Monotonicity with volume of entropy and of mean entropy for translationally invariant systems as consequences of strong subadditivity

2001 J. Phys. A: Math. Gen. 34 365

(http://iopscience.iop.org/0305-4470/34/3/304)

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

J. Phys. A: Math. Gen. **34** (2001) 365–382 www.iop.org/Journals/ja PII: S0305-4470(01)08377-9

Monotonicity with volume of entropy and of mean entropy for translationally invariant systems as consequences of strong subadditivity

Amanda R Kay and Bernard S Kay

Department of Mathematics, University of York, York YO10 5DD, UK

E-mail: bsk2@york.ac.uk

Received 4 October 1999, in final form 17 October 2000

Abstract

We consider some questions concerning the monotonicity properties of entropy and of mean entropy for translationally invariant states on translationally invariant systems (classical lattice, quantum lattice and quantum continuous). By taking the property of *strong subadditivity*, which for quantum systems was proven rather late in the historical development, as one of four primary axioms (the other three being simply positivity, subadditivity and translational invariance), we are able to obtain results, some new, some proved in a new way, which appear to complement in an interesting way results proved around 30 years ago on limiting mean entropy and related questions. In particular, we prove that as the sizes of boxes in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ increase in the sense of set inclusion, (1) their mean entropy decreases monotonically and (2) their entropy increases monotonically. We include a proof of (2) based on the notion of *m-point correlation entropies*, which we introduce and which generalize the notion of *index of correlation* (see e.g. Horodecki R 1994 *Phys. Lett.* A **187** 145). We mention a number of further results and questions concerning monotonicity of entropy and of mean entropy for more general shapes than boxes and for more general translationally invariant (/homogeneous) lattices and spaces than $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$. We also obtain some further results on monotonicity of entropy in these more general situations by adjoining a fifth axiom, which embodies yet another general property of entropy (which we call the '*strong triangle inequality*').

PACS numbers: 0367, 0365, 0570

1. Introduction

This paper is concerned with some general monotonicity properties of entropy and of mean entropy for translationally invariant states on translationally invariant infinite Euclidean lattice and continuum quantum (and, in the lattice case, also classical) systems, and with some possible generalizations of these properties to other homogeneous lattices and continua.

To explain the sorts of question in which we are interested here, and the approach we take to answering them, it will be useful to first recall what was learned about the related question of the existence of *limiting mean entropy* in the mid-1960s. A simple variant of this question (see corollary 1 below) is the question whether the mean entropy of a box in a Euclidean lattice or continuum tends to a definite limit as the lengths of each of its sides tend to infinity. Here, by the *mean entropy* of a (finite) box, we simply mean its entropy divided by its volume. (The reader should be warned that in the literature referred to here, no particular phrase is attached to this concept, and the term 'mean entropy' is used instead to denote what we call here 'limiting mean entropy'.) This variant had been proven to be true both for classical systems by Robinson and Ruelle in [1] and for quantum lattice systems by Lanford and Robinson in [2]. However, there were important reasons for wanting to prove variants of this result which involved more general shapes than boxes, such as the variant known as '(limiting) mean entropy in the sense of van Hove' [1]. This had been proven in the classical case in the Robinson–Ruelle paper [1] as a consequence of a general property called *strong subadditivity* (SSA). The Lanford–Robinson paper [2] put forward the conjecture that SSA holds also in the quantum case but, in the absence of a proof of this, could not immediately establish limiting mean entropy in the sense of van Hove. (It was in fact first proven for quantum systems by Araki and Lieb [3].) In fact, six years were to pass before SSA was finally established for quantum systems by Lieb and Ruskai [4].

Here we recall that, if ρ_{123} is a density operator on a Hilbert space which is given to us as a triple tensor product of three preferred Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, and if ρ_2, ρ_{12} and ρ_{23} denote its partial traces over $\mathcal{H}_1 \otimes \mathcal{H}_3$, $\mathcal{H}_3$ and $\mathcal{H}_1$ respectively, then the quantum version of SSA can be written as

$$S(\rho_{123}) + S(\rho_2) \leqslant S(\rho_{12}) + S(\rho_{23}) \tag{1}$$

where for any density operator ρ, $S(\rho)$ denotes its (von Neumann) entropy, $-\operatorname{Tr}(\rho \log \rho)$.

In the present paper, we are concerned not with the existence of limiting mean entropy as such but with more detailed questions related to the behaviour of the mean entropy of sequences of finite regions which are growing in size in the sense of set-inclusion. In particular, we are interested in answering the question under which circumstances the mean entropy of such sequences decreases monotonically. We shall also recall what is already known, and discuss what further can be proved, about the question whether the entropy of such sequences of regions increases monotonically. It turns out that the SSA property has immediate implications for both these questions and our purpose is to explore what these implications are and, in particular to use SSA to derive a number of general results on these questions.

Our first result is that, for translationally invariant states on translationally invariant infinite Euclidean lattice and continuum quantum systems, the *mean entropy* of boxes *decreases* monotonically as the boxes increase in size in the sense of set inclusion.

As a natural complement to this result, we shall include a proof (in fact, two alternative proofs) of the result that, for translationally invariant states on translationally invariant infinite Euclidean lattice and continuum quantum systems, SSA implies that the *entropy* of boxes *increases* monotonically, again as the boxes increase in size in the sense of set inclusion.

We shall also present some partial results, and pose a number of open questions, on the more problematic question to what extent these results can be generalized to more general shapes than boxes in translationally invariant infinite Euclidean lattice and continuum quantum systems, and also to certain appropriate types of shape in more general translationally invariant (/homogeneous) lattices and spaces than the usual infinite Euclidean lattices and spaces.

We now explain our basic setting and list our results in detail. We begin with the following discrete and continuous versions of the standard definition of a translationally invariant quantum system (see, for example, [2, 5]). In the case of a lattice, $\mathbb{Z}^\nu$, we define a *region* Λ to be a finite subset. In the case of a continuum, $\mathbb{R}^\nu$, we define a *region* Λ to be a measurable subset with finite volume. In either case, there is an assignment of a separable Hilbert space $\mathcal{H}_\Lambda$ to each region, satisfying, in the continuum case, the additional condition that this assignment be the same for any two regions which differ by a region of zero volume. Further, this assignment is required to satisfy the compatibility condition that, if two regions Λ_1 and Λ_2 are disjoint, then $\mathcal{H}_{\Lambda_1 \cup \Lambda_2} = \mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$, where, in the continuum case, two regions are said to be *disjoint* if their intersection has zero volume. We define a *state* mathematically to consist of a family $\{\rho_\Lambda\}$ of density operators (positive trace class operators with trace 1) on the Hilbert spaces $\mathcal{H}_\Lambda$ which are compatible in the sense that, for disjoint Λ_1 and Λ_2,

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\Lambda_2}(\rho_{\Lambda_1 \cup \Lambda_2}) \tag{2}$$

where for any region Λ, Tr_Λ denotes the partial trace over $\mathcal{H}_\Lambda$.

We remark that it is well known that, in the case where the single-site space is finite, classical lattice systems can be regarded as special cases of quantum lattice systems, where the density matrices representing the state are simultaneously diagonal, and so any result for a quantum lattice system will also be true for such a classical lattice system. However, our results below are not applicable to classical continuous systems since the property POS below fails (see [1]) in this case.

In this paper we shall mainly confine our interest to situations where not only the quantum system, but also the *state* is translationally invariant. This means that for all regions Λ and all translations τ from the relevant translation group ($\mathbb{Z}^\nu$ for lattice systems or $\mathbb{R}^\nu$ for continuous systems) there exists a unitary operator $U(\tau, \Lambda)$ from $\mathcal{H}_\Lambda$ to $\mathcal{H}_{\tau(\Lambda)}$ such that

$$\rho_{\tau(\Lambda)} = U(\tau, \Lambda)\rho_\Lambda U(\tau, \Lambda)^{-1}. \tag{3}$$

Given any state on a translationally invariant quantum system, we define the *entropy of a region* Λ to be the von Neumann entropy of ρ_Λ, i.e.

$$S(\Lambda) \stackrel{\mathrm{def}}{=} -\mathrm{Tr}(\rho_\Lambda \log \rho_\Lambda). \tag{4}$$

We remark that, in general, there is no guarantee that this will be well defined; it could of course be 'infinite'. But, for simplicity, we shall tacitly assume throughout this paper that it is always well defined and finite. Note however, that, even if this assumption were false, many of the results below would still hold when suitably modified to allow S to take its values in $[0, \infty) \cup \infty$.

When regarded as a function on regions, with values in $[0, \infty)$, the entropy of a region is known to satisfy a number of general properties [6], i.e. properties which hold in any state. In this paper, we shall mainly focus on:

POS: *positivity*.

$$S(\Lambda) \geqslant 0 \qquad \text{for all} \quad \Lambda.$$

SA: *subadditivity*.
If Λ_1 and Λ_2 are disjoint, then

$$S(\Lambda_1 \cup \Lambda_2) \leqslant S(\Lambda_1) + S(\Lambda_2).$$

SSA: *strong subadditivity*.

$$S(\Lambda_1 \cup \Lambda_2) + S(\Lambda_1 \cap \Lambda_2) \leqslant S(\Lambda_1) + S(\Lambda_2).$$

POS follows immediately from (4). SSA follows immediately from (1), (2) and (4). SA is just a special case of SSA but we prefer to list it as a separate property. Furthermore, if our state is translationally invariant, it follows immediately from (3) and (4) that:

TRANS: *translational invariance*.

For any element τ of the relevant translation group

$$S(\Lambda) = S(\tau(\Lambda)).$$

As we discussed above, SSA (or rather the more general statement (1)) has the status of a difficult theorem [4], but in spite of this, the game we wish to play from now on is to regard POS, SA, SSA and TRANS as axioms and to see what one can easily prove about the class of functions $\Lambda \mapsto S(\Lambda)$ from regions of $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ to the positive reals which obey these axioms. Clearly, any result we prove from these axioms will be a general property of the entropy of translationally invariant states on $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$. In sections 1–4, none of our results require more than these axioms; some of our results (as we shall indicate) do not require all of them. In section 5, we shall briefly consider the effect of adjoining to these axioms one further axiom, embodying yet another general property of entropy (i.e. $S\triangle$, see below).

We begin to play this game by defining the *mean entropy* $\bar{S}$ of a non-zero-volume region Λ by

$$\bar{S}(\Lambda) \stackrel{\text{def}}{=} \frac{S(\Lambda)}{|\Lambda|}$$

where $|\Lambda|$ denotes, in the lattice case, the number of lattice points contained in Λ and, in the continuum case, the volume of Λ.

We also define the notion of *box* regions, Λ_a, $a = (a_1, \ldots, a_\nu)$, where $a_1, \ldots, a_\nu$ are positive integers (in the lattice case) or positive real numbers (in the continuum case) by

$$\Lambda_a \stackrel{\text{def}}{=} \{x \in \mathbb{Z}^\nu \text{ or } \mathbb{R}^\nu : 0 < x_i \leqslant a_i \text{ for } i = 1, \ldots, \nu\}.$$

These have $|\Lambda_a| = \prod_{i=1}^{\nu} a_i$.

With these two definitions, we shall prove in sections 2 and 4 that, both in the lattice and continuum cases, and for arbitrary dimension ν, the axioms POS, SA, SSA and TRANS imply:

Theorem 1. $\Lambda_a \subset \Lambda_b \Rightarrow \bar{S}(\Lambda_a) \geqslant \bar{S}(\Lambda_b)$.

Theorem 2. $\Lambda_a \subset \Lambda_b \Rightarrow S(\Lambda_a) \leqslant S(\Lambda_b)$.

By POS and the elementary result from real analysis that any monotonic sequence which is bounded below has a limit, we immediately have from theorem 1 the corollary:

Corollary 1. *Given any infinite sequence of boxes* $\Lambda_{a(i)}$, $i = 1, 2, \ldots,$ *which increase in size in the sense of set inclusion,*

$$\lim_{i\to\infty} \bar{S}(\Lambda_{a(i)})$$

exists, and equals $\inf_i \bar{S}(\Lambda_{a(i)})$.

The special case of this where every edge length of $\Lambda_{a(i)}$ tends to infinity as i tends to infinity, reproduces the result of Lanford and Robinson [2] in the case where the box-sides tend to infinity monotonically.

We have found a number of intriguing hints that it may be possible to considerably generalize theorem 1 both to settings which involve a class of shapes more general than boxes and to translationally (and rotationally etc) invariant lattices and continua more general than $\mathbb{Z}^\nu$ and $\mathbb{R}^\nu$ including finite lattices such as the 'lattice circle with circumference N' (see before equation (20)). In section 3 we outline a number of partial results concerning such possible generalizations, some positive, some negative, and pose a number of (as far as we are aware) open questions.

Theorem 2 is not really a new result. In one dimension a stronger result (see equation (21)) is in fact well known, and, once one has established theorem 2 in this one-dimensional case, it is trivial to extend it to arbitrary dimension. Moreover, it is of course obvious that, in the classical lattice case (with finite single-site space but now without any assumption of translational invariance) the entropy of a region can only increase as one enlarges the region. Nevertheless, for completeness, we collect together in section 4 what is known about theorem 2 in one dimension and explicitly point out the extension to arbitrary dimension. We also include, in an appendix, a new alternative proof of the one-dimensional case of theorem 2 which we feel may be of independent interest because of a connection with recent work (see e.g. [7]) on 'quantum information theory'. This alternative proof involves the concept (which we introduce) of *m-point correlation entropies*, which, in a sense we explain, refine the notion of *index of correlation* (see e.g. [7]). Section 4 ends with a compendium of further questions, examples and partial results which concern the extent to which it may be possible to strengthen theorem 2 and/or generalize it to settings which involve a class of shapes more general than boxes and to translationally (and rotationally etc) invariant lattices and continua more general than $\mathbb{Z}^\nu$ and $\mathbb{R}^\nu$.

There are at least two further known general properties of entropy which would naturally lend themselves to being incorporated into our axiomatic scheme: namely, the inequality

$$S(\rho_1) \leqslant S(\rho_{12}) + S(\rho_2) \tag{5}$$

which, together with subadditivity constitutes the *triangle inequality* as stated in the paper [3] of Araki and Lieb, but which, in the present paper, we shall refer to by itself with this name, and the inequality

$$S(\rho_1) + S(\rho_3) \leqslant S(\rho_{12}) + S(\rho_{23}) \tag{6}$$

which was proven by Lieb and Ruskai in the same paper [4] in which strong subadditivity was proved and which was regarded, in that paper, as the second half of a property which was called there 'strong subadditivity' (the first part being what we call 'strong subadditivity' in this paper). See also [6] and [8]. Here, we propose to call (6) the *strong triangle inequality* since the triangle inequality can, in fact, be regarded as a special case of it. In both (5) and (6) we have assumed a similar setting and notation to that spelled out above before equation (1).

In the case of the translationally invariant quantum systems studied in this paper, one can incorporate these properties by augmenting our existing set of axioms POS, SA, SSA, TRANS with the following further axioms:

$\triangle$: *triangle inequality.*
For $\Lambda_1 \subset \Lambda_2$,

$$S(\Lambda_2 \backslash \Lambda_1) \leqslant S(\Lambda_2) + S(\Lambda_1).$$

$S\triangle$: *strong triangle inequality.*

$$S(\Lambda_1 \backslash \Lambda_2) + S(\Lambda_2 \backslash \Lambda_1) \leqslant S(\Lambda_2) + S(\Lambda_1).$$

We remark that $\triangle$ is actually a special case of $S\triangle$ (just as (5) is a special case of (6)) but, just as we did for SA and SSA above, we prefer to list it as a separate axiom. We also remark that one can easily see that $\triangle$ implies POS (i.e. $S(\Lambda) \geqslant 0$). We further remark that, in the case of box regions in the one-dimensional lattice $\mathbb{Z}$ and the one-dimensional continuum $\mathbb{R}$, and in the presence of POS, TRANS, SA and SSA, $\triangle$ and $S\triangle$ do not actually contain any new information (see section 5). On the other hand, it seems that, in higher dimensions and for more general lattices and continua, more can be proven by including $\triangle$ and $S\triangle$. We have begun to explore this and in a final section (section 5) we point out two ways in which, in the presence of POS and TRANS, $S\triangle$ may be used to obtain a number of results which throw a considerable amount of light on the 'compendium of further questions, examples, and partial results' at the end of section 4 which we mentioned above.

To end this introduction, we briefly consider what would happen if we were to drop the assumption of translational invariance, i.e. if we were to drop our axiom TRANS. It is well known and easy to see that theorems 1 and 2 would then fail. In fact, already on the tensor product of two Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, it is easy to find a total density operator ρ_{12} which is mixed (i.e. has non-zero entropy) but for which one of the reduced density operators, say ρ_1, is pure (i.e. has zero entropy) so that $\bar{S}(\rho_{12}) > \bar{S}(\rho_1)$. In fact, this is achieved by taking $\rho_{12} = \rho_1 \otimes \rho_2$, where ρ_2 is mixed. Also, it is well known (since it was first proven by Araki and Lieb in [3]) that, given any choice of density operator ρ_1 on a given Hilbert space $\mathcal{H}_1$, one can always find a Hilbert space $\mathcal{H}_2$ and a total density operator ρ_{12} on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ which is pure and such that ρ_1 coincides with its reduced density operator on $\mathcal{H}_1$ whereupon $S(\rho_1) > S(\rho_{12})$. However, we remark, on the one hand, that theorem 3, which we prove in section 3, tells us that, even in the absence of TRANS, theorem 1 continues to hold for general states on Euclidean lattices and continua in a suitable 'averaged' sense. On the other hand, in section 4, we shall give an example of a *translationally invariant* state (but now on a finite lattice) for which monotonicity of entropy fails. This example is elucidated further by theorem 4 in section 5.

Throughout this paper, we adopt the convention that a quantity is called 'positive' if it is greater than *or equal to* zero, that a sequence is called 'monotonically increasing' if each term after the first is greater than *or equal to* its predecessor, etc.

2. Proof of theorem 1

We shall treat in turn the four cases of the one-dimensional lattice, the ν-dimensional lattice, the one-dimensional continuum and the ν-dimensional continuum.

Case 1 (One-dimensional lattice). In this case, a box region, $\Lambda_{(n)}$, is simply a set consisting of the first n natural numbers for some natural number n. Writing $S(n)$ instead of $S(\Lambda_{(n)})$ for ease of notation, it follows from SA and TRANS that

$$S(q + r) \leqslant S(q) + S(r) \tag{7}$$

and from SSA and TRANS that

$$S(q + r + t) + S(r) \leqslant S(q + r) + S(r + t) \tag{8}$$

where $q, r, t \in \mathbb{N}$.

The statement of our theorem in this case amounts to the statement that the mean entropy $\bar{S} = S(n)/n$ is monotonically decreasing. We prove this by establishing the proposition

$$\frac{S(n)}{n} \geqslant \frac{S(n+1)}{n+1} \tag{9}$$

with the following simple inductive argument. First notice that a special case of (7) is the statement that $S(2) \leqslant 2S(1)$. This establishes proposition (9) in the case $n = 1$. Next, on the assumption that proposition (9) is true for $n = p$, we have, by (8) in the case $r = p$ and $q = t = 1$, that

$$S(p + 2) \leqslant S(p + 1) + S(p + 1) - S(p)$$

$$\leqslant 2S(p+1) - \frac{p}{p+1}S(p+1)$$
$$= \frac{p+2}{p+1}S(p+1)$$

which implies that (9) is true for $n = p+1$. We conclude that (9) is true and hence that theorem 1 is true in the case of a one-dimensional lattice.

We remark that an alternative proof of case 1 can be had by noting that SSA (or rather SA in the case of the first two terms and SSA in the case of the rest) implies (in fact is equivalent to) the monotonic decrease of the 'sequence of differences' $S(1), S(2) - S(1), S(3) - S(2), \ldots$. In fact, quite generally, one has the following easily proved lemma from real analysis:

Lemma 1. *If $S(1), S(2), S(3)\ldots$ is any sequence of positive numbers for which the sequence of differences is monotonically decreasing, then the sequence $S(1)/1, S(2)/2, S(3)/3, \ldots$, must be monotonically decreasing.*

In a classical context, this argument goes back as least as far as McMillan's work [9] on information theory. Note however that it is the first proof we gave above which is more closely related to the proofs of the generalizations we shall discuss in section 3.

Case 2 (ν-dimensional lattice). With a similar change in notation to that used above, we now need to prove

$$\frac{S(a_1, \ldots, a_\nu)}{a_1 \ldots a_\nu} \geqslant \frac{S(b_1, \ldots, b_\nu)}{b_1 \ldots b_\nu} \tag{10}$$

where $a_i, b_i \in \mathbb{N}$ and $a_i \leqslant b_i$ for $i = 1, \ldots, \nu$. We first notice that the function $S_{a_2,\ldots,a_\nu}(\cdot) \stackrel{\text{def}}{=} S(\cdot, a_2, \ldots, a_\nu)$, from the natural numbers to $\mathbb{R}$, clearly satisfies (7) and (8). Thus, by case 1, we have

$$\frac{S(a_1, a_2, \ldots, a_\nu)}{a_1 a_2 \ldots a_\nu} \geqslant \frac{S(b_1, a_2, \ldots, a_\nu)}{b_1 a_2 \ldots a_\nu}.$$

We next notice that, in a similar way to above, the function $S_{b_1,a_3,\ldots,a_\nu}(\cdot) \stackrel{\text{def}}{=} S(b_1, \cdot, a_3, \ldots, a_\nu)$ also satisfies (7) and (8). Thus, by applying case 1 again, we have

$$\frac{S(b_1, a_2, a_3, \ldots, a_\nu)}{b_1 a_2 a_3 \ldots a_\nu} \geqslant \frac{S(b_1, b_2, a_3, \ldots, a_\nu)}{b_1 b_2 a_3 \ldots a_\nu}.$$

One may clearly continue in this way, arriving at (10) after a total of ν such steps.

Case 3 (One-dimensional continuum). In this case, a box region, $\Lambda_{(x)}$, is simply a real interval $(0, x]$. Writing $S(x)$ instead of $S(\Lambda_{(x)})$ we now need to prove

$$\frac{S(y)}{y} \geqslant \frac{S(x)}{x} \tag{11}$$

for $y \leqslant x$.

We first argue that (11) holds on the rationals. For any two rationals x and y, let c be their common denominator and define the function $S_c(\cdot)$, taking its argument from the natural numbers, by $S_c(n) \stackrel{\text{def}}{=} S(n/c)$. This function satisfies (7) and (8) of case 1 and thus $S_c(n)/n$ and hence $S(n/c)/(n/c)$ are monotonically decreasing by the argument given there, thus establishing (11) for these x and y. To extend (11) to the reals, it then clearly suffices to prove that $S(x)$ is continuous. This follows immediately from the following lemmas and POS:

Lemma 2 (Lieb). *$S(x)$ is weakly concave i.e. for positive real numbers x and y, $S((x+y)/2) \geqslant S(x)/2 + S(y)/2$.*

Lemma 3. *If a real-valued function on a real interval is weakly concave and bounded below, then it is necessarily continuous.*

To prove lemma 2, first note that if $x = y$ the statement is trivially true. Otherwise, assume without loss of generality that $y < x$. The result then follows from (8) in the case established above where q, r and t are real, by identifying $q = t = (x - y)/2$ and $r = y$. We remark that this is essentially the same as an argument given in section IIIA of [8], where it is attributed to E H Lieb. Lemma 3 (or rather the alternative statement with 'convex' substituted for 'concave' and 'bounded above' substituted for 'bounded below') is proved in [10]. We remark that this is the only place where we use POS. In particular, POS is unnecessary for cases 1 and 2.

Case 4 (ν-dimensional continuum). This case can be established from case 3 by an argument similar to that used above to go from case 1 to 2.

This completes the proof of theorem 1. We remark that it can be helpful to visualize the steps in the above proof using a geometrical picture in which lattice points are identified with ν-dimensional continuum cubes of side 1. In detail, one identifies the particular lattice point $(1, \ldots, 1)$ with the particular continuum cube $\Lambda_{(1,\ldots,1)}$ and extends this identification by identifying the general lattice point $(a_1, \ldots, a_\nu)$, $a_1, \ldots, a_\nu \in \mathbb{Z}$, with the result of translating the cube $\Lambda_{(1,\ldots,1)}$ by the vector $(a_1 - 1, \ldots, a_\nu - 1)$. We also remark that, in the continuum case, theorem 1 can trivially be extended from the case of nested box regions to nested parallelepiped regions with parallel faces (by simply 'squashing' the boxes in the theorem).

3. Remarks about possible generalizations of theorem 1

We now discuss two different directions in which one can attempt to generalize theorem 1.

Firstly, one can ask whether theorem 1 generalizes to more general shapes than boxes (or parallelepipeds).

We remark that, for any given pair of nested regions, such a decrease-of-mean-entropy result clearly has a better chance of holding amongst states which are invariant, not only under the translation group of section 1, but also under rotations and reflections. This can be taken into account in our axioms by replacing 'translation group' by 'full symmetry group' in the axiom TRANS and from now on, whenever the discussion involves non-box regions, we shall assume this replacement to have been made.

We have found a few pairs of shapes which go beyond the box shapes (and parallelepiped shapes—cf the second remark at the end of section 2) of theorem 1 for which one can prove from our axioms that such a decrease-of-mean-entropy result holds. For example, in $\mathbb{Z}^2$ we can prove inequalities such as

$$\frac{S\left(\begin{smallmatrix}\square & \\ \square & \square\end{smallmatrix}\right)}{3} \leqslant \frac{S(\square\!\square)}{2} \tag{12}$$

where we are now using an obvious notation suggested by the first remark at the end of section 2.

Equation (12) may easily be proven from the special cases

$$S(\square\!\square) \leqslant S(\square) + S(\square)$$

$$S\left(\begin{smallmatrix}\square & \\ \square & \square\end{smallmatrix}\right) + S(\square) \leqslant S(\square\!\square) + S(\square\!\square)$$

of SA and SSA in an entirely analogous way to the way we established case 1 of theorem 1 from equations (7) and (8) in section 2 in the case that $q = r = t = 1$.

However, many of the cases where we have been able to prove such a result turn out to refer just to consecutive figures in a one-dimensional 'chain' of figures. For example the

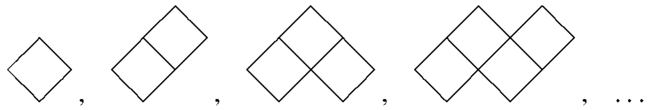


Figure 1. Chain of figures.

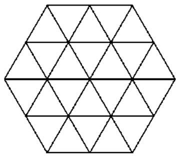


Figure 2. Hexagon figure.

case (12) illustrated above clearly easily extends to a more general inequality which refers to an arbitrary pair of successive figures in the chain shown in figure 1.

In an earlier draft of this paper we mentioned that we had been unable either to prove or disprove either of the two candidate inequalities

$$\frac{S(\boxminus\square\square)}{4} \overset{?}{\leqslant} \frac{S(\boxminus\square)}{3} \tag{13}$$

$$\frac{S(\boxminus\square\square)}{4} \overset{?}{\leqslant} \frac{S(\square\square\square)}{3}. \tag{14}$$

However, Bernhard Baumgartner [11] has since found, for each of these candidate inequalities, a symmetry-invariant state for which it is false. (Below, see the remark after corollary 2, we shall show that it is nevertheless impossible to find a single symmetry-invariant state in which both these candidate inequalities are simultaneously false.)

To summarize the above discussion, while there are some non-boxlike nested pairs for which it holds, decrease of mean entropy cannot hold unrestrictedly and the question:

Question 1. *For which pairs of regions, R_1, R_2, in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ satisfying $R_1 \subset R_2$, is it necessarily the case that $\bar{S}(R_1) \geqslant \bar{S}(R_2)$?*

seems to be a non-trivial one which requires further investigation.

In an attempt to throw some light on this question, it may be worth focusing on the more specific question:

Question 2. *For which pairs of similar regions R_1, R_2 in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ satisfying $R_1 \subset R_2$, is it necessarily the case that $\bar{S}(R_1) \geqslant \bar{S}(R_2)$?*

Of course, we know from theorem 1 that we can answer question 2 positively for the case of similar boxes and parallelepipeds. But we have found it difficult to find other pairs of similar shapes for which we can prove anything. In fact, we have not even been able to answer question 2 in the case of discs in $\mathbb{R}^2$ with increasing radii. However, we *have* been able to answer question 2 positively in some cases involving similar isosceles triangles in the plane and also in the case of two regular hexagons in the plane (i.e. $\mathbb{R}^2$) with diameters in the ratio two to one.

We now explain in detail how we treat the latter situation, making reference to figure 2.

Denoting the smaller hexagon made of six small triangles by H and the diamond region made of two small triangles by D, we begin by noting that the mean entropy of H is less than or equal to the mean entropy of D. This follows immediately once one notices that H can be viewed as the disjoint union of three copies of D, since by applying SA (twice) we have $S(H) \leqslant S(D) + S(D) + S(D)$, which implies that

$$\frac{S(H)}{6} \leqslant \frac{S(D)}{2}. \tag{15}$$

Next, imagine that the vertices of the central small hexagon in figure 2 are numbered (say) clockwise, starting at some particular vertex, from 1 to 6 and regard the large hexagon as the union of six copies of H, which we shall call $H_1, \ldots, H_6$, centred respectively at each of these six vertices. Also define the sequence of figures $F_1 = H_1$, $F_2 = F_1 \cup H_2$, $F_3 = F_2 \cup H_3$, etc so that F_6 is our large hexagon. We may then argue that each of these figures F_n, taken successively, has a mean entropy less than or equal to that of H. The first step in this argument proceeds by first noticing that F_2 consists of the union of two copies of H whose intersection is a copy of D and hence by SSA and TRANS that $S(F_2) + S(D) \leqslant S(H) + S(H)$. This is easily combined with the inequality (15) to conclude that $S(F_2)/10 \leqslant S(H)/6$. The subsequent steps proceed along similar lines, each using the result of the previous step, along with inequality (15) and the facts that (a) F_i consists of the union of the figure F_{i-1} and a copy of the figure H and (b) the intersection of F_{i-1} and the same H is a copy of D. After the fourth step we have the result that $S(F_5)/22 \leqslant S(H)/6$. For the final step we note that F_6 is the union of the figure F_5 and a copy of the figure H, but this time one sees that the intersection of these figures is a new figure G (composed of four small triangles). To derive the final result that $S(F_6)/24 \leqslant S(H)/6$ we now need, instead of (15), the result that $S(H)/6 \leqslant S(G)/4$. This can easily be shown by using SSA and TRANS to establish that $S(H) + S(D) \leqslant 2S(G)$ and combining this with (15).

Besides the above specific examples, we can also prove (say for a lattice $\mathbb{Z}^\nu$, and continuing to interpret lattice points as cubes and to refer to unions of cubes as 'figures') the general result which says that monotonicity of mean entropy does hold in a certain averaged sense:

Theorem 3. *The mean entropy of a figure $\mathcal{F}(n)$ composed of n cubes ($n \geqslant 2$) is less than or equal to the average of the mean entropies of all the (connected or disconnected) figures contained in $\mathcal{F}(n)$ which are composed of $n - 1$ cubes.*

We remark that theorem 3 and corollary 2 below actually only assume SA and SSA. In particular, the symmetry-invariance axiom TRANS is not required in any form.

Proof of theorem 3. First we introduce some new notation. Labelling the cubes of $\mathcal{F}(n)$ by the integers $1, \ldots, n$ we let $\mathcal{F}(n; i, j, \ldots)$ denote the figure that is formed from the figure $\mathcal{F}(n)$ by taking away its ith, jth, ... cubes. Then the statement of theorem 3 amounts to

$$\frac{S(\mathcal{F}(n))}{n} \leqslant \frac{1}{n} \sum_j \frac{S(\mathcal{F}(n; j))}{n-1}. \tag{16}$$

We prove this inequality by induction on n. First, (16) is true for all figures $\mathcal{F}$ with $n = 2$ by SA. Next, we assume that (16) is true for all figures $\mathcal{F}$ with $n = p$ cubes. Taking any figure $\mathcal{F}(p + 1)$, we note that $\mathcal{F}(p + 1; i)$ consists of just p cubes, so by our assumption

$$\frac{S(\mathcal{F}(p+1; i))}{p} \leqslant \frac{1}{p} \sum_{j \neq i} \frac{S(\mathcal{F}(p+1; i, j))}{p-1}. \tag{17}$$

Also, for $j \neq i$, SSA implies that

$$S(\mathcal{F}(p+1)) \leqslant S(\mathcal{F}(p+1; i)) + S(\mathcal{F}(p+1; j)) - S(\mathcal{F}(p+1; i, j)). \tag{18}$$

Summing (18) for $j = 1, \ldots, p+1$, with $j \neq i$, leads to

$$pS(\mathcal{F}(p+1)) \leqslant pS(\mathcal{F}(p+1; i)) + \sum_{j \neq i} S(\mathcal{F}(p+1; j)) - \sum_{j \neq i} S(\mathcal{F}(p+1; i, j)).$$

Combining this with (17) we have

$$\begin{aligned} pS(\mathcal{F}(p+1)) &\leqslant pS(\mathcal{F}(p+1; i)) + \sum_{j \neq i} S(\mathcal{F}(p+1; j)) - (p-1)S(\mathcal{F}(p+1; i)) \\ &= \sum_{j} S(\mathcal{F}(p+1; j)). \end{aligned}$$

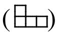
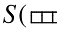
Dividing this last equation by $p(p+1)$ shows that (16) is true for $n = p+1$. □

This theorem also leads to the natural corollary:

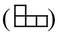
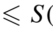
Corollary 2.

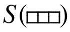
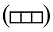
$$\frac{S(\mathcal{F}(n))}{n} \leqslant \frac{\max_j S(\mathcal{F}(n; j))}{n-1}.$$

Thus the mean entropy of a figure on a lattice is less than or equal to the mean entropy of at least one of its subfigures composed of one less cube. Returning to an example discussed above, we see that this remark implies that the mean entropy of the figure  is less than or equal to the mean entropy of one of its four subfigures each composed of three cubes. In fact, again assuming the axiom TRANS, one can prove, by an alternative route, that the mean entropy of  is less than or equal to the mean entropy of one of its two *connected* subfigures composed of three cubes; i.e. that one of the two inequalities (13) and (14) is actually true, but we cannot say which one. This can be done e.g. by first noting that, by SSA,

$S($$) + S($$) \leqslant S($$) + S($$).$

Combining this with (12) we have

$S($$) \leqslant S($$) + \frac{1}{3}S($$).$ (19)

But, we must have *either* $S($$) \leqslant S($$)$ *or* $S($$) \leqslant S($$)$. Thus, we conclude from (19) that (in any given state) one of the inequalities (13) and (14) must hold.

A second direction in which one can attempt to generalize theorem 1 is suggested by the fact that the basic setting of section 1 clearly generalizes to more general lattices than $\mathbb{Z}^\nu$ and to more general homogeneous spaces than $\mathbb{R}^\nu$ such as circles in one dimension and spheres and tori in higher dimensions. One can thus ask to what extent theorem 1 generalizes to such settings, where TRANS is now replaced by invariance under the relevant symmetry group. As far as more general lattices are concerned, we remark that the hexagon example discussed above could be regarded as an example concerning a triangular lattice. For the case of the one-dimensional circle and higher-dimensional tori, it is easy to see that the obvious analogue of theorem 1 still holds. For example, on the one-dimensional 'lattice circle of circumference N' (which one can think of as the subset of the unit circle consisting of the angles $2m\pi/N$, $m = 1, \ldots, N$) one easily shows, by a close analogue to the argument in case 1 of section 2, that

$$\frac{S(m_1)}{m_1} \geqslant \frac{S(m_2)}{m_2} \qquad \text{for} \quad m_1 \leqslant m_2 \tag{20}$$

where $S(m)$, $m \leqslant N$, denotes the entropy of m consecutive lattice sites. The obvious analogue of this in the continuum case also clearly holds. It is natural to ask the following specific question (and the obvious counterparts to this question in higher dimensions) concerning a possible generalization of this result, in the continuum case, to the two-sphere:

Question 3. *Does the mean entropy of a disc drawn on the surface of a sphere decrease monotonically as the solid angle subtended at the centre increases?*

But, just as for discs in $\mathbb{R}^2$, we have been unable to answer this question.

4. Proof and discussion of theorem 2

We begin by recalling that, in the one-dimensional lattice and continuum cases, Wehrl [8] has actually proven a stronger statement than theorem 2. In the lattice case, this amounts to the inequality

$$S(n+1) - S(n) \geqslant \inf_p \bar{S}(p) \tag{21}$$

where we use the same notation as in case 1 of section 2. In view of theorem 1, we of course know that we can replace the infimum in (21) by the limit $p \to \infty$. We now reproduce Wehrl's proof, specialized to this lattice case. In the appendix to this section, we give an alternative proof which we feel may be of interest because it uses some ideas related to some recent work on 'quantum information theory'.

Proof of (21). By equation (8) in the case $r = n$, $q = m$ and $t = 1$, we have

$$S(n+1) - S(n) \geqslant S(n+1+m) - S(n+m)$$

for any integers m and n. Summing over m from 1 to M where M is an arbitrary integer, we obtain

$$S(n+1) - S(n) \geqslant \frac{1}{M}\sum_{m=1}^{M}[S(n+1+m) - S(n+m)]$$

which 'telescopes' to

$$\begin{aligned} S(n+1) - S(n) &\geqslant \frac{1}{M}[S(n+1+M) - S(n+1)] \\ &\geqslant \frac{n+1+M}{M}\inf_p \frac{S(p)}{p} - \frac{S(n+1)}{M}. \end{aligned}$$

Equation (21) clearly follows immediately from this on taking the limit $M \to \infty$. □

Turning to the proof of theorem 2 itself, in view of the axiom POS, (21) clearly implies the same inequality with the right-hand side replaced by zero, which is equivalent to theorem 2 in the one-dimensional lattice case and one easily sees that this can then be proven to extend to $\mathbb{Z}^\nu$ and $\mathbb{R}^\nu$ in a similar way to the way in which we proved cases 2, 3 and 4 of theorem 1 from case 1 of theorem 1 in section 2. Thus we have arrived at a proof of theorem 2 by taking as our starting point a stronger result in a special case.

We next comment further on the one-dimensional lattice case. First we point out that while theorem 2 in this case is clearly a weaker statement than the result (21), in the presence of SSA and POS, it is actually very easy to recover (21) as an easy consequence of theorem 2. In fact, theorem 2 and SSA AND POS easily imply the even stronger result that

$$\lim_{n\to\infty}[S(n+1) - S(n)] = \lim_{p\to\infty}\bar{S}(p). \tag{22}$$

To see this, recall (cf before lemma 1 in section 2) that in the one-dimensional lattice case SSA is equivalent to the statement that the 'sequence of differences' $S(1)$, $S(2)-S(1)$, $S(3)-S(2), \ldots$ is monotonically decreasing while theorem 2 is equivalent to the statement that each of the terms in this 'sequence of differences' is positive and use the following easily proved lemma from real analysis:

Lemma 4. *If $S(1), S(2), S(3), \ldots$ is any sequence of positive numbers such the terms of its sequence of differences are each positive and decrease monotonically, then*

$$\lim_{n\to\infty}[S(n+1) - S(n)] = \lim_{p\to\infty} \bar{S}(p).$$

The chain of reasoning we have just followed to arrive at (22) resembles that given in [9, 12] for the classical lattice case. There is also a proof of (22) by van Enter [13] in the quantum lattice case, which proceeds somewhat differently: It does not assume monotonicity. Instead it requires the existence of limiting mean entropy to have been established first and uses the triangle inequality (i.e. axiom $\triangle$ mentioned towards the end of the introduction).

We remark that (22) can easily be extended to

$$\lim_{n\to\infty}[S(n+m) - S(n)] = m \lim_{p\to\infty} \bar{S}(p) \tag{23}$$

for any $m < n$ and $m, n \in \mathbb{N}$. Concerning the one-dimensional continuum case, there appear to be two natural counterparts to (22), (23). First, by using methods similar to those used in the proof of case 3 of section 2, we can show that

$$\lim_{x\to\infty}[S(x+y) - S(x)] = y \lim_{w\to\infty} \bar{S}(w) \tag{24}$$

for any $y < x$ and $y, x \in \mathbb{R}^+$. Alternatively, under the additional assumption that $S(z)$ is uniformly differentiable, it is not difficult to show, from (24), that

$$\lim_{x\to\infty} \left.\frac{\mathrm{d}S}{\mathrm{d}z}\right|_{z=x} = \lim_{w\to\infty} \bar{S}(w). \tag{25}$$

It would seem of interest to search for counterparts to (22)–(25) in dimensions higher than one, but we shall not pursue this in this paper.

Next we discuss further the general statement of theorem 2. First, just as we asked in section 3 whether theorem 1 generalizes to more general shapes than boxes, it seems of interest to ask:

Question 4. *Given any pair of regions R_1, R_2 in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ satisfying $R_1 \subset R_2$, is it necessarily the case that $S(R_1) \leqslant S(R_2)$?*

Or, in the case where the answer to this turns out to be 'no', in analogy to question 2, one can ask:

Question 5. *Given any pair of similar regions R_1, R_2 in $\mathbb{Z}^\nu$ or $\mathbb{R}^\nu$ satisfying $R_1 \subset R_2$, is it necessarily the case that $S(R_1) \leqslant S(R_2)$?*

We shall discuss a way of obtaining some partial answers to these questions in section 5 below.

One might also wonder whether theorem 2 generalizes to more general lattices than $\mathbb{Z}^\nu$ or to more general homogeneous spaces than $\mathbb{R}^\nu$. On this question, we remark that the prospects for generalization cannot be as straightforward as they were for theorem 1. (See however section 5 below.) In particular, we should emphasize that the proof of theorem 2 is unlike the proof of theorem 1 in that it makes essential use of the fact that the full system is infinite. If we drop this property, it is easy to find a counterexample. In fact, already in the case where one replaces the one-dimensional lattice $\mathbb{Z}$ by the lattice circle of circumference N (cf section 3) it is easy to see that monotonicity of *entropy* fails. In fact, it is then easy to have translationally invariant states ('pure total states') for which $S(N) = 0$ while $S(m) > 0$ for some $m < N$. This is to be contrasted with the result (20) that monotonicity of *mean* entropy continues to hold in this case. An amusing example of this is provided by the case where each point around our circle corresponds to a quantum system with Hilbert space $\mathcal{H} = \mathbb{C}^2$ and the pure total state,

$\rho = |\Psi\rangle\langle\Psi|$, is the generalized GHZ [14] state on the N-fold tensor product of $\mathcal{H}$ with itself where

$$\Psi = \frac{1}{\sqrt{2}}|\cdots\uparrow\uparrow\uparrow\cdots\rangle + \frac{1}{\sqrt{2}}|\cdots\downarrow\downarrow\downarrow\cdots\rangle$$

and $|\uparrow\rangle$ and $|\downarrow\rangle$ are a choice of orthonormal basis for $\mathcal{H}$. Clearly, in this case, we would have $S(m) = \log 2$ whenever $m < N$, but $S(N) = 0$!

There are a number of remarks worth making about this example (see also section 5). First, one might, at first sight, worry that a similar counterexample might be possible in the case that one were to replace the lattice circle of circumference N by $\mathbb{Z}$, in contradiction with theorem 2. For, after all, the case of $\mathbb{Z}$ can, picturesquely, be thought of as the case of 'lattice circle with $N = \infty$'. Of course the total state would now have to be understood (see e.g. [15]) say as a positive linear functional on the appropriate C^* algebra of local observables, but one still expects there to be examples of pure (i.e. extremal) total states for which (with any reasonable extension of the notion of 'entropy' to the algebraic framework) one would have '$S(\infty) = 0$' while $S(n)$ for some $n < \infty$ were non-zero. Indeed, one knows that, if the single-site Hilbert space is finite dimensional, there are 'plenty' of such pure states by the result of Fannes *et al* [16] that the pure translationally invariant states are weak-$*$ dense in the set of all translationally invariant states. But any worry that this might contradict theorem 2 is dispelled on realizing that theorem 2 only states that $S(n)$ will monotonically increase with n as long as n is *finite*! Nevertheless, one might still regard such examples as *puzzling* and we shall return to this issue in section 5.

Next, we remark that, amusingly, the state which formally corresponds to the generalized GHZ state of the above lattice-circle example in the case '$N = \infty$' and which thus formally 'looks like' a vector state on an infinite tensor product of $\mathbb{C}^2$ with itself is, when regarded as a positive linear functional on the appropriate C^* algebra, easily seen to be a mixed state. In fact, it would clearly be appropriate to assign to it an entropy of $\log 2$. Thus this particular state not only fails (as it must!) to give a counterexample to theorem 2 on $\mathbb{Z}$ but is actually not even of the 'puzzling' type mentioned above.

5. Some implications of the strong triangle inequality

Further light can be thrown on at least two of the matters discussed in the previous section by invoking the further axiom $S\triangle$, mentioned towards the end of the introduction, as we now discuss.

First of all, in the case of the one-dimensional lattice $\mathbb{Z}$ or continuum $\mathbb{R}$, it follows from $S\triangle$ and TRANS that

$$S(n) + S(m) \leqslant S(n+p) + S(m+p) \tag{26}$$

from which, setting $n = m$, one immediately obtains an alternative, and very quick, proof of monotonicity. Of course, in this (infinite) one-dimensional lattice case, (26) is clearly implied by POS, SA, SSA and TRANS since we know from section 4 that those axioms imply monotonicity of entropy, which clearly immediately implies (26). (In fact (26) is clearly equivalent to the monotonic increase of $S(n)$.) However, it is interesting to notice that, in the case of our lattice circle of circumference N, we know, by invoking $S\triangle$ that (26) will still hold true whenever $n + m + p \leqslant N$. This now gives us genuinely new information, which seems to be logically independent from the information which can be obtained by the use of POS, SA, SSA and TRANS alone. In particular, it allows us to conclude the interesting result that, on the lattice circle of circumference N, while, as we illustrated with our GHZ example above, the entropy $S(n)$ may well not increase montonically through the full range $0 \leqslant n \leqslant N$, it *must* necessarily increase monotonically throughout the first half of this range. In fact, we have:

Theorem 4. *For the lattice circle of circumference* N*,*

(a) $S(n) \leqslant S(n+1)$ *for* $0 \leqslant n \leqslant (N-1)/2$
(b) if $S(N) = 0$*, then* $S(n) = S(N-n)$*,*

part (b) being an easy consequence of $\triangle$. Hence, if the lattice circle is in a total pure state, then the entropy is necessarily monotonically increasing over the first half of the range, and monotonically decreasing over the second half of the range. The obvious continuum analogue of theorem 4 also clearly holds.

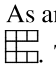
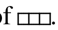
We remark that theorem 4 may be invoked to shed some light on the puzzle, raised in the previous section, concerning the existence of translationally invariant pure states on $\mathbb{Z}$ with non-zero, monotonically increasing entropies $S(n)$. In the case that such a state was the ground state of a translationally invariant Hamiltonian with a short-range interaction, it would seem reasonable to expect that it would arise, in a suitable sense, as the limit, as $N \to \infty$, of the sequence of (presumably pure) ground states for the corresponding Hamiltonians on the lattice-circle of circumference N. Denoting the entropies of the original state on $\mathbb{Z}$ now by $S(n, \infty)$ and the entropies of the corresponding states on the lattice circles of circumference N (say with $N \geqslant n$) by $S(n, N)$ we thus expect in particular that

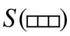
$$S(n, \infty) = \lim_{N \to \infty} S(n, N).$$

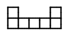
If one grants this equation, one arrives at the following 'physical' understanding of how it can be that $S(n, \infty)$ increases monotonically, the point being that, even though, according to theorem 4, each $S(n, N)$ rises and then falls as n increases, the 'turnover place' $(n = N/2)$ where it stops rising and starts falling will occur at larger and larger n as N increases.

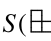
Secondly, the property $S\triangle$ can be used to shed more light on questions 4 and 5. We can use a similar method to that used for showing that $S\triangle$ implies the monotonicity of entropy in one dimension, to show that increase of entropy holds for certain pairs of general figures Λ_1 and $\Lambda_1 \backslash \Lambda_2$ in the lattice $\mathbb{Z}^\nu$. These pairs of figures are those for which there exist τ_1 and τ_2 from the enhanced (see the beginning of section 3) invariance group such that $\Lambda_1 = \tau_1(\Lambda_2)$ and $\Lambda_1 \backslash \Lambda_2 = \tau_2(\Lambda_2 \backslash \Lambda_1)$. For such figures, using TRANS, $S\triangle$ reduces to

$$S(\Lambda_1 \backslash \Lambda_2) \leqslant S(\Lambda_1). \tag{27}$$

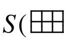
As an example of this, take the case where Λ_1 and Λ_2 are each copies of , with $\Lambda_1 \cup \Lambda_2 =$ . Then $\Lambda_1 \backslash \Lambda_2$ and $\Lambda_2 \backslash \Lambda_1$ are each copies of . By (27) we then have

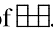
$S($$) \leqslant S($$).$ (28)

In a similar way, with the same shaped Λ_1 and Λ_2, but with $\Lambda_1 \cup \Lambda_2 =$  so that $\Lambda_1 \backslash \Lambda_2$ and $\Lambda_2 \backslash \Lambda_1$ are each copies of , one sees that

$S($$) \leqslant S($$).$ (29)

Interestingly, the pairs of figures in the inequalities (28) and (29) are just those for which decrease of mean entropy (see (13) and (14)) could not be proven. Indeed the method illustrated above for finding pairs of figures for which increase of entropy holds seems to be more prolific than the method illustrated in section 3 (cf after (12)) for finding pairs of figures for which decrease of mean entropy holds. However there are still problem cases, the first being the case

$S($$) \overset{?}{\leqslant} S($$).$ (30)

Our method fails to prove this inequality since there is no figure $\Lambda_1 \cup \Lambda_2$ such that Λ_1 and Λ_2 are copies of  and $\Lambda_1 \backslash \Lambda_2$ and $\Lambda_2 \backslash \Lambda_1$ are copies of . On the other hand, we are unaware of any example which would disprove it. (Note that such an example would, of course, have to be 'genuinely non-classical'.)

6. Epilogue

One immediate consequence of POS and SSA is that, if two regions each have zero entropy, then both their intersection and their union must also have zero entropy. This might be expressed by saying: 'If a state is pure on each of two regions, it must be pure on both their union and intersection.'

Amongst other things, this remark further illuminates one of the heuristic remarks (concerning theorem 6.4 of [17]) made in a paper [17] by Kay and Wald on quantum field theory in curved spacetime (see pp 55, 99 and 105 of [17]), namely, that it is impossible for a state to be pure on each of two 'double-wedge regions' [17] but mixed on their intersection. In fact, one of the motivations for the present research was a desire to elucidate that remark.

With an extension of the reasoning behind the above remark, another result which one can easily derive, now from our full set of axioms POS, SA, SSA and TRANS, is:

Theorem 5. *In both lattice and continuum cases, and for arbitrary dimension ν, if the entropy of any box is zero the entropy of all boxes is zero.*

One may prove this either as an immediate consequence of theorems 1 and 2, or as an easy direct consequence of the axioms.

Acknowledgments

We thank Tony Sudbery for pointing out to us the GHZ example mentioned at the end of section 4. We thank Aernout van Enter for valuable comments on an earlier version of this paper and for drawing to our attention the references [9, 12, 13]. We further thank Aernout van Enter for drawing to our attention the reference [16] and the puzzle (mentioned towards the end of section 4, and discussed further in section 5) concerning pure translationally invariant states with monotonically increasing entropy. We also thank Bernhard Baumgartner for comments on an earlier draft of this paper and especially for informing us of the progress he recently made (see after equations (13), (14)) in answering some of the questions we had posed there. Additionally, we thank Raymond Streater and also an anonymous referee for a number of further helpful remarks. ARK thanks the EPSRC for a research studentship.

Appendix to section 4

In this appendix we give an alternative proof of (21). We begin by introducing the notion of the *m-point correlation entropies* of a translationally invariant state.

To motivate this definition, we first recall the notion of the *index of correlation* (see for example [7], where it is discussed in an abstract setting concerning states on tensor products of Hilbert spaces). In the case of a one-dimensional quantum lattice system we can interpret this as the difference between the entropy of the union of n consecutive lattice points (or, in our alternative interpretation, cubes) and the sum of their individual entropies:

$$I_n \stackrel{\text{def}}{=} nS(1) - S(n). \tag{A1}$$

By using SA $n-1$ times, it is easy to show that I_n is positive.

Our new notion of *m-point correlation entropies* may be regarded as designed so as to provide a new way of writing the index of correlation I_n as a sum of positive terms, each of which concerns $m \leqslant n$ lattice points. Namely, we define the *m-point correlation entropies* by

$$S_m^c \stackrel{\text{def}}{=} \begin{cases} 2S(1) - S(2) & m = 2 \\ 2S(m-1) - S(m-2) - S(m) & m \geqslant 3. \end{cases} \tag{A2}$$

Note that S_m^c is positive by SA for $m = 2$ and by SSA for $m \geqslant 3$. An easy calculation then shows that

$$I_n = \sum_{m=2}^{n} (n+1-m) S_m^c. \tag{A3}$$

By (A1) and (A3), we can write the entropy of n consecutive lattice points as

$$S(n) = nS(1) - \sum_{m=2}^{n} (n+1-m) S_m^c. \tag{A4}$$

We note that by adding an extra lattice point onto a region of n consecutive lattice points, the entropy increases by $S(1)$ i.e. the entropy of one lattice point, but decreases by S_i^c (for $i = 2, \ldots, n+1$). Thus it is natural to think of S_i^c as a measure of the degree of correlation of a chain of lattice points of length i over and above the correlations involving subchains of length j where $j < i$.

To prove (21), we first note that all the terms in the sum in (A4) are positive. Thus for any $n > N$, removing the last $n - N$ terms gives us the inequality

$$S(n) \leqslant nS(1) - \sum_{m=2}^{N} (n+1-m) S_m^c$$

from which we have

$$\frac{S(n)}{n} \leqslant S(1) - \frac{1}{n} \sum_{m=2}^{N} (n+1-m) S_m^c.$$

Replacing the left-hand side by its infimum and then taking the limit $n \to \infty$, we deduce that

$$\inf_p \bar{S}(p) \leqslant S(1) - \sum_{m=2}^{N} S_m^c.$$

Substituting the expression for S_m^c given in equation (A2) into the right-hand side of this inequality, one finds that the sum 'telescopes' with all but two of its $3(N-2)+3$ terms cancelling and one is left with (21).

References

[1] Robinson D W and Ruelle D 1967 Mean entropy of states in classical statistical mechanics *Commun. Math. Phys.* **5** 288–300

[2] Lanford O E and Robinson D W 1968 Mean entropy of states in quantum-statistical mechanics *J. Math. Phys.* **9** 1120–5

[3] Araki H and Lieb E H 1970 Entropy inequalities *Commun. Math. Phys.* **18** 160–70

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

[5] Ruelle D 1969 *Statistical Mechanics* (New York: Benjamin)

[6] Lieb E H 1975 Some convexity and subadditivity properties of entropy *Bull. Am. Math. Soc.* **81** 1–13

[7] Horodecki R 1994 Informationally coherent quantum systems *Phys. Lett.* A **187** 145–50

[8] Wehrl A 1978 General properties of entropy *Rev. Mod. Phys.* **50** 221–60
[9] McMillan B 1953 The basic theorems of information theory *Ann. Math. Stat.* **24** 196–219
[10] Pólya G and Szegö G 1972 *Problems and Theorems in Analysis I* (Berlin: Springer)
[11] Baumgartner B 1999 Private communication
[12] Brascamp H J 1971 Equilibrium states for a one dimensional lattice gas *Commun. Math. Phys.* **21** 56–70
[13] van Enter A C D 1982 On a question of Bratteli and Robinson *Lett. Math. Phys.* **6** 289–91
[14] Greenberger D M, Horne M A and Zeilinger A 1989 Going beyond Bell's theorem *Bell's Theorem, Quantum Theory and Conceptions of the Universe* ed M Kafatos (Dordrecht: Kluwer)
[15] Bratteli O and Robinson D W 1987 and 1996 *Operator Algebras and Quantum Statistical Mechanics* 2nd edn, vols 1 and 2 (Berlin: Springer)
[16] Fannes M, Nachtergaele B and Werner R F 1992 Abundance of translation invariant pure states on quantum spin chains *Lett. Math. Phys.* **25** 249–58
[17] Kay B S and Wald R M 1991 Theorems on the uniqueness and thermal properties of stationary, nonsingular, quasifree states on spacetimes with a bifurcate Killing horizon *Phys. Rep.* **207** 49–136