A partial ordering of sets, making mean entropy monotone



2002 J. Phys. A: Math. Gen. 35 3163

(http://iopscience.iop.org/0305-4470/35/14/302)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 3163–3182 PII: S0305-4470(02)27099-7

# A partial ordering of sets, making mean entropy monotone

**Bernhard Baumgartner**

Institut für Theoretische Physik, Universität Wien, Boltzmanngasse 5, A-1090 Vienna, Austria

E-mail: baumgart@ap.univie.ac.at

Received 18 July 2001, in final form 4 February 2002
Published 29 March 2002
Online at stacks.iop.org/JPhysA/35/3163

**Abstract**

Consider a state of a system with several subsystems. The entropies of the reduced state on different subsystems obey certain inequalities, provided there is an equivalence relation, and a function measuring volumes or weights of subsystems. The entropy per unit volume or unit weight, the mean entropy, then decreases with respect to an order relation of the subsystems, defined in this paper. In the context of statistical mechanics a lattice system is studied in detail, and a decrease of mean energy is deduced for blow-up sequences of regular and irregular octagons. Examples referring to quantum information, with strict relations for mean entropies of subsystems in multipartite systems are also presented.

PACS numbers: 05.50.+q, 03.67.−a, 02.10.Ab

## 1. Introduction

*Entropy* is a key concept in several areas, from thermodynamics to dynamical systems. The study presented here refers in detail to systems in statistical mechanics, but it also offers links to its use in the context of information and quantum information theory. The problem studied in this paper, *comparing entropies of subsystems*, is presented in the context of statistical mechanics first, then it is stated in a formulation referring to (quantum) information theory. In both fields worked out examples are shown. The presentations are rather abstract, but yet not as general as could be possible. It turns out that the methods which are used are very flexible, so they may be applied also under different circumstances. This is remarked on at some places, especially in section 8. Also the logic of the patterns of definitions, axioms and deductions are analysed.

Consider a quantum *lattice system*, defined with a state on the inductive limit of local algebras of observables. Assume that a density matrix and an entropy $S(A)$ is assigned to each restriction of this state to any local algebra $\mathcal{A}_A$ belonging to a finite subset $A$ of the lattice. If the state is invariant under translations, and if $B$ is a '*translate*' of $A$ (meaning that there is a translation, mapping $A$ to $B$), we know that $S(B) = S(A)$. Now we are interested in relations between entropies assigned to different sets, with possibly different measures $\mu(A)$. For some time there has already been knowledge, see propositions 6.2.25 and 6.2.38 in [BR81], and [ABK01] with references therein, of a subadditivity, a strong subadditivity, a triangle inequality, and a 'strong triangle inequality' for entropies, and of the existence of the van Hove limit of the *mean entropy*, the entropy per unit volume (also called 'entropy density')

$$s(A) = S(A)/\mu(A). \tag{1}$$

All these assertions are true when the local algebras $\mathcal{A}_A$ are represented on local Hilbert spaces $\mathcal{H}_A$ with the product property

$$\mathcal{H}_D = \mathcal{H}_A \otimes \mathcal{H}_C, \qquad \text{if} \quad D = A \cup C \quad \text{and} \quad A \cap C = \emptyset, \tag{2}$$

and the density matrices fulfil the compatibility condition

$$\rho_A = \mathrm{Tr}_C\, \rho_D, \qquad \text{if} \quad D = A \cup C \quad \text{and} \quad A \cap C = \emptyset, \tag{3}$$

where $\rho_A$ represents the state on the local algebra $\mathcal{A}_A$. See [LR73a] for a listing of fundamental properties of entropy related to the product property of algebras, from which all the above assertions follow. See also [W78] for properties of entropy, containing the first exact considerations on monotonicity in volume in translation invariant systems.

Then it has been established for mean entropy in [ABK98] that $s(A) \geqslant s(B)$, if $A$ and $B$ are rectangular boxes or parallelepipeds with $A \subset B$. In the same electronic 'paper' this rule of monotone decrease has been extended to some specific simple shapes which go beyond the box and parallelepiped shapes, and the question was raised, whether this relation of monotonicity can be extended to *all* pairs of sets with $A \subset B$. This 'question 1' of [ABK98] has been answered by me in e-mail discussions in the negative, constructing counterexamples to the two candidate inequalities mentioned there. In the appendix of the present paper I give counterexamples close to physical applications. The sets of the 'candidate inequalities' are presented in (54).

The printed paper [ABK01] is actually a revision of [ABK98], which (amongst other important things) takes the counterexamples into account, and contains a new 'question 1': for which pairs of regions A, B in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ satisfying $A \subset B$, is it necessarily the case that $s(A) \geqslant s(B)$? Now we give a positive answer, and we extend the monotonicity property of $s(A)$ from rectangular boxes to more general volumes.

To enable this quest we define an *order relation between sets*, adapted to imply monotone decrease of mean entropy, as is stated in theorem 5. The formalism of this method is not tied to lattice systems. All we need is an *equivalence relation* between sets, corresponding to an invariance property of the state, as for example translation invariance. To define the mean entropy we need, of course, a measure. Then we will need nothing more but the strong subadditivity of the entropy as a function of sets,

$$S(A \cup B) \leqslant S(A) + S(B) - S(A \cap B), \tag{4}$$

which is a consequence of the strong subadditivity proven in [LR73b], and of the product property (2) together with the compatibility (3). This is also true for fermionic systems, see [AM01]. We remark that we do not need positivity of the entropy. So our results are also true for classical continuous systems.

In the context of *quantum information* as in [NC00] and in [B98], consider information *stored* in an assembly of several elementary quantum systems. Consider also subsystems of the total system, defined by subsets of the total set of elementary systems. Each subsystem $A$, and the information it carries, is represented on a Hilbert space $\mathcal{H}_A$ with a density matrix $\rho_A$, obeying the product property (2) and the compatibility condition (3). We remark that entropy in such systems is related to measures of entanglement [B98, B01].

We assume that to any subsystem $A$ a weight $\mu(A) > 0$ is assigned. Typically, it may be the number of qubits. These weights fulfil the conditions for a measure on the family of subsets $\{A, B, \ldots\}$, that is: $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$. Now we are interested not only in the von Neumann entropy $S(A)$, but also in the *mean entropy*, the entropy per unit weight, defined in (1). When can we be sure that $s(A) \geqslant s(B)$? This question can now be answered with the help of the strong subadditivity (4) and the order relation defined in section 2, using nothing more but a knowledge about pairs of subsets, for which $s(A') = s(A)$ holds with certainty. So we assume, moreover, that this system is equipped with an equivalence relation on the family of subsets, which is compatible with the assignment of a mean entropy, so that '$A'$ equivalent to $A$' implies $s(A') = s(A)$. It is not necessarily compatible with the assignment of a weight, but it will be so in the examples. Typically, the equivalence relations are consequences of symmetries of the system and its state.

The definition of partial ordering, $A \prec B$, is given in axiomatic form, with axioms defining the process of how to construct the order relation step by step. We take some care also of the logic in this process; examples demonstrate both the use of these axioms, using subsets of a two-dimensional lattice, and their mutual logical independence, that is, that no axiom is a consequence of the others.

Each step in the creation of such an order relation is consistent with *monotonicity* of the mean entropy. This fact is a consequence of strong subadditivity (4); it is proven in section 4 and stated in theorem 5: if the equivalence of two sets implies that they have the same mean entropy, then mean entropy is monotone non-increasing with respect to the partial ordering.

We present an application of this method in the context of statistical mechanics, creating an order relation for convex octagons in $\mathbb{Z}^2$. Their ordering is related to simple inequalities for the pieces of the boundaries, as stated in definition 6 and in theorem 7. It implies the monotone decrease of mean entropy, and it includes the ordering of *blow-up* sequences of octagons.

Then three examples of a multipartite four qubit system with entangled pure states are presented. Three *different* prescribed symmetries imply *different* order relations. With specific entangled states we demonstrate the correctness of both the predicted decrease and the possible non-decrease of mean entropy in the increasing sequences of subsystems.

We also mention other fields, where entropies of subsystems are of interest, and we remark on several inequalities for average values of mean entropy.

## 2. The partial ordering

Consider a *space* (set) $\mathcal{M}$, with a family $\mathcal{F}$ of subsets which is closed under forming unions and intersections of pairs, and an equivalence relation $\sim$ on $\mathcal{F}$. A partial ordering $\prec$ on $\mathcal{F}$ is defined by the following *axioms*:

(I) if $A \prec B$, $A \sim A'$ and $B \sim B'$, then also $A' \prec B$ and $A \prec B'$ hold (*invariance*);
(II) if $A \cap B$ is empty, and if either $A \sim B$ or $A \prec B$ holds, then $A \prec D = A \cup B$;
(III) if $A \cap B = C \prec B$, and if either $A \sim B$ or $A \prec B$ holds, then $A \prec D = A \cup B$;
(IV) if $A \cap B = C \prec D = A \cup B$, and if either $A \sim B$ or $A \prec B$ holds, then $A \prec D$;
(V) if $B \cap C = A \prec B$ and also $A \prec C$ hold, then $A \prec D = B \cup C$, even if $B$ is not comparable with $C$ (i.e. neither $B \sim C$, nor $B \prec C$, nor $C \prec B$ hold);
(VI) if $A \prec B$ and $B \prec D$, then also $A \prec D$ holds (*transitivity*).

To enable a precise referencing, we subdivide the second, third and fourth axioms:

(II) (a) $A \cap B = \emptyset$ and $A \sim B \Rightarrow A \prec D = A \cup B$

(b) $A \cap B = \emptyset$ and $A \prec B \Rightarrow A \prec D = A \cup B$

(III) (a) $A \cap B = C \prec B$ and $A \sim B \Rightarrow A \prec D = A \cup B$

(b) $A \cap B = C \prec B$ and $A \prec B \Rightarrow A \prec D = A \cup B$

(IV) (a) $A \cap B = C \prec D = A \cup B$ and $A \sim B \Rightarrow A \prec D$

(b) $A \cap B = C \prec D = A \cup B$ and $A \prec B \Rightarrow A \prec D$.

The condition of uncomparability is included in the fifth axiom to keep the axioms mutually independent, but this condition can be ignored in applications:

**Lemma 1.** *Let $A = B \cap C$, $A \prec B$, $A \prec C$ and either $B \sim C$ or $B \prec C$. Then $A \prec D = B \cup C$.*

**Proof.** Use axioms IIIa and VI if $B \sim C$, axioms IIIb and VI if $B \prec C$. □

On the other hand, allowing equivalence of $B$ with $C$ in axiom V would make the order relation stated in axiom IIIa a consequence of axioms I, V and VI.

There are, in principle, two different ways of using these axioms. One possibility would be to 'import' an order relation between sets, 'dictating' it, and then extending it. In this paper, however, we are concerned with the other way of starting from scratch, *creating* the ordering out of the equivalence relation, by forming unions and checking for order relations concerning the intersections. We state an immediate consequence of using this creative process:

**Lemma 2.** *If the ordering is created in a countable number of steps, then $A \prec D$ implies that there exist sets $A^{(1)} \cdots A^{(N)}$ which are equivalent to $A$, such that*

$$D = \bigcup_{n=1}^{N} A^{(n)}. \tag{5}$$

**Proof.** We may order the sequence of logical steps leading to $A \prec D$ in such a way, that in the $n$th step, establishing $\ldots \prec B^{(n)}$, only the relations established in the former steps are needed. Going backwards from $A \prec D$, we see that $D$ and each $B^{(n)}$ has to be represented as the union of two sets, $B^{(\nu)}$ and $B^{(\mu)}$, which are either equivalent to $A$ or are ordered like $A \prec B^{(\nu)}$. Then we proceed by induction. The basis of the creative process is the axiom IIa. This is the only one which does not need a pair of sets which are already in an order relation. So $B^{(1)} = A^{(1)} \cup A^{(2)}$. We assume that $A \prec B^{(\nu)}$ for $\nu < n - 1$ has been established. Now either a similar procedure as in step one may occur in step $n$, or the existence of one or two of the former relations is needed. Checking the axioms it is obvious that the property (5) propagates from $B^{(\nu)}$ to $B^{(n)}$ in the process of establishing $A \prec B^{(n)}$. □

It is to be remarked that (5) is not a sufficient condition for $A \prec D$. In the appendix we present a counterexample, combined with $s(A) \ngeq s(D)$.

The following may in applications apply to the empty set:

**Corollary 3.** *If a set $Z$ is not equivalent to any other set, then no relation $Z \prec D$ can be established.*

## 3. Logical independence of the axioms and examples of applications

The axiom of invariance has a special status, just enabling full power for the other axioms in creating the order relation. Especially, axioms IIb and IVb would have no meaning at all without this invariance. Without it, order relations can only be established between sets where one of them is a subset of the other. This can be seen by inspection of the 'creative' axioms II–V. So the independence of this axiom is obvious.

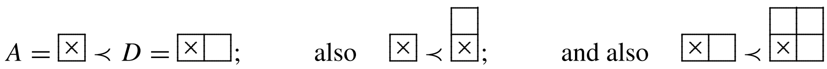
Now we demonstrate the independence of the other axioms II to VI—with one exception—on examples for each one of them, where an ordered pair would *not* be comparable without it, i.e. we present each time an order relation $A \prec D$, which cannot be created without this axiom. The exception is axiom IIb. If the ordering is created in a countable number of steps, starting with axiom IIa, then the assertion of axiom IIb is a consequence of axiom V in its generalized form. We state this strange fact as lemma 4. Nevertheless we present two simple examples of how to use axiom IIb.

**Lemma 4 (deduction of IIb).** *If the ordering is created as stated above by axioms I . . . VI, the assertion of axiom IIb is a consequence of the other axioms.*

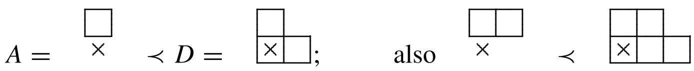
**Proof.** Axiom IIb assumes $A \prec B$ and $A \cap B = \emptyset$. By lemma 2, $B$ is a union of sets $A^{(n)}$ which are all equivalent to $A$. Now define $C = A \cup A^{(1)}$, observe $A \prec C$ by IIa, and use axiom V. Note that lemma 1 allows us to ignore the existence or non-existence of relations between $B$ and $C$. □

*Examples*

In our system the two-dimensional lattice $\mathbb{Z}^2$ is the space $\mathcal{M}$, the finite subsets of $\mathbb{Z}^2$ form the family $\mathcal{F}$, and the discrete group of two-dimensional translations maps each set $A$ to all equivalent sets $A' \sim A$. To make the notation short, we denote those points of $\mathbb{Z}^2$ which we need like the squares of a chessboard, i.e. $a1$ instead of (0, 0), $c4$ instead of (2, 3) etc. In the pictorial presentation of these examples the points are represented by squares, and the point $a1$ (which may also be outside $A$) is marked with a cross.

(IIa) 

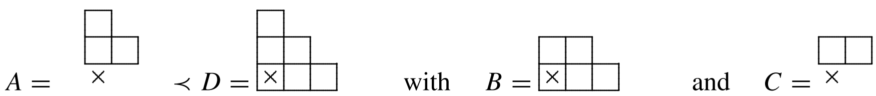
$A = \{a1\} \sim B = \{b1\}$, so
$A \prec D = A \cup B = \{a1, b1\}$;
also $\{a1\} \prec \{a1, a2\}$; and also $\{a1, b1\} \prec \{a1, a2, b1, b2\}$.

(IIb) 

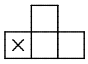
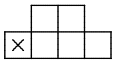
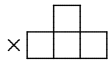
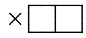
$A = \{a2\} \prec D$
$D = \{a1, b1, a2\}$,
with $B = \{a1, b1\}$,
and $A \prec B$ by the order created above and the invariance property.
Also $A = \{a2, b2\} \prec D = \{a1, b1, c1, a2, b2\}$, with $B = \{a1, b1, c1\}$ and the ordering $A \prec B$ created in the following example.

(IIIa) 

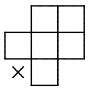
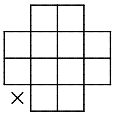
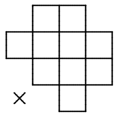
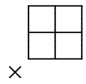
$A = \{a1, b1\} \prec D$
$D = \{a1, b1, c1\}$, with
$B = \{b1, c1\}$, $C = \{b1\}$,
and with the order $C \prec A \sim B$ created with IIa.

(IIIb) 

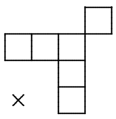
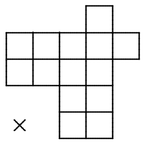
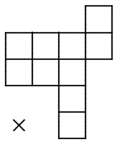
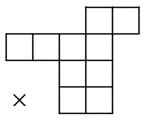
$A = \{a1, b1, a2\} \prec D$
$D = \{a1, b1, c1, a2, b2, a3\}$, with
$B = \{a1, b1, c1, a2, b2\}$,
$C = \{a2, b2\}$,
and the order $C \prec B$ created with IIb and IIIa.

(IVa) $A =$  $\prec D =$  with $B =$  and $C =$ 

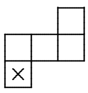
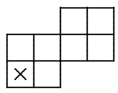
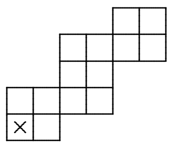
$A = \{a1, b1, c1, b2\} \prec D$
$D = \{a1, b1, c1, d1, b2, c2\}$, with
$B = \{b1, c1, d1, c2\}$, $\quad C = \{b1, c1\}$.
The order $C \prec D$ can easily be seen to follow from IIa and IIb.

(IVb) $A =$  $\prec D =$  with $B =$  and $C =$ 

$A = \{a2, b1, b2, b3, c2, c3\} \prec D$ with
$D = \{a2, a3, b1, b2, b3, b4, c1, c2, c3, c4, d2, d3\}$,
$B = \{a3, b2, b3, b4, c1, c2, c3, c4, d2, d3\}$,
$C = \{b2, b3, c2, c3\}$.
Relation $A \prec B$ is a consequence of axiom IIIa, using $B = A' \cup A''$, where $A' = \{b2, c1 \ldots\}$ and also $A'' = \{a3, b2 \ldots\}$ are translates of $A$, and $A' \cap A'' = \{b2, c3\} \prec A$ by applications of IIa and IIb. The relation $C \prec D$ can be proven with IIa and IIIa.

(V) $A =$  $\prec D =$  with $B =$  and $C =$ 

$A = \{a3, b3, c1, c2, c3, d4\} \prec D$ with
$D = \{a3, a4, b3, b4, c1, c2, c3, c4, d1 \ldots d5, e4\}$,
$B = \{a3, a4, b3, b4, c1 \ldots c4, d4, d5\}$,
$C = \{a3, b3, c1, c2, c3, d1 \ldots d4, e4\}$.
Relation $A \prec B$ is a consequence of axiom IVa, using $B = A \cup A'$, where $A' = \{a4, b4 \ldots\}$ is a translate of $A$, and $A \cap A' = \{c2, c3\} \prec B$ by repeated application of IIa and IIb. The analogue procedure proves $A \prec C$.

(VI) $A =$  $\prec$ $B =$  $\prec$ $D =$ 

$A = \{a1, a2, b2, c2, c3\} \prec D$, with
$D = \{a1, a2, b1, b2, c2, c3, c4, d2, d3, d4, e4, e5, f4, f5\}$,
and with the intermediate set
$B = \{a1, a2, b1, b2, c2, c3, d2, d3\}$.
The relation $A \prec B$ is established with IVa, because of $B = A \cup A'$, where $A' = \{b1, b2 \ldots\}$ is a translate of $A$, and $A \cap A' \prec B$. Relation $B \prec D$ is established with either IIIa or IVa, because of $D = B \cup B'$, where $B' = \{c3, c4 \ldots\}$ is a translate of $B$, and $B \cap B' \prec B$ or $B \cap B' \prec D$.

To see the necessity of the axioms, their logical independence, one has to see, in each case, that the order relation does not follow from the application of another axiom instead. Here lemma 2 helps. It is easy in each example to find out all the sets $A'$, $A''$, ..., which are equivalent to $A$ and subsets of $D$, and then to test all the chains $A \prec ?A' \cup A'' \prec ? \ldots \prec ?D$. For the examples to IIIb and IVb there are several ordered chains, but one always has to use the special axiom which is just under inspection.

## 4. Monotonicity of mean entropy

**Theorem 5 (monotone decrease).** *Assume that equivalent sets have the same mean entropy. Then mean entropy is monotone decreasing:*

$$A \prec D \Rightarrow s(A) \geqslant s(D). \tag{6}$$

**Proof.** We write the strong subadditivity of the entropy as a relation for the mean entropy:

$$s(A \cup B) \leqslant \lambda_A s(A) + \lambda_B s(B) - \lambda_C s(C) \tag{7}$$

where we denote $A \cap B = C$, as in the axioms II, III and IV, and use

$$\lambda_A = \frac{\mu(A)}{\mu(A \cup B)}, \qquad \lambda_B = \frac{\mu(B)}{\mu(A \cup B)}, \qquad \lambda_C = \frac{\mu(C)}{\mu(A \cup B)}. \tag{8}$$

Note that $0 \leqslant \lambda_C \leqslant \min\{\lambda_A, \lambda_B\} \leqslant 1$ and $\lambda_A + \lambda_B = 1 + \lambda_C$. In the case $A \cap B = \emptyset$, we define $\lambda_C s(C = \emptyset) = 0$, in accordance with the simple subadditivity of the entropy (which would be sufficient to check axiom II),

$$S(A \cup B) \leqslant S(A) + S(B). \tag{9}$$

Now we proceed in an inductive way along the process of creating the ordering, similar as we did in the proof of lemma 4. We check each axiom, and assume that the inequality (6) holds for the pairs of sets which are assumed to be in an order relation already.

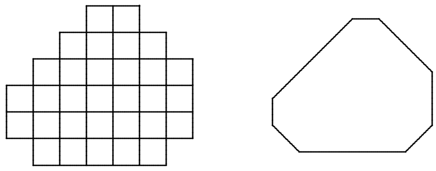
(I) Equivalent sets have the same mean entropy, so the inequalities are invariant.
(II) We have $\lambda_A + \lambda_B = 1$, $s(B) \leqslant s(A)$, hence $s(D) \leqslant \lambda_A s(A) + \lambda_B s(B) \leqslant s(A)$.
(III) Now $s(C) \geqslant s(B)$, $s(B) \leqslant s(A)$, so $s(D) \leqslant \lambda_A s(A) + \lambda_B s(B) - \lambda_C s(C) \leqslant \lambda_A s(A) + (\lambda_B - \lambda_C) s(B) \leqslant (\lambda_A + \lambda_B - \lambda_C) s(A) = s(A)$.
(IV) Here $s(C) \geqslant s(D)$, implying $s(D) \leqslant \lambda_A s(A) + \lambda_B s(B) - \lambda_C s(D) \Rightarrow s(D) \leqslant (\lambda_A s(A) + \lambda_B s(B))/(1 + \lambda_C)$, and $s(B) \leqslant s(A) \Rightarrow s(D) \leqslant s(A)$.
(V) Here we have $s(D = B \cup C) \leqslant \lambda_B s(B) + \lambda_C s(C) - \lambda_A s(A)$, $s(B) \leqslant s(A)$, $s(C) \leqslant s(A)$ and $\lambda_B + \lambda_C = 1 + \lambda_A$, which obviously implies $s(D) \leqslant s(A)$.
(VI) Transitivity is obvious.

□

It occurs as a surprise that we had not to assume equivalent sets to have the same measure, as it is naturally the case in all our examples.

## 5. Octagonal sets in a two-dimensional lattice

We extend the monotonicity of mean entropy, known for rectangular boxes, to the family $\mathcal{O}$ of convex octagonal sets. $\mathcal{O}$ is defined as consisting of all those finite convex subsets of $\mathbb{Z}^2$, whose boundaries are horizontal, vertical and diagonal lines, with some exceptions: oblique rectangles are excluded. (The necessity of this exclusion is demonstrated at the end of the appendix.) But oblique straight lines are elements of $\mathcal{O}$. We characterize each set $A$ in $\mathcal{O}$ by the coordinates of the most left point in the lowest line, and by a sequence of eight non-negative integers, $(m, n, p, \ldots, u)$, which describe its boundary $\boldsymbol{b}_A$. The length of the lower horizontal boundary is $m$, the other boundary lines follow counterclockwise, with the lengths $n\sqrt{2}$, $p, \ldots, u\sqrt{2}$. The following drawings show a set $A$, each point in $\mathbb{Z}^2$ represented by a square, and its boundary $\boldsymbol{b}_A = (4, 1, 2, 2, 1, 3, 1, 1)$:



Note that the boundary lines are included in the set.

Since the boundary is a closed curve, the parameters have to obey

$$u + m + n = q + r + s, \tag{10}$$

$$n + p + q = s + t + u, \tag{11}$$

and for *each* sequence of eight non-negative numbers obeying these relations there exists a boundary. The length of any boundary line (or of more boundary lines) may be zero. Then two (or more) corners, otherwise at the endpoints of this boundary line, coalesce into one. Octagons which are 'lines' have two pieces of boundary (representing two opposite directions of the boundary curve). A special boundary is $\boldsymbol{b}_E = (0, 0, \ldots, 0)$, characterizing elementary sets, the 'atoms'.

Again we define a set $A'$ to be equivalent to $A$, if and only if $A'$ is a translate of $A$. An equivalence class of octagons is obviously characterized by the boundary, which is common to all of the members of the equivalence class. Moreover, and to show this is the goal of this section, the ordering of the octagons corresponds to an ordering of the boundaries.

**Definition 6 (ordering of boundaries).** *Boundary $\boldsymbol{b}_A = (m_A, \ldots, u_A)$ is said to be piecewise shorter than $\boldsymbol{b}_B = (m_B, \ldots, u_B)$, denoted as $\boldsymbol{b}_A < \boldsymbol{b}_B$, if $\boldsymbol{b}_A \neq \boldsymbol{b}_B$ and $m_A \leqslant m_B, n_A \leqslant n_B, \ldots, u_A \leqslant u_B$.*

**Theorem 7 (ordering of octagons).** *Two octagons, $A$ and $D$, elements of $\mathcal{O}$, are in the order relation $A \prec D$, if and only if $\boldsymbol{b}_A < \boldsymbol{b}_D$.*

This ordering includes *blow-up sequences*: let $A_\nu$ be characterized by the boundary $\nu\boldsymbol{b}_A = (\nu m, \nu n, \ldots, \nu u)$, with $\nu$ a positive integer. Then $\nu < \pi$ implies $A_\nu < A_\pi$.

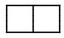
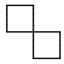
**Proof.** One direction, assuming $A \prec D$, is easy. By lemma 2, $D$ is a union of some translates of $A$. So each boundary line of $D$ contains the corresponding boundary line of at least one of these translated sets and thus cannot be shorter. The proof of the assertion in the other logical direction fills the remaining part of this section. □

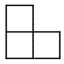
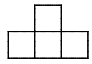
Consider the circumference of a boundary

$$\ell(D) = m_D + p_D + r_D + t_D + \sqrt{2}(n_D + q_D + s_D + u_D). \tag{12}$$

Each finite interval in $\mathbb{R}$ contains only a finite number of such lengths, so they can all be numbered and ordered as $\ell_1 < \ell_2 < \cdots < \ell_N < \cdots$. The proof will be by *induction on the circumference* $\ell(D)$, that is, by induction on N. Since $\ell(E) = 0$ holds for the atoms only, with no other boundaries piecewise shorter than $\boldsymbol{b}_E$ existing, there is nothing to prove at the start of the induction. The induction hypothesis is that $\boldsymbol{b}_A < \boldsymbol{b}_B$ and $\ell(B) < \ell(D)$ imply already $A \prec B$. Now, if $\boldsymbol{b}_A < \boldsymbol{b}_D$, and if, concerning the ordering, there are other boundaries in between, $\boldsymbol{b}_A < \cdots \boldsymbol{b}_B < \boldsymbol{b}_D$, then $\ell(B) < \ell(D)$, and, because of the transitivity, it remains to prove $B \prec D$ only. In other words, we assume that there are no other boundaries between $\boldsymbol{b}_A$ and $\boldsymbol{b}_D$.

Since the linear relations (10) and (11) hold for both $\boldsymbol{b}_A$ and $\boldsymbol{b}_D$, they hold also for the piecewise difference $\boldsymbol{b}_D - \boldsymbol{b}_A = (m_D - m_A, \ldots)$, and this difference can be interpreted as boundary $\boldsymbol{b}_M$ of a '*molecule*' $M$, a set which is in the ordering one level directly above the atoms, with no other set in between. It is not difficult to find out that there are exactly twelve equivalence classes of molecules in $\mathcal{O}$, which can be grouped into four types. Representatives of these types are:

$M_1 = \{a1, b1\} =$ , $M_2 = \{a2, b1\} =$ ,

$M_3 = \{a1, a2, b1\} =$ , $M_4 = \{a1, b1, b2, c1\} =$ .

Each of these sets represents an equivalence class and, together with the equivalence classes defined by its one rotated version or three rotated versions, it represents a type.

There is a relation between the sets $M_i$ and the process of constructing $D$ out of $A$.

**Definition 8 (convolution of sets).** *Let $A = \{\boldsymbol{a}_i, 1 \leqslant i \leqslant I\}$, $C = \{\boldsymbol{c}_j, 1 \leqslant j \leqslant J\}$, then the convolution of A with C is*

$$A * C = \{\boldsymbol{a}_i + \boldsymbol{c}_j, 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J\}. \tag{13}$$

See also the remark following the conjecture 12 in the next section.

**Lemma 9 (convolution is addition).** *If A and M are octagonal sets in $\mathcal{O}$, but not orthogonal oblique lines, then $D = A * M \in \mathcal{O}$ and $\boldsymbol{b}_D = \boldsymbol{b}_A + \boldsymbol{b}_M$.*

**Proof.** It is not difficult to see that $D$ is a convex set. (This would not be true if $A$ and $M$ were oblique orthogonal lines, so we excluded this case; and we have no oblique rectangular boxes in $\mathcal{O}$.) Consider a boundary line $\{\boldsymbol{a}_i, i \in \text{Line}_A\}$ of $A$ and the corresponding boundary line $\{\boldsymbol{x}_j, j \in \text{Line}_M\}$ of $M$, for example, the lower horizontal boundary lines, with lengths $m_A$ and $m_M$. (If its length happens to be zero, it consists of one point, for example, the lowest point in the set.) Then $\{\boldsymbol{a}_i + \boldsymbol{x}_j, i \in \text{Line}_A, j \in \text{Line}_M\}$ is the corresponding boundary line of $A * M$, and its length is $m_A + m_M$. □

The convolution of $A$ with an atom produces a translate of $A$,

$$A * \{\boldsymbol{x}_i\} = A^{(i)} \sim A, \tag{14}$$

so, for $M = \bigcup_i \{\boldsymbol{x}_i\}$,

$$A * M = \bigcup_i A^{(i)}. \tag{15}$$

Now we are ready to apply the axioms and perform the step of induction.

**Proposition 10.** *Let $\boldsymbol{b}_A < \boldsymbol{b}_D$, and $A * M_i = D$, for $A \in \mathcal{O}$ and one of the twelve molecules $M_i$. Assume that $\boldsymbol{b}_C < \boldsymbol{b}_B$ and $\ell(B) < \ell(D)$ together imply $C \prec B$. Then $D \in \mathcal{O}$ and $A \prec D$ holds.*

**Proof.** If $A$ is an atom, then $A * M_i = M_i$ or $A * M_i \sim M_i$, and $A \prec M_i$ by axiom IIa.

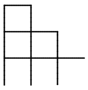
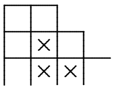
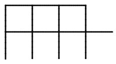
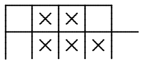
For larger sets we have to treat different cases differently, but since the system is invariant under 90° rotations, it is sufficient to prove this proposition with the molecules $M_1 \dots M_4$. With $(m, n, \dots, u)$ we denote the boundary of $A$. Assume that the point $a1$ is the origin.

*Convolution with $M_1$*:

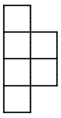
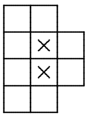
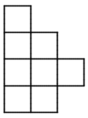
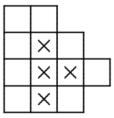
$A * M_1$ is the union of $A$ with $A'$, where $A'$ is $A$ translated by one step to the right. We have to investigate whether $C = A \cap A'$ is empty or whether it can be compared with $A$ or with $D$. This depends on the form of $A$ in the lowest and in the highest region: if $A$ is not a vertical or oblique line, we say that the bottom of $A$ is

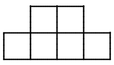
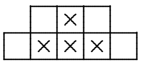
◇ *flat*, if $m \geqslant 1$;
◇ *sharp*, if $m = 0$, and either $u = 0$ and $n \geqslant 1$ or vice versa;
◇ *rectangular*, if $m = 0$, $u \geqslant 1$ and $n \geqslant 1$.

Analogously we denote the top, with $u, m, n$ replaced by $q, r, s$. The following pictures show the sharp top of some $A$ and the corresponding top of $D$ with $C$ inside marked with crosses, and the same for some $A$ with a flat top.

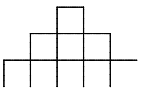
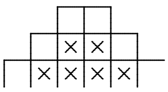
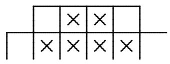
$A_{sharp} =$ , $D =$ ; $A_{flat} =$ , $D =$ . (16)

- *A is a vertical or oblique line.* $A \cap A' = \emptyset$ and $A \prec D$ by axiom IIa.
- *Top and bottom are sharp.* $A$ is either a parallelepiped, or an equilateral triangle, where the vertical line is the 'base', or a trapezoid, also with the vertical line as the base. We have $m_C = 0$ and also $r_C = 0$, all the nonvanishing adjacent border lines of $C$ are shorter than the corresponding border lines of $A$, so $C$ can be compared with $A$: $\boldsymbol{b}_C < \boldsymbol{b}_A$, which implies $C \prec A$. With axiom IIIa we infer $A \prec D$.

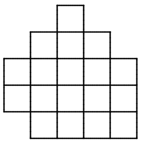
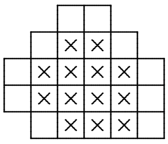
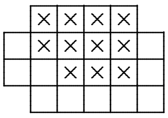
  Example: $A =$  $\prec D =$ , with $C$ inside.
- *Bottom is flat, top is sharp, or vice versa.* If $m \geqslant 1$, then $m_C = m - 1$; the top of $C$ is sharp with adjacent border lines shorter than those at $A$, as in the above case. Other border lines of $C$ have the same lengths as the corresponding lines of $A$. So $C \prec A$, and we again use axiom IIIa. If $m = 0$ but $r \geqslant 1$, the same argument is true, with top and bottom exchanged. Example: $A =$  $\prec D =$ , with $C$ inside.
- *Top and bottom are flat.* (Horizontal lines and rectangles are here included.) Both $m_C = m - 1$ and $r_C = r - 1$ hold, with the other border lines of $C$ with the same lengths as those of $A$. So again $C \prec A$ holds and axiom IIIa is applicable.

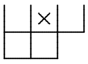
  Example: $A =$  $\prec D =$ , with $C$ inside.
  Now, if $m = r = 1$ and $p = t = 0$, the set $C$ is an oblique rectangle, not in $\mathcal{O}$. It is nevertheless true that $C \prec A$, as is shown under the item 'Top and bottom are rectangular'.

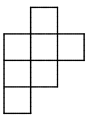
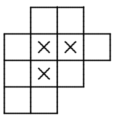
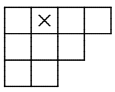
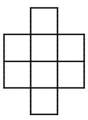
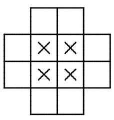
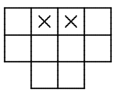
If one of the extremal regions is rectangular, $C$ cannot be compared with $A$. Either $m_C = 1$ while $m = 0$, or $r_C = 1$ while $r = 0$, or both. The picture shows the rectangular top of some $A$, the top of $D$ with $C$ inside, $C$ marked with crosses, and the top of a set $F$ which is needed in the proof.

$A_{rectangular} =$ , $D =$ , $F =$ . (17)

So, to prove $A \prec D$ one has to use axiom IVa, and establish the ordering $C \prec D$ first. In any case one has to establish $C \prec D$ to construct a set $F \in \mathcal{O}$, with $F \subset D$, $\ell(F) < \ell(D)$, and $\boldsymbol{b}_C < \boldsymbol{b}_F$, so that $C \prec F$ by assumption (the induction hypothesis to prove the theorem). Then one has to use a set $C' \sim C$, such that $D = C' \cup F$ and $G = C' \cap F \prec F$ (or $G = \emptyset$), which implies $C \prec D$ by axiom IIIb (or by axiom IIb).

- *Top is rectangular, bottom flat (or vice versa).* $F$ is the union $C^- \cup C \cup C^+$, where $C^-$ is $C$ translated one unit to the left, $C^+$ is $C$ translated one unit to the right. So $F = C * \{z1, a1, b1\}$, with $z1$ denoting the point left of $a1$. We discuss the case with the rectangular top, the other case is analogous by reflection. We move $C$ one unit upwards, denote this set as $C'$, and observe $D = F \cup C'$, see (17). The simplest case is just that of the example to axiom IVa in section 3. In this case $F \cap C' = \emptyset$, so axiom IIb makes the proof of $C \prec D$ complete. For any larger $A$, $C$ is also larger and $G = F \cap C' \neq \emptyset$. By inspection of (17) one sees that $r_G = r_F = 3$, $q_G = q_C - 1$, $s_G = s_C - 1$. The other border lines have the same lengths as those of $C$. Since $\boldsymbol{b}_C < \boldsymbol{b}_F$, we have $\boldsymbol{b}_G < \boldsymbol{b}_F$, and $G \prec F$.
  Example:
  $A =$  $\prec D(C \text{ inside}) =$ , $F =$ , $G$ inside.
- *Top is rectangular, bottom sharp (or vice versa).* $F = C * H$, where $H = \{z0, z1, a0, a1, b1\}$ ($a0$ denotes the point below $a1$, etc) is equivalent to the bottom of $D$:

  $\mathit{bottom}(D) =$  (18)

  $F \in \mathcal{O}$ since $C \in \mathcal{O}$ and $H \in \mathcal{O}$. The remaining procedure is as above: we construct $C' = C * \{a2\}$, $G = C' \cap F$, so that $D = C' \cup F$, and observe $r_G = r_F = 3$, $q_G = q_C - 1$, $s_G = s_C - 1$. The other border lines have the same lengths as those of $C$. Since $\boldsymbol{b}_C < \boldsymbol{b}_F$, we have $\boldsymbol{b}_G < \boldsymbol{b}_F$, and $G \prec F$. This implies $C \prec D$.
  Example:
  $A =$  $\prec D(C \text{ inside}) =$ , $F =$ , $G$ inside.
- *Top and bottom are rectangular.* The bottom of $D$ with $C$ inside is like the bottom in (17) upside down. $H$ is defined accordingly as $H = \{z1, a1, b1, a0\} \in \mathcal{O}$. It follows that $F = C * H \in \mathcal{O}$, $C \prec F$. Again we use $C' = C * \{a2\}$, and perform the same procedures as in the last two cases.
  Example:
  $A =$  $\prec D(C \text{ inside}) =$ , $F =$ , $G$ inside.

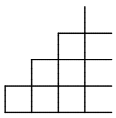
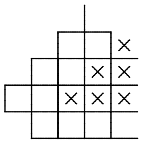
Since it is needed at another place, we observe that the same procedure can be applied also if $A$ is an oblique rectangle, because also in this case $C \in \mathcal{O}$ and $D \in \mathcal{O}$.

*Convolution with $M_2$*:

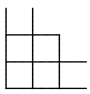
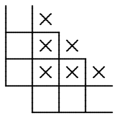
$A * M_2$ is the union of $A' = A * \{b1\}$ with $A'' = A * \{a2\}$. The case $A' \cap A'' = \emptyset$ occurs if $A$ is a horizontal or a vertical line, or a narrow oblique strip perpendicular to $M_2$, where the border of $A$ has $q = u = 0$, and either $m = 0$ and $t = 1$ or $m = 1$ and $t = 0$. In these cases axiom IIa applies. Note that using oblique lines rectangular to $M_2$ is forbidden (see the remark in the proof of lemma 9).

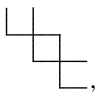
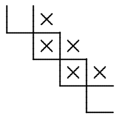
If $C = A' \cap A''$ is not the empty set, it is always a precursor of $A$ in the ordering. To see this, one has to check the lower left end and the upper right end of $A$. We define the types for the lower left end, with obvious analogues for the upper right ends by symmetry. The pictures show again each type of end for some $A$, and the corresponding end of $D$ with $C$ inside marked with crosses.

- *Sharp ends.* Either $m \geqslant 2$, $s \geqslant 1$, $t = u = 0$, so that $m_C = m - 1$, $t_C = u_C = 0$,

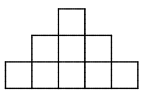
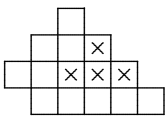
  $A_{sharp} =$ , $D =$ , (19)

  or $t \geqslant 2$, $n \geqslant 1$, $m = u = 0$, and $t_C = t - 1$, $m_C = u_C = 0$.
  The length $s_C$ (or $n_C$) is also shorter than $s$ (or $n$); how much shorter it is depends also on the other end.
- *Rectangular ends.* $u = 0$, $m \geqslant 1$, $t \geqslant 1$, so $u_C = 0$, $m_C = m - 1$, $t_C = t - 1$.

  $A_{rectangular} =$ , $D =$ . (20)
- *Flat ends.* $u \geqslant 1$, so $u_C = u - 1$, $m_C = m$, $t_C = t$.

  $A_{flat} =$ , $D =$ . (21)

If no end is sharp, both $n_C = n$ and $s_C = s$. Otherwise at least one of these border lines is shorter than those of $A$. In any case, $\boldsymbol{b}_C < \boldsymbol{b}_A$, so $C \prec A$, and axiom IIIa applies, proving $A \prec D$.

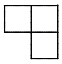
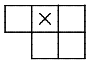
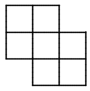
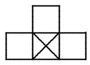
Example: $A =$  $\prec D =$ , $C$ inside.

*Convolution with $M_3$*:

- *Lines:* If $A$ is a vertical line, $(A * \{b1\}) \cap (A * \{a1, a2\}) = \emptyset$, and axiom IIb applies, using $D = (A * \{b1\}) \cup (A * \{a1, a2\})$ and $A \prec (A * \{a1, a2\})$. For lines in other directions, the same argument applies, with the three points from $M_3$ appearing in appropriate permutations.

In the other cases, the types of the bottom and of the left side of $A$ dictate the way of proof, whether axiom IIIb can be used, or whether IVb is necessary. The types of the left side are defined in the same way as the types of the bottom, rotated by 90°: *Flat* means $t \geqslant 1$, *rectangular* means $t = 0$, $m \geqslant 1$, $u \geqslant 1$, and *sharp* means $t = 0$ with either $s = 0$, $r \geqslant 1$, $u \geqslant 1$, or $u = 0$, $m \geqslant 1$, $s \geqslant 1$.

- *Left side is flat.* $t \geqslant 1$. Define $B = A * M_1$. By assumption $A \prec B$, since $\ell(B) < \ell(D)$. Observe $D = A' \cup B$, with $A' = A * \{a2\}$. Then $C = A' \cap B \prec B$ has to be shown. If $s \geqslant 1$, then $s_C = s - 1$, otherwise $t_C = t - 1$. On the right side. If $q \geqslant 1$, then $q_C = q - 1$; if $q = 0$, $p \geqslant 1$, then $p_C = p - 1$; if $p = q = 0$, then $n \geqslant 1$, $n_C = n - 1$. In all cases, $r_C \leqslant r$, so $\boldsymbol{b}_C \leqslant \boldsymbol{b}_B$ and $C \prec B$. By symmetry of the system and of the molecule $M_3$ under reflections at a diagonal, the proof for the cases with flat bottom, $m \geqslant 1$, is analogous. The example for the axiom IIIb is an example for this case. (What there has been termed $A$ is here $A'$.)
- *Left side is sharp.* Since the cases with $m \geqslant 1$ can be treated in the way described above, we assume $m = 0$. Also the lines have already been considered, so we have here $u \geqslant 1$, $s = 0$, $r \geqslant 1$. Also here $B = A * M_1$, $A'$ and $C$ defined as above. On the left side we observe $u_C = u - 1$, on the right side the inequalities for different cases are as above. So again $C \prec B$ holds, and axiom IIIb applies.

  Example: $A =$  $\prec B =$ , with $C$ inside, $D =$  .
- *Left side and bottom are rectangular.* It is necessary to use $B = A * M_2$. Also in this case $A \prec B$. We have to investigate on $C = A \cap B$. The left side and the bottom are: $t_C = 1$, $m_C = 1$, $u_C = u - 1$, $s_C = s - 1$, $n_C = n - 1$. The other border lines have the same lengths as in $A$. Form $F = C * H$, with $H = \{z1, a1, a2, b1\} =$ . (The point $\{a1\}$ is marked with a cross.) Represent $D = (C * \{a2\}) \cup F$, observe $C \prec F$. So, by axiom IIIb, $C \prec D$, and axiom IVb implies $A \prec D$. The example for axiom IVb is an example for this case.

*Convolution with $M_4$*:

Define $B = A * \{z1, a1, b1\}$, $A' = A * \{a2\}$. So $D = A' \cup B$. If $A$ is a horizontal line, $A' \cap B = \emptyset$ and axiom IIb applies. For the other cases it is easy to see $C = A' \cap B \prec B$: observe $r_C \leqslant r_B$. For the left side, consider the sequence $s, t, u$ of border lengths in $A$. The first one, which is not zero, is shorter by one in the set $C$. The analogue is true for the right side, considering $q, p, n$ and $q_C, p_C, n_C$. Obviously $m_C = m$. So $C \prec B$, and axiom IIIb applies. □

## 6. Remarks on continuum systems

Consider a translation invariant state on the inductive limit of local algebras $\{\mathcal{A}_A\}$ on bounded convex measurable subsets $\{A\}$ of $\mathbb{R}^2$. Assume that the system fulfils the product property (2) and the compatibility condition (3). Impose moreover a continuity condition, which enables an approximation of an octagon $A$ in $\mathbb{R}^2$ and the state restricted to $\{\mathcal{A}_A\}$ by sets $A_\nu$ and the states restricted to them, where the $A_\nu$ are unions of small elementary squares, as shown in the figures of section 5 in this paper. Then the statement of theorem 7 can be carried over to the octagons in $\mathbb{R}^2$. One might moreover think of an extension to higher dimensions, establishing an ordering of some sets in $\mathbb{Z}^d$ and $\mathbb{R}^d$.

But these possible results are somehow unsatisfactory. Why just the octagons? If the monotonicity of mean entropy is true for sequences of octagons it is probably true for all kinds of convex sets. So I state now a conjecture (related to 'question 2' of [ABK01]):

**Conjecture 11 (blow-up sequences).** *Consider a system with all the properties stated above. Let A be a convex bounded measurable set in* $\mathbb{R}^d$*, and define* $\lambda A = \{\lambda \boldsymbol{x}; \boldsymbol{x} \in A, \lambda > 0\}$*. Then* $s(\lambda_1 A) \geqslant s(\lambda_2 A)$ *if* $\lambda_1 < \lambda_2$*.*

The route to a proof of this statement might involve a stronger property:

**Conjecture 12 (ordering convex sets).** *Consider a system with all the properties stated above. Let A and B be convex bounded measurable sets in* $\mathbb{R}^d$*, and define*

$$D = A * B = \{\boldsymbol{x} + \boldsymbol{y}; \boldsymbol{x} \in A, \boldsymbol{y} \in B\}, \tag{22}$$

*then mean entropy is decreasing, in the sense that* $s(A) \geqslant s(D)$ *and* $s(B) \geqslant s(D)$*.*

I remark that $A * B$, defined here as 'convolution of sets', is simply denoted as $A + B$ in mathematical investigations on convex sets and polytopes, and known as 'Minkowski sum'.

## 7. A multipartite system with quantum information

Consider Alice and Bob, Cassandra and Diomedes, each one in the possession of one qubit. The total system of four qubits is in a pure state. We consider different examples of symmetry for this system, and construct the order relations of the subsystems. With special states we demonstrate both the decrease of mean entropy, as it is predicted by theorem 5, and the possibilities of an increase, when this theorem does not apply.

In *case 1* we assume two 'reflection' symmetries. One reflection exchanges all the nationalities, while keeping the sexes:

$$A \leftrightarrow C, \qquad B \leftrightarrow D. \tag{23}$$

The other symmetry operation exchanges simultaneously all the sexes, while keeping the cultural connections:

$$A \leftrightarrow B, \qquad C \leftrightarrow D. \tag{24}$$

Then there is of course the 'product' of these two generating operations:

$$A \leftrightarrow D, \qquad B \leftrightarrow C. \tag{25}$$

These symmetry operations create an equivalence relation in an obvious way: two sets are equivalent, if there exists a symmetry operation, transforming the one into the other as, for example, $\{A, B\} \leftrightarrow \{C, D\}$. So all the subsets of one party, $\{A\}, \{B\}, \ldots$, to each of which we ascribe the measure $\mu = 1$, are equivalent. $\{A\} \sim \{B\} \sim \ldots$. Also all the subsets of three, with $\mu = 3$, form one equivalence class. But the subsets of two, with $\mu = 2$, form three distinct classes:

$$\{A, B\} \sim \{C, D\}, \qquad \{A, C\} \sim \{B, D\}, \qquad \{A, D\} \sim \{B, C\}. \tag{26}$$

Because of axiom IIa, sets such as $\{A\}$ precede all the larger sets, and, by the same axiom, each set of two precedes the full set $\{A, B, C, D\}$. Axiom IVa implies an order relation between the sets of three and the full set, such as $\{A, B, C\} \prec \{A, B, C, D\}$. So there are ordered sequences such as

$$\{A\} \prec \{A, B\} \prec \{A, B, C, D\}, \tag{27}$$

$$\{A\} \prec \{A, B, C\} \prec \{A, B, C, D\}. \tag{28}$$

Equivalent sequences can be found by exchanging equivalent sets. But no other order relations exist, no set of two can be compared with a set of three.

Now we present pure entangled states with this prescribed symmetry: the pair $\{A, B\}$ is in the EPR–Bohm–Bell (EPR–B–B) state $(|00\rangle + |11\rangle)/\sqrt{2}$. The pair $\{C, D\}$ is in the same state, and the product state of these two entangled states gives the pure state for the entire system. The mean entropies are, represented by one representative for each class,

$$\begin{aligned} &s(\{A\}) = \ln 2, \\ &s(\{A, B\}) = 0, \qquad s(\{A, C\}) = s(\{A, D\}) = (1/2)\ln 2, \\ &s(\{A, B, C\}) = (1/3)\ln 2, \\ &s(\{A, B, C, D\}) = 0. \end{aligned} \tag{29}$$

As another pure entangled state with the same symmetry we may choose the product of EPR–B–B states for the pairs $\{A, C\}$ and $\{B, D\}$. And still another one with $\{A, D\}$ and $\{B, C\}$ as the entangled pairs. Each of these states gives mean entropy zero to one class of pairs, the class of the entangled pairs, *lower* than the mean entropy of the sets of three. Comparisons between the other subsets show the *decrease* of mean entropy as predicted by theorem 5.

In *case 2* we consider the same set of four qubits, but another equivalence relation, this time connected with the symmetry under 'rotation',

$$A \to B, \qquad B \to C, \qquad C \to D, \qquad D \to A. \tag{30}$$

Again there is one equivalence class of the single parties, and one class with all the sets of three. The parties of two now form two classes,

$$\{A, B\} \sim \{B, C\} \sim \{C, D\} \sim \{A, D\}, \tag{31}$$

$$\{A, C\} \sim \{B, D\}. \tag{32}$$

The ordering relations of case 1 hold also in this case, for the same reasons. As a new relation comes in this case

$$\{A, B\} \prec \{A, B, C\}, \tag{33}$$

due to an application of axiom IIIa. But the set $\{A, C\}$ is in no relation to the sets of three. As an entangled pure state having this symmetry we consider the product state of the EPR–B–B states of the pairs $\{A, C\}$ and $\{B, D\}$. The mean entropies of pairs are

$$\begin{aligned} &s(\{A, B\}) = s(\{B, C\}) = \cdots = (1/2)\ln 2, \\ &s(\{A, C\}) = s(\{B, D\}) = 0. \end{aligned} \tag{34}$$

The mean entropies of the other sets are the same as in (29). We see a *decrease* of mean entropy as prescribed by theorem 5, and an *increase* when enlarging the set $\{A, C\}$ to $\{A, B, C\}$, in accordance with the fact that there is no ordering relation between these two sets.

In *case 3* we consider the four qubits totally symmetric under permutations. Now all the pairs are equivalent to each other, in addition to the equivalence relations of the case above. With this high symmetry the order relation is total, instead of only partial as before:

$$\{A\} \prec \{A, B\} \prec \{A, B, C\} \prec \{A, B, C, D\}, \tag{35}$$

where each set may be replaced by another one with the same number of qubits. None of the states considered before shows this symmetry. But the GHZ-state $(|0000\rangle + |1111\rangle)/\sqrt{2}$ does, and mean entropy is decreasing, as it has to be, predicted by theorem 5:

$$\begin{aligned} &s(\{A\}) = \cdots = \ln 2 > s(\{A, B\}) = \cdots = (1/2)\ln 2 \\ &> s(\{A, B, C\}) = \cdots = (1/3)\ln 2 > s(\{A, B, C, D\}) = 0. \end{aligned} \tag{36}$$

## 8. Discussion

In this paper we present an investigation on *qualitative* and *strict* comparison of mean entropies of systems. Most interesting is here the comparison of a system $B$ with a subsystem $A$. There are, contrary to first guess, cases with an increase $s(B) > s(A)$, as is shown in the appendix. So we established conditions on the systems definitely implying *non-increase* $s(B) \leqslant s(A)$, Other existing investigations show that an increase is in some sense rather exceptional.

In [P93, S96] the formula for the *unitary average* of entropy

$$\langle S(A)\rangle = \sum_{k=N/m+1}^{N} \frac{1}{k} - \frac{m(m-1)}{2N} \tag{37}$$

was found, related to a quantum system $A$ which is the subsystem of an $N$-dimensional system $C$ in a pure state, where $m = \dim(A)$ is bounded as $m \leqslant \sqrt{N}$. 'Dimension' of the system denotes the dimension of the associated Hilbert space, and $N/\dim(A)$ is assumed to be an integer. The average is here formed over all unitary transformations of the large system. No geometry, no set-relations and no equivalence relations are involved. Since the system $A'$ complementary to $A$, that is $\mathcal{H}_C = \mathcal{H}_A \otimes \mathcal{H}_{A'}$, is in a state with the the same entropy as $A$, we extend this formula, dropping the inequality on $m$, defining $m = \min\{\dim(A), N/\dim(A)\}$ instead. The logarithm of the dimension can serve as a measure,

$$\mu(A) = \ln(\dim(A)), \tag{38}$$

in accordance with the usual association of the 'union' of systems with the tensor product of the Hilbert spaces. Analysing the functions in (37) we find

**Corollary 13 (decrease of unitarily averaged mean entropy).** *Consider a large system in a pure quantum state. The average mean entropy of a subsystem A,*

$$\mu(A)^{-1}\langle S(A)\rangle \tag{39}$$

*is a decreasing function of the dimension of A.*

Another form of averaging is studied in [ABK01] by considering set-relations, but still without any symmetry: Let $E_1 \ldots E_N$ be disjoint sets, all with the same measure, which may be chosen as $\mu(E_k) = 1$. Consider the class of $N! - 1$ different sets formed by the unions of some $E_k$, with associated local Hilbert spaces and compatible states. Now form the *subset-average* of entropy, forming the average over the class of sets with measure $n$ (unions of $n$ different $E_k$),

$$S_n = \binom{N}{n}^{-1} \sum_{A,\mu(A)=n} S(A). \tag{40}$$

Theorem 3 of [ABK01] states that $(N-1)^{-1} S_{N-1} \leqslant N^{-1} S_N$. A simple combinatorial argument leads to

**Corollary 14 (decrease of the subset-average).** *Consider a large system associated with a large set composed of N disjoint 'elements' $E_k$, each with measure* 1. *The subset-average of mean entropy $S_n/n$ of all subsets composed of n elements is a non-increasing function of n.*

The monotonicity of mean entropy may find applications in studies on the thermodynamic limit. In this paper it has already been proven, as an extension of theorems on sequences of increasing parallelepipeds, that the mean entropies of a blow-up sequence of octagons in a two-dimensional lattice is monotonically non-increasing, and conjectures on further extensions have been stated. More ambitious would be an implementation into theories of *themodynamic* *depth*, [LP88], *statistical complexity*, *structure* and *patterns* [FC98, CF01, SC01]. The studies in these fields have up to now been concerned mainly with time sequences of information. But there are relations to studies on the thermodynamic limit, and an extension to theories on spatio-temporal patterns is predicted as a future direction, see [CF01].

## Appendix. Negative answers to a question of A Kay and B Kay

The system is again a two-dimensional lattice. To avoid a discussion of the thermodynamic limit we consider some large box, with periodic boundary conditions for the interaction. The local Hilbert space for a single 'elementary' site is $\mathbb{C}^2$. The system is now actually a classical discrete system, since we use only commuting operators: for each site $\alpha$ we use only the unit operator $\mathbb{I}$ and

$$\sigma_\alpha \cong \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{41}$$

Then we use products:

$$\sigma_A = \left(\bigotimes_{\alpha\in A} \sigma_\alpha\right) \otimes \left(\bigotimes_{\alpha\notin A} \mathbb{I}\right). \tag{42}$$

Since the system is essentially a classical system, each state of such a system may be reinterpreted as a double periodic state of the infinite system on the lattice $\mathbb{Z}^2$: any state is a probability measure on the set of configurations. Each configuration of the box can infinitely often be repeated periodically to give a configuration of the infinite system. Transferring a probability measure, which is invariant under translations of the box = torus, to this family of periodic configurations gives a translation invariant state of the infinite system.

We stay close to possible applications in physics, considering states defined by short-range interactions, symmetric under translations. The states in the large box $G$ are represented by density matrices

$$\rho = \mathrm{e}^{-\beta H} / \operatorname{Tr} \mathrm{e}^{-\beta H}, \tag{43}$$

$H$ being a Hamiltonian acting in $\mathcal{H}_G$,

$$H = \sum_{A\in\mathcal{G}} \sigma_A, \tag{44}$$

where $\mathcal{G}$ denotes any family of sets which is invariant under the symmetry group of periodic translations in the box $G$. This family may also be invariant under reflections and 90° rotations, if desired. The empty set is no element of $\mathcal{G}$.

To enable a simple discussion of the restriction to a smaller set $D$, we consider very high temperatures, i.e. small $\beta$. The partition function $\operatorname{Tr} \mathrm{e}^{-\beta H}$ and all the expectation values $\operatorname{Tr} \sigma_B \rho$ are analytic functions of $\beta$.

$$Z(\beta) = \operatorname{Tr} \mathrm{e}^{-\beta H} = Z(0) + \mathrm{O}(\beta^2) \qquad \text{with } Z(0) = 2^{\mu(G)} \tag{45}$$

$$\operatorname{Tr} \sigma_B \rho = -Z(0)^{-1} \beta \operatorname{Tr} \sigma_B H + \mathrm{O}(\beta^2) \qquad \text{for } B \neq \emptyset. \tag{46}$$

Note that $\operatorname{Tr} \sigma_B \sigma_A = 1$ if $B = A$, zero otherwise. Therefore, from the Hamiltonian only the $\sigma_A$ with sets $A$ being subsets of $D$ contribute to the expectation values of $\sigma_B$ to first order in $\beta$, when $B \subset D$. So the restriction of the state to the region $D$ can be represented by a density matrix

$$\rho_D = \mathrm{e}^{-\beta H_D} / \operatorname{Tr} \mathrm{e}^{-\beta H_D}, \tag{47}$$

with a Hamiltonian $H_D$ acting in $\mathcal{H}_D$,

$$H_D = \sum_{A \in \mathcal{G}(D)} \sigma_A + \tilde{H}_D, \tag{48}$$

with $\|\tilde{H}_D\| = \mathrm{O}(\beta)$, and where the family of interaction-multiplets is denoted as

$$\mathcal{G}(D) = \{A \in \mathcal{G}, A \subset D\}. \tag{49}$$

Now we can expand the mean entropy. The leading order is the same, $s_0 = \ln 2$, for each set. The terms of order $\beta$ cancel, so we are interested in the terms of second order.

$$\begin{aligned} s(D) &= -\mu(D)^{-1} \operatorname{Tr} \rho_D \ln \rho_D \\ &= \mu(D)^{-1} (\ln \operatorname{Tr} \mathrm{e}^{-\beta H_D} + \beta \operatorname{Tr}[H_D \mathrm{e}^{-\beta H_D}] / \operatorname{Tr} \mathrm{e}^{-\beta H_D}) \\ &= \mu(D)^{-1} (\ln \operatorname{Tr} \mathbb{I} - \tfrac{1}{2}\beta^2 \operatorname{Tr} H_D^2 / \operatorname{Tr} \mathrm{e}^{-\beta H_D}) + \mathrm{O}(\beta^3) \\ &= s_0 - \tfrac{1}{2}\beta^2 \mu(D)^{-1} \operatorname{Tr}\Bigg[\sum_{A \in \mathcal{G}(D)} \sigma_A\Bigg]^2 \Bigg/ \operatorname{Tr} \mathbb{I} + \mathrm{O}(\beta^3). \end{aligned} \tag{50}$$

Observe that $\operatorname{Tr} \sigma_A \sigma_{A'} = 0$ for $A \neq A'$, and that $\sigma_A^2 = \mathbb{I}$. So the calculation of the entropy $S(D)$ up to order $\beta^2$ amounts to counting how many interacting pairs or multiplets of sites are contained in the set $D$, that is, the number of sets in $\mathcal{G}(D)$, which we denote as

$$n(D) = |\mathcal{G}(D)|. \tag{51}$$

Mean entropy is thus

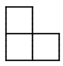
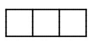
$$s(D) = s_0 - \tfrac{1}{2}\beta^2 n(D)/\mu(D) + \mathrm{O}(\beta^3). \tag{52}$$

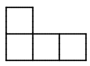
And we summarize the considerations as

**Lemma 15 (high-temperature states).** *Let $B$ and $D$ be finite subsets of $\mathbb{Z}^2$, their measures denoted as $\mu(B)$ and $\mu(D)$. Consider the lattice spin system with a translational invariant state defined as in (43) and (44). Let $s(B)$ and $s(D)$ denote the mean entropies, $n(B)$ and $n(D)$ the numbers of covered interaction multiplets, as defined in (49) and (51). Then there exists $\beta_C > 0$, such that for each $\beta \in (0, \beta_C)$ the following implication is true:*

$$n(B)/\mu(B) > n(D)/\mu(D) \Longrightarrow s(B) < s(D). \tag{53}$$

Now we are ready to present some examples and counterexamples. First we give a negative answer to 'question 1' of [ABK98]. Consider the sets

$B = \{a1, a2, b1\} =$ , $\qquad C = \{a1, b1, c1\} =$ ,

$D = \{a1, a2, b1, c1\} =$ , (54)

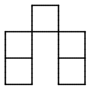
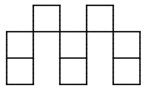
where $B \subset D, C \subset D$, but $B \not\prec D, C \not\prec D$. In comparing $B$ with $D$, we consider an interaction of diagonally nearest neighbours, $\mathcal{G}$ consisting of $\{a1, b2\}$,$\{a2, b1\}$ and its translates. Since $n(B) = n(D) = 1$, we have $s(B) < s(D)$.

In comparing $C$ with $D$, consider $\mathcal{G}$ as consisting of triples like $C$ itself and its translates, and possibly also of the rotated triples and its translates. Since $n(C) = n(D) = 1$, we have $s(C) < s(D)$.

The last example can easily be generalized to many cases, where a larger set does *not* have lower mean entropy: if $C \subset D$, but no other set $C'$ equivalent to $C$ is a subset of $D$, consider $\mathcal{G} = \{C' \sim C\}$ as the set of interacting multiplets and the interaction Hamiltonian defined as in (44).

We may now extend the set of questions posed in [ABK01] and ask:
*Question:* Considering two sets $C \subset D$, is it *necessary* that $D$ can be represented as a union of sets which are equivalent to $C$, to guarantee $s(C) \geqslant s(D)$ for any invariant state?

This property $D$ being a union of sets equivalent to $C$, is, alas, not sufficient. Here is the counterexample promised after the proof of lemma 2. Consider the sets

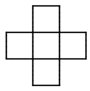
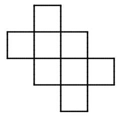
$C = \{a1, a2, b3, c1, c2\} =$ , and $D = C \cup C' =$ ,

with $C' = \{c1, c2, d3, e1, e2\}$ being a translate of $C$. $\mathcal{G}$ consists of pairs of nearest neighbours, $A = \{a1, a2\}$ and its translates, possibly also of the rotated pairs, $B = \{a1, b1\}$ and its translates. The comparison gives

$$\frac{n(C)}{\mu(C)} = \frac{2}{5} > \frac{n(D)}{\mu(D)} = \frac{3}{8}. \tag{55}$$

So $s(C) < s(D)$ for small $\beta \neq 0$. As it has to be, because of theorem 5, there is also no order relation between $C$ and $D$. This comes from the intersection $C \cap C'$ being in no order relation, neither with $C$, nor with $D$. Moreover, this example can be extended. Translating $C$ again and again, always the same distances, one can form chains of different length, $D_N = C \cup C' \cup \ldots \cup C^{(N)}$. With the same method, calculating $n(D^{(N)})/\mu(D^{(N)})$, one can prove that $s(D^{(N)})$ is *increasing* in $N$.

With a state, which is *not* rotation invariant, it can be demonstrated, that oblique rectangles cannot be ordered in the same way as the other orthogonal sets. Consider

$C = \{a2, b1, b2, b3, c2\} =$ , and $D = C \cup C' =$ ,

with $C' = \{z2, a1, a2, a3, b2\}$, a translate of $C$.
$\mathcal{G}$ consists of pairs of diagonal neighbours, $A = \{a1, b2\}$ and its translates. The comparison gives again (55) and $s(C) < s(D)$ for small $\beta \neq 0$.

## Acknowledgments

I thank Amanda Kay and Bernard Kay for stimulating e-mail discussions, Hajime Moriya for telling me about the work in progress of H Araki and himself, and also for pointing out some misprints in the first draft. I thank the referee for hints to other fields of interest, and I thank Elliott H Lieb, Heide Narnhofer, Helmuth Urbantke and Jacob Yngvason for interest in this study.

## References

[BR81] Bratteli O and Robinson D W 1981 *Operator Algebras and Quantum Statistical Mechanics* vol 2 (Berlin: Springer)

[LR73a] Lieb E H and Ruskai M B 1973 A fundamental property of quantum-mechanical entropy *Phys. Rev. Lett.* **30** 434–6

[W78] Wehrl A 1978 General properties of entropy *Rev. Mod. Phys.* **50** 221–60

[ABK98] Kay A R and Kay B S 1998 Monotonicity with volume of entropy and of mean entropy for translationally invariant systems as consequences of strong subadditivity *Preprint* arXiv math-ph/9812008

[ABK01] Kay A R and Kay B S 2001 Monotonicity with volume of entropy and of mean entropy for translationally invariant systems as consequences of strong subadditivity printed version *J. Phys. A: Math. Gen.* **34** 365–82

[LR73b] Lieb E H and Ruskai M B 1973 Proof of the strong subadditivity of quantum-mechanical entropy *J. Math. Phys.* **14** 1938–41

[AM01] Araki H and Moriya H 2001 personal communication (paper in preparation)

[NC00] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

[B98] Bennett C H 1998 Quantum information *Phys. Scr.* T **76** 210–17

[B01] Bennett C H *et al* 2001 Exact and asymptotic measures of multipartite pure state entanglement *Phys. Rev.* A **63** 012307/1–12

[P93] Page D N 1993 Average entropy of a subsystem *Phys. Rev. Lett.* **71** 1291–4

[S96] Sen S 1996 Average entropy of a subsystem *Phys. Rev. Lett.* **77** 1–3

[LP88] Lloyd S and Pagels H 1988 Complexity as thermodynamic depth *Ann. Phys.* **188** 186–213

[FC98] Feldman D P and Crutchfield J P 1998 Measures of statistical complexity: Why? *Phys. Lett.* A **238** 244–52

[CF01] Crutchfield J P and Feldman D P 2001 Regularities unseen, randomness observed: levels of entropy convergence *Preprint* arXiv cond-mat/0102181

[SC01] Shalizi C R and Crutchfield J P 2001 Computational mechanics: pattern and prediction, structure and simplicity *J. Stat. Phys.* **104** 817–79