phys.* **1978**, *50*, 221–260.
6. Guiasu, S.; Shenitzer, A. The Principle of Maximum Entropy. *Math. Intell.* **1985**, *7*, 42–48.
7. Ben-Naim, A. *Entropy Demystified*; World Scientific Publishing, Singapore, 2008.
8. Ben-Naim, A. *A Farewell to Entropy: Statistical Thermodynamics Based in Information*; World Scientific Publishing, Singapore, 2008.
9. Curado, E.; Tsallis, C. Generalized Statistical Mechanics: Connection with Thermodynamics. *J. Phys. A* **1991**, *24*, L69–L72.
10. Tsallis, C. Nonadditive Entropy: The Concept and its Use. *Europ. Phys. Journ. A* **2009**, *40*, 257–266.
11. Maczek, A. *Statistical Thermodynamics*; Oxford University Press: Oxford, UK, 1998.
12. Wilde, D.J.; Beightler, C.S. *Foundations of Optimization*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1967.
13. Fowler, R.H.; Kapitza, P.; Mott, N.F.; Bullard, E.C. *Mixtures - The Theory of the Equilibrium Properties of Some Simple Classes of Mixtures, Solutions and Alloys*; Oxford at the Clarendon Press, 1952.
14. Abrams, D.S.; Prausnitz, J.M. Statistical Thermodynamics of Liquid Mixtures: A New Expression for the Excess Gibbs Energy of Partly or Completely Miscible Systems. *AIChE J.* **1975**, *21*, 116–128.
15. Bronneberg, R.; Pfennig, A. MOQUAC, a New Expression for the Excess Gibbs Energy based on Molecular Orientations. *Fluid Phase Equilib.* **2013**, *338*, 67–77.