Entropy, majorization and thermodynamics in general probabilistic theories

Howard Barnum

University of New Mexico
Albuquerque, NM

hnbarnum@aol.com

Jonathan Barrett

Department of Computer Science, University of Oxford
Oxford, U.K.

jonathan.barrett@cs.ox.ac.uk

Marius Krumm

Department of Theoretical Physics, University of Heidelberg, Heidelberg, Germany

M.Krumm@stud.uni-heidelberg.de

Markus P. Müller

Department of Theoretical Physics, University of Heidelberg, Heidelberg, Germany

Department of Applied Mathematics, University of Western Ontario, London ON, Canada

Department of Philosophy, University of Western Ontario, London ON, Canada

Perimeter Institute for Theoretical Physics, Waterloo, Canada

markus@mpmueller.net

Much progress has recently been made (see for example [1, 2]) in understanding the fine-grained thermodynamics and statistical mechanics of microscopic quantum physical systems, using the fundamental idea of thermodynamics as a particular type of *resource theory*. A resource theory is a theory that governs which state transitions, whether deterministic or stochastic, are possible in a given theory, using specified means. This depends on the kind of state transformations that are allowed in the theory, and in particular on which subset of them are specified as "thermodynamically allowed".

In this note, we lay some groundwork for the resource theory of thermodynamics in general probabilistic theories. We describe simple, but fairly strong, postulates under which the structure of systems in a theory gives rise to a natural notion of spectrum, allowing definition of entropy-like quantities and relations such as majorization analogous to those involved in fine-grained quantum and classical thermodynamics. These are candidates for governing which transitions are thermodynamically possible under specialized conditions (such as a uniform ("microcanonical") bath), and for figuring in definitions of free-energy-like quantities that constrain state transitions under more general conditions. In further work [3] we will investigate the extent to which they do so, under reasonable assumptions about which transformations are thermodynamically allowed. Since the postulates are shared by quantum theory, and the spectrum of quantum states is deeply involved in determining which transitions are possible, we expect that these postulates, supplemented with further ones (including a notion of energy), allow the development of a thermodynamics and statistical mechanics fairly similar to the quantum one.

Our main result, proven in Section 4, is that under certain assumptions, implied by but weaker than the conjunction of Postulates 1 and 2 of [4], the outcome probabilities of a finegrained measurement are majorized by those for a spectral measurement, and therefore the "spectral entropy" is the measurement entropy (and therefore concave). It also allows other entropy-like quantities, based on Schur-concave functions, to be defined. Our first assumption, described in Section 2 is a purely convex abstraction of the spectral decomposition of density matrices: that every state has a decomposition, with unique probabilities, into perfectly distinguishable pure (i.e. extremal) states. The spectral entropy (and analogues using other Schur-concave functions) can be defined as the entropy of these probabilities. Another assumption, *projectivity* (Section 3.1), abstracts aspects of the projection postulate in quantum theory; together with *symmetry of transition probabilities* it ensures the desirable behavior of the spectral entropic quantities that follows from our main result.

In Sections 5 and 6 we note that projectivity on its own implies a spectral expansion for *observables* (our additional spectrality assumption is for *states*), and also note the equivalence of the premises of our theorem on spectra to a strong kind of self-duality, known as *perfection*, of the state space.

Section 7 contains another main result of this work. Using spectrality, and some assumptions about the thermodynamic cost of certain processes that are partially motivated by our other postulates, especially projectivity, we generalize von Neumann's argument that the thermodynamic entropy of a quantum system is its spectral entropy, to generalized probabilistic systems satisfying spectrality. We then consider the prospect of embedding this result in a broader thermodynamics of systems satisfying relevant properties including the ones used in the present work, as well as others. Among the other useful properties, *Energy Observability*, which was used in [4] to narrow down the class of Jordan algebraic theories to standard complex quantum theory, can provide a well-behaved notion of energy to play a role in a fuller thermodynamic theory, and an ample automorphism group of the normalized space, acting transitively on the extremal points or, even stronger, satisfying Strong Symmetry ([4]), may enable reversible adiabatic processes that can be crucial to thermodynamic protocols.

While our postulates are strong and satisfied by quantum theory, it is far from clear that, even supplemented by energy observability, they constrain us to quantum theories: in [4] the strong property of no higher-order interference was used, along with the properties of Weak Spectrality, Strong Symmetry, and Energy Observability, to obtain complex quantum theory as the unique solution. While it is possible that latter three properties alone imply quantum theory, this would be a highly nontrivial result and we consider it at least as likely that they do not.

In the special case of assuming Postulates 1 and 2 of [4], a proof of our main theorem (Theorem 4.7) and several further results have been obtained in one of the authors' Master thesis [5]. We will elaborate on this, and in particular on the physics as detailed in von Neumann's thought experiment (cf. Section 7), elsewhere [3]. Note also the very closely related, but independent work of Chiribella and Scandolo [6, 7]. The main difference to our work is that (in most cases) they assume a "purification postulate" (among other postulates), and thus rely on a different set of axioms than we do. General-probabilistic thermodynamics has also been considered in [8, 9], where entanglement entropy and its role in the black-hole information problem has been analyzed. We hope that these different ways of approaching generalized thermodynamics will help to identify the main features of a probabilis-

tic theory which are necessary for consistent thermodynamics, and thus lead to a different, possibly more physical understanding of the structure of quantum theory.

1 Systems

In this section, we recall the general notion of system that we will use as an abstract model of potential physical systems in a theory, and define several properties of a system that in following sections will be related to the existence of a spectrum. We make a standing assumption of finite dimension throughout the paper except where it is explicitly suspended (notably in Appendix 5).

A system will be a triple consisting of a finite-dimensional regular cone $A_+ \subset A$, a distinguished regular cone $A_+^\sharp \subseteq A_+^*$ ($A_+^* \subset A^*$ being the cone dual to A_+), and a distinguished element u in the interior of $A_+^\sharp$. The normalized states are the elements $x \in A_+$ for which $u(x) = 1$; the set Ω of such states is compact and convex, and forms a base for the cone. Measurement outcomes, called *effects*, are linear functionals $e \in A_+^\sharp$ taking values in $[0,1]$ when evaluated on normalized states; a measurement is a (finite, for present purposes) set of effects that add up to u. Below, we will assume that $A_+^\sharp = A_+^*$, although we are investigating whether this can be derived from our other assumptions. Allowed dynamical processes on states will usually be taken to be *positive maps*: linear maps T such that $TA_+ \subset A_+$. Such a map is an *order-automorphism* if $TA_+ = A_+$, and *reversible* if $T\Omega = \Omega$. An *order-isomorphism* $T : A \to B$ between ordered vector spaces has $TA_+ = B_+$.

An *extremal ray* of a cone A_+ is a ray $\rho = \mathbb{R}_+ x$, for some nonzero $x \in A_+$, such that no $y \in \rho$ is a nonnegative linear combination of distinct elements of A_+ not in ρ. Equivalently, it is the set of nonnegative multiples of an extremal state of Ω. A cone is *reducible* if $A = A_1 \oplus A_2$, a nontrivial vector space direct sum, and all extremal rays of A_+ are contained either in A_1 or A_2, and irreducible if it is not reducible. Information about which of the summands A_i of a reducible cone a state lies in can be considered essentially classical; A_i are like "superselection sectors".

2 Spectrality

Axiom WS: ("Weak Spectrality") Every state ω has a convex decomposition $\omega = \sum_i p_i \omega_i$ into perfectly jointly distinguishable pure states.

Axiom S: ("Spectrality" (or "Unique Spectrality")). Every state has a decomposition $\sum_i p_i \omega_i$ into perfectly jointly distinguishable pure states. If it has more than one such decomposition, the two use the same probabilities. In other words, if $\omega = \sum_{i=1}^N p_i \omega_i = \sum_{i=1}^M q_i \rho_i$, where both ω_i and ρ_i are sets of perfectly distinguishable pure states, $p_i, q_i \geq 0$, $\sum_i p_i = \sum_i q_i = 1$, then there is a permutation σ such that for each i, $\sigma(p_i) = q_i$.

A priori Axiom **S** is stronger than Axiom **WS**. Later in this note we will give an example of a weakly spectral, but not uniquely spectral, system.

Note that **WS** is Postulate 1 of [4]. Postulate 2 of [4] is Strong Symmetry: that every set of mutually distinguishable pure states can be taken to any other such set of the same size by a reversible transformation (affine automorphism of Ω.) **WS** and Strong Symmetry together imply Axiom **S** and Axiom **P** (that is, "projectivity" as defined below). The converse is probably not true. Indeed, Postulates 1 and 2 of [4] imply the very strong property of *perfection*, which, as we note in Section 6 is equivalent to Axioms **S**, **P**, and Symmetry of Transition Probabilities.

3   Projective and perfect systems

3.1   Projectivity

We call a finite-dimensional system *projective* if each face of Ω is the positive normalized part of the image of a *filter*. Filters are defined in [4] to be normalized, bicomplemented, positive linear projections $A \rightarrow A$. This is equivalent to being the dual of a *compression*, where the latter is as defined in [13]. Normalization just means that they are contractive in (do not increase) the base norm, which for $x \in A_+$ is just $u(x)$. P a projection means $P^2 = P$, positive means $PA_+ \subseteq A_+$, complemented means there is another positive projection P' with $\mathrm{im}\,P \cap A_+ = \ker P' \cap A_+, \ker P \cap A_+ = \mathrm{im}\,P' \cap A_+$, and bicomplemented means complemented with complemented adjoint. It can be shown that filters are neutral: if $u(x) = u(Px)$ ("x passes the filter with certainty") then $Px = x$ ("x passes the filter undisturbed"). The complement, P', is unique. The projections $P : X \mapsto QXQ$ of quantum theory, where Q is a projector onto a subspace of the Hilbert space, are examples of filters. The existence of filters might be important for informational protocols such as data compression, or for thermodynamic protocols or the machinations of Maxwell demons. In finite dimension, a system is projective in this sense if and only if it satisfies the standing hypothesis of [13], Ch. 8.

A system is said to satisfy Axiom **P** ("Projectivity") if it is projective. The effects $u \circ P$, for filters P, are called *projective units*.

Proposition 3.1 ([13], Theorem 8.1)**.** *For a projective state space, the lattice of faces is complete and orthomodular. The filters and the projective units, being in one-to-one correspondence with faces, can be equipped with an orthocomplemented lattice structure isomorphic to that of the faces. For orthogonal faces F and G, $u_{F \vee G} = u_F + u_G$.*

The relevant orthocomplementation is the map $P \rightarrow P'$ described in the definition of filter above; by Proposition 3.1 it transfers to the lattices of faces and of projective units.

3.2   Self-duality and perfection

A regular cone A_+ is said to be *self-dual* if there exists an inner product on A such that $A_+ = \{y \in A : \forall x \in A_+, \ (y,x) \geq 0\}$. (We sometimes refer to the RHS of this expression, even when A_+ is not self-dual, as the *internal dual cone* of A_+ relative to the inner product $(.,.).$) This is equivalent to the existence of an order isomorphism $\varphi : A^* \rightarrow A$ such that bilinear form $\langle ., \varphi(.) \rangle$ is an inner product

on A. It is stronger than just isomorphism of A_+ to A^*, since we may have $\varphi(A_+^*) = A_+$ without the nondegenerate bilinear form $\langle ., \varphi(.) \rangle$ being positive definite.

Definition 3.2. *A cone $A_+ \subset A$ is called* perfect *if we can introduce a fixed inner product on A such that each face of the cone (including the cone itself) is self-dual with respect to the restriction of that inner product to the span of the face.*

For such cones, the orthogonal (in the self-dualizing inner product) projection P_F onto the span of a face is a positive linear map [14, 15]. It is clearly bicomplemented. If the system has a distinguished order unit, with respect to which P_F is normalized, then P_F is a filter.

Definition 3.3. *A* perfect system *is one whose state space A_+ is perfect and for which each of the orthogonal projections P_F onto* lin F *is normalized.*

It follows from this definition that the projections P_F of a perfect system are filters, hence a perfect system is projective.

Question 1. *For a perfect cone, is there always a choice of order unit that makes it projective?*

Question 2. *Are there perfect cones that can be equipped with order units that make them projective in inequivalent ways?*

One may investigate these questions by looking at an analogue of *tracial states* (Def. 8.1 and the remark following it in [13]). In an appropriate setting (which includes systems with spectral duality ([13], Def. 8.42) in general, and is equivalent to projective systems in the finite-dimensional case) a tracial state is one that is central, i.e. such that $(P + P')\omega = \omega$ for all filters P. Equivalently it is the intersection of Conv $(F \cup F')$ for all projective faces F. The conditions $(P + P')\omega = \omega$ are linear, so this defines a subspace of the state space; in a self-dual cone, it also gives a subspace of the observables, and it is natural to ask whether the order unit lies in that subspace, and whether, indeed, in an irreducible self-dual projective cone the tracial states are just the one-dimensional space generated by the order unit. If so, that suggests that we consider an analogue of the notion of the linear space generated by tracial elements, for perfect cones: the linear space spanned by states (or effects) such that $(P_F + P_{F'})\omega = \omega$ for every orthogonal projection P_F onto the span of a face F. We call the nonnegative elements of such a linear space in a perfect cone *orthotracial*.

Proposition 3.4. *A system with A_+ a perfect cone and an orthotracial element e in the interior of A_+ taken as the order unit is the same thing as a perfect projective system.*

Proof: Let e be orthotracial and in the interior of A_+. Orthotraciality of e says $P_F e + P_{F'} e = e$, i.e. $e - P_F e = P_{F'} e$. Since in any perfect cone the orthogonal projections onto spans of faces are positive, $P_{F'} e \geq 0$; hence $e - P_F e \geq 0$, i.e. $P_F e \leq e$, i.e. P_F is normalized. Conversely, in a perfect projective system the order unit is orthotracial (as well as tracial). We have already pointed out that the orthogonal projectors P onto spans of faces are compressions/filters in this context; from $p + p' = u$, and $p := Pu, p' := P'u$, we have $(P + P')u = u$ for all filters P. ∎

Question 3. *Is an orthotracial state automatically in the interior of A_+?*

4   Measurements, measurement entropy, and majorization

Definition 4.1. *Let Axiom* **WS** *hold. A* spectral measurement *on state ω is a measurement that distinguishes the pure states ω_i appearing in a convex decomposition of ω.*

Consider a system satisfying Axiom **S** whose normalized state space Ω has maximal number of perfectly distinguishable pure states N. Call a function $f : \mathbb{R}^n \to \mathbb{R}$ *symmetric* if $f(\mathbf{x}) = f(\sigma(\mathbf{x}))$ for any permutation σ. For any symmetric function $f : \mathbb{R}^n \to \mathbb{R}$, we define another function, $f : \Omega \to \mathbb{R}$, by $f(\omega) = f(\mathbf{p})$ where $\mathbf{p}$ are decomposition probabilities for f. By symmetry and unique spectrality this is independent of the choice of decomposition, so our claim that it defines a function is legitimate. Define the functions $\lambda^\downarrow : \Omega \to \mathbb{R}^n$ and $\lambda^\uparrow : \Omega \to \mathbb{R}^n$ to take a state and return the decreasingly-ordered and increasingly-ordered decomposition probabilities, respectively, of ω. Then $f(\omega) = f(\lambda^\downarrow(\omega)) = f(\lambda^\uparrow(\omega))$.

Definition 4.2. *For $x, y \in \mathbb{R}^n$, $x \prec y$, "x is majorized by y", means that $\sum_{i=1}^{k} x_i^\downarrow \leq \sum_{i=1}^{k} y_i^\downarrow$ for $k = 1, ..., n-1$, and $\sum_{i=1}^{n} x_i^\downarrow = \sum_{i=1}^{n} y_i^\downarrow$.*

If the first condition holds, and the second holds with $\leq$ in place of equality, we say x is *lower weakly majorized* by y, $x \prec_w y$.

We can extend the majorization relation to the set of all "vectors" (i.e. $1 \times n$ row matrices) of finite length (n not fixed) by padding the shorter vector with zeros and applying the above definition.

Theorem 4.3. *An $n \times n$ matrix M is doubly substochastic iff $y \in \mathbb{R}_+^n \implies (My \in \mathbb{R}_+^n \,\&\, My \prec_w y)$.*

This is C.3 on page 39 of [10].

Definition 4.4. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* Schur-concave *if for every $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\mathbf{v}$ majorizes $\mathbf{w}$ implies $f(\mathbf{v}) \leq f(\mathbf{w})$.*

Proposition 4.5. *Every concave symmetric function is Schur-concave.*

An effect is called *atomic* if it is on an extremal ray of A_+ (the cone of normalized effects) and is the maximal normalized effect on that ray. It is equivalent, in the projective context, to say it is an atom in the orthomodular lattice of projective units. In a projective state space the projective units are precisely the extremal points of the convex compact set $[0, u]$ of effects. Indeed, Proposition 8.31 of [13] states that "if the standing hypothesis of this chapter is satisfied [which it is in finite dimensional projective state spaces] and $A = V^*$, then the projective units are w^*-dense in the set of extreme points of A_1^+", the latter being the set of effects.

In projective state spaces, elements of A^* have a spectral expansion in terms of mutually orthogonal projective units. We may take these unit to be atomic, but then the expansion coefficients may be degenerate. We can also choose the expansion so that the coefficients are nondegenerate (the projective units no longer being necessarily atomic); the nondegenerate expansion is unique. In Appendix 5 these facts are shown to be the finite-dimensional case of [13], Corollary 8.65. We leave open whether the expansion is unique in the stronger sense analogous to that in (Unique)

Spectrality, that the probabilities in any expansion into atomic effects (though not necessarily the expansion itself) are unique.

One wonders whether analogous things hold for the state space A of a projective system. Weak Spectrality probably does hold (and perhaps some analogue of weak spectral decomposition for arbitrary elements of A therefore follows, i.e. one without uniqueness). But there are clear counterexamples to the conjecture that projectivity implies (Unique) Spectrality. Strictly convex sets in finite dimension are spectral convex sets in the sense of [13], Ch. 8, and therefore normalized state spaces of projective systems. But one can easily construct one even in two affine dimensions in which there is a state with two distinct convex decompositions into perfectly distinguishable ("antipodal") pure states, having very different probabilities. A non-equilateral isosceles triangle that has been perturbed ("puffed out") to be strictly convex (and therefore spectral [13], and the base for a projective system) does the trick. One can even construct an example (not strictly convex, but still spectral in the sense of [13], Ch. 8) in which there is a state with convex decompositions into different numbers of perfectly distinguisable pure states. See Theorem 8.87 of [13]. For the special case of the family of sets constructed in that theorem, illustrated in their Fig. 8.1, the "triangular pillow" (an equilateral triangle puffed into a third dimension), the state at the barycenter of the equilateral triangle (with vertices the three pure states in the "equatorial plane") can be written as the sum of $1/3$ times each of the three vertices, or of $1/2$ times the "north pole" plus $1/2$ times the "south pole". It would be nice to know whether or not this state space is self-dual.

For a projective state space, every atomic effect takes the value 1 on a *unique* normalized state, which is extremal in Ω, called $\widetilde{e}$, and every extremal normalized state takes the value 1 on a unique atomic effect, called $\widehat{\omega}$. $\widehat{}$ and $\widetilde{}$ are 1-1 maps of the atomic effects onto the pure states and vice versa, and are each others' inverses. For a pair of states ω, σ, $\widehat{\omega}(\sigma)$ is sometimes called the *transition probability* from σ to ω.

Definition 4.6 (Symmetry of Transition Probabilities)**.** *A system is said to satisfy* Symmetry of Transition Probabilities *(or Axiom* **STP***) if for any pair of pure states* ω, σ, $\widehat{\omega}(\sigma) = \widehat{\sigma}(\omega)$.

Theorem 4.7. *Let a system satisfy Unique Spectrality, Symmetry of Transition Probabilities, and Projectivity. Then for any state* ω *it holds that for any fine-grained measurement* $e_1, ..., e_n$, *the vector* $\mathbf{p} = [e_1(\omega), ..., e_n(\omega)]$ *of probabilities of the measurement outcomes is majorized by the vector of probabilities of outcomes for a spectral measurement on* ω.

Lemma 4.8. *For a system satisfying Projectivity,* ω *is perfectly distinguishable from* σ *if, and only if,* $\text{Face}(\omega) \subseteq \text{Face}(\sigma)'$.

Proof: [of Lemma 4.8] It follows straightforwardly from the definition of filters and their complements, that for P the filter associated with $\text{Face}\sigma$ and P' the filter associated with $\text{Face}(\sigma')$, the projective units $u \circ P$ and $u \circ P'$ distinguish σ from ω.

Proof: [of Theorem 4.7] Let $\omega = \sum_j p_j \omega_j$ be a convex decomposition, with ω_j pure and perfectly distinguishable. Then p_j are the outcome probabilities for a spectral measurement on ω. Write the effects e_i of an arbitrary fine-grained measurement as $e_i = c_i \pi_i$, where $0 < c_i \leq 1$ and π_i are atomic. Then the outcome probabilities for this measurement, made on ω, are

$$q_i = \sum_j p_j c_i \pi_i(\omega_j) \,. \tag{1}$$

That is, $q = Mp$ where $M_{ij} = c_i\pi_i(\omega_j)$. So $\sum_{i=1}^{M} M_{ij} = \sum_i c_i\pi_i(\omega_j) = u(\omega_j) = 1$. That is, M is row-stochastic. Also $\sum_{j=1}^{N} M_{ij} = c_i\pi_i(\sum_j \omega_j)$. By Symmetry of Transition Probabilities this is equal (using the fact that $\widetilde{}$ and $\widehat{}$ are inverses) to $c_i\sum_j \widehat{\omega}_j(\widetilde{\pi}_i)$. Lemma 4.8 tells us that $\sum_j \widehat{\omega}_j = u$ so this is equal to: $c_i u(\widetilde{\pi}_i) = c_i$. So, $\sum_{j=1}^{N} M_{ij} = c_i$. $c_i \leq 1$, so M is column-substochastic.

So M is doubly substochastic. Letting $R \geq N$ be the number of outcomes of the finegrained measurement, we pad p with $R - N$ zeros and pad M on the right with $R - N$ zero columns to obtain a doubly substochastic matrix $\tilde{M}$. Then $\tilde{M}\tilde{p} = q$, so by Theorem 4.3 $q \prec_w p$. Since $\sum_i p_i = \sum_i q_i = 1$, lower weak majorization implies majorization, $q \prec p$.   ∎

Corollary 4.9. *In a perfect system satisfying Axiom* **S***, for any state ω the outcome probabilities for any fine-grained measurement on ω are majorized by those for a spectral measurement on ω. In particular, this is so for systems satisfying Postulates 1 and 2 of [4].*

The first statement holds because, as we will show in Section 6, perfect systems are the same thing as projective systems satisfying **STP**. The second sentence holds because Postulates 1 and 2 of [4] imply both **P** and **S**. While we shall see that perfection implies weak spectrality, we do not know whether it implies **S**, so **S** had to be included in the premise of the Corollary.

Corollary 4.10. *Let $\omega' = \int_K d\mu(T)T_\mu(\rho)$, where $d\mu(T)$ is a normalized measure on the compact group K of reversible transformations. Then $\omega \preceq \omega'$.*

Proof: Let e_i be the spectral measurement on ω'. Then $e_i(\omega') = \int_K d\mu(T)e_i(T(\rho))$. For any state σ, write **p** for the vector whose i-th entry is $e_i(\sigma)$. Then the spectrum of ω' is $\mathbf{p}(\omega)$, and $\mathbf{p}(\omega) \equiv \int_K d\mu(T)\mathbf{p}(T(\omega))$ but $\mathbf{p}(T(\omega))$ is just the vector of probabilities for a measurement on ρ in the frame $\{T^\dagger e_i\}$, hence is majorized by the spectrum of ω. A limit of convex combinations of such things, for example the spectrum of ω', also majorizes the spectrum of ω.   ∎

Definition 4.11. *The* measurement entropy $S_{meas}(\omega)$ *of a state ω of a system A is defined to be the infimum, over finegrained measurements, of the entropy of the outcome probabilities of the measurement.*

Recall that $S(\omega)$ is defined, for any Axiom-**S** theory, as the entropy of the probabilities in any convex decomposition of ω. From the definitions of $S(\omega)$ and $S_{meas}(\omega)$, Theorem 4.7 and the Schur-concavity of entropy we immediately obtain:

Proposition 4.12. *For any state $\omega \in \Omega_A$, of a system satisfying Axioms S, P, and STP, equivalently (see the next two sections) satisfying Axiom S and perfection, $S(\omega) = S_{meas}(\omega)$.*

From this we immediately obtain, by Theorem 1 in [11] (concavity of the measurement entropy; there is probably a similar theorem in [12]), that $S(\omega)$ is concave in any system satisfying Axioms **S**, **P**, and **STP** (and hence in any system satisfying Postulates 1 and 2 (WS and Strong Symmetry) of [4]). Similarly, under these assumptions any entropy-like quantity constructed by applying a Schur-concave function χ to the spectrum will be the same as the infimum of χ over probabilities of measurement outcomes, and if χ is concave, so will be the function $\omega \mapsto \chi(spectrum(\omega))$.

5 Spectral expansions of observables in projective systems

In proving Theorem 6.1 in Section 6 (that any (finite-dimensional) projective system satisfying symmetry of transition probabilities is perfect) we will use the following, which is of interest in its own right.

Proposition 5.1. *In a projective system each $a \in A^*$ is a linear combination $\sum_{i=1}^n \lambda_i p_i$ of mutually orthogonal projections p_i in the $\mathscr{P}$-bicommutant of a. We can always choose the expansion so that the coefficients p_i are nondegenerate (i.e. $i \neq j \implies p_i \neq p_j$), and then the expansion is unique.*

We call the unique expression for *a* as a nondegenerate linear combination of mutually orthogonal projectors its *spectral expansion*. As shown in Appendix A, Proposition 5.1 is the finite-dimensional case of:

Proposition 5.2 ([13], Corollary 8.65). *If A and V are in spectral duality, then each $a \in A$ can be approximated in norm by linear combinations $\sum_{i=1}^n \lambda_i p_i$ of mutually orthogonal projective units p_i in the $\mathscr{P}$-bicommutant of a.*

As noted in the discussion following the definition of spectral duality, Definition 8.42 of [13], by Theorem 8.72 of [13], in finite dimension their property of spectral duality is equivalent to all exposed faces of Ω being projective, i.e. in our terminology, to the system (A, u) being projective. The $\mathscr{C}$-bicommutant of *a* is the set of all compressions compatible with *a* (i.e. $a = Pa + P'a$) and with all compressions compatible with *a*. Compatibility of two compressions is equivalent to commutation. Compatibility of projective units is defined as compatibility of the associated compressions; the definition of $\mathscr{P}$-bicommutant is exactly analogous.

6 For projective systems, symmetry of transition probabilities is perfection

The following theorem shows that we may replace the conjunction of Projectivity and Symmetry of Transition Probabilities with the the property of perfection. Parts of our proof are modeled after the proof of Lemma 9.23 of [13], but with different assumptions: finite dimension makes certain things simpler for us, but our premise involves only projectivity and symmetry of transition probabilities, not the additional property of purity-preservation by compressions that figures in said Lemma. After proving the theorem, we realized that essentially the same result, stated in somewhat different terms, was proved by H. Araki in [16]. We include our proof in Appendix B.

Theorem 6.1. *Let A, u be a finite-dimensional projective system satisfying Symmetry of Transition Probabilities. Then there is a unique positive linear map $\varphi : A^* \to A$ such that $\varphi(x) = \widehat{x}$ for each atom $x \in \mathscr{P}$. (x, y), defined by $(x, y) := \langle x, \varphi(y) \rangle$, is an inner product on A^*, with respect to which compressions are symmetric, i.e.:*

$$(Pa, b) = (a, Pb). \tag{2}$$

Hence A_+^* (and so also A_+) is a perfect self-dual cone, so the system A, u is a perfect system.

Corollary 6.2. *For projective systems, symmetry of transition probabilities is equivalent to perfection.*

Proof: Theorem 6.1 gives one direction; the other direction, that perfect projective systems satisfy STP, is near-trivial (cf. [4]).

Corollary 6.3. *For a projective system satisfying STP any element $x \in A$ has a "spectral" expansion $x = \sum_i \lambda_i \omega_i$, with $\lambda_i \in \mathbb{R}$ and ω_i mutually orthogonal pure states in Ω.*

This follows from the spectral expansion of observables (Proposition 5.1) and Theorem 6.1 since the latter implies self-duality. The uniqueness properties of the expansion of elements of A^* (cf. discussion following Proposition 5.1) also hold for elements of A_+ since the expansion in the state space will be the image of the expansion of Proposition 5.1 under the order-isomorphism $\varphi : A^* \to A$.

Question 4. *Does perfection imply the stronger uniqueness properties embodied in Axiom* **S***?*

If the triangular pillow based state space (see Sec. 4) is perfect, then it does not.

7   Filters, compressions, and von Neumann's argument for entropy

We have given assumptions that imply the existence of a spectral entropy and related quantities with operational interpretation in terms of probabilities of measurement entropy, and majorization properties, such as Corollary 4.10 that, in the quantum case, play a crucial role in thermodynamic resource theory. We would like to use the spectrum and associated entropic quantities and majorization relations in a generalized thermodynamic resource theory. In this section we take a step in this direction by extending von Neumann's argument that $S(\rho)$ is the thermodynamic entropy in quantum theory, to systems whose internal state space is a more general GPT state space satisfying Axiom **S** and Axiom **P**.

Von Neumann's argument is that a reversible process, making use of a heat bath at temperature T, exists between a system with density matrix ρ, and a system in a definite pure state, and that this process overall involves doing work $-kT \operatorname{tr} \rho \ln \rho$ in the forward direction. His argument involves a system with both quantum and classical degrees of freedom, e.g. a one-molecule ideal gas, and the direct heat exchange and doing of work involves classical degrees of freedom (specifically, expanding or contracting the volume occupied by a gas, while in contact with the heat reservoir).

Consider an arbitrary state $\omega = \sum_i q_i \omega_i$, where q_i, ω_i are a convex expansion of the state into a set of N perfectly distinguishable pure states ω_i. Axiom **S** ensures that such expansions exist and that they uniquely define (as in the preceding section, except with different units corresponding to taking ln instead of the base-2 log) $S(\omega) := -\sum_i q_i \ln q_i$. By Lemma 4.8 Axiom **P** ensures that there are atomic projective units π_i that form a measurement distinguishing the states ω_i (that is to say, $\sum_i \pi_i = u$, and $\pi_i(\omega_j) = \delta_{ij}$). There are associated filters P_i such that $P_i \omega_j = \delta_{ij} \omega_j$.

Assumption: *if such filters exist, they allow us, at no thermodynamic cost, to take a cylinder of volume V, containing such a particle in equilibrium at temperature T and separate it into N cylinders*

C_i of volume V, still at temperature T, such that C_i contains a system in state ω_i, with probability p_i, in which case the other cylinders are empty.

We may think of this separation process as realizing the measurement by using the instrument $\{P_i\}$, with the classical measurement outcome recorded as which cylinder the system is in. Because of projectivity, it is consistent to assume that this is possible, and that the state of the system in cylinder i (i.e., conditional on measurement outcome i) is still ω_i due to the neutrality property of filters. von Neumann's argument involves instead semipermeable membranes, allowing particles whose internal state is $|i\rangle\langle i|$ to pass (from either direction), whilst reflecting particles whose internal state is supported on the subspace of Hilbert space orthogonal to $|i\rangle\langle i|$. The use of analogous semipermeable membranes, in the GPT case, which behave differently for systems in face F_i (i.e., whose state "is" ω_i) than for particles whose internal state "is" in F_i', will allow us to ultimately separate each of the mutually distinguishable states ω_i into its own box or cylinder. We may, if we like, represent such a procedure by a transformation on a tensor product of a classical state space and the internal state space of the particle, for example:

$$x \otimes \omega \mapsto (x \oplus 0) + P_i \omega \otimes (x \oplus 1) \otimes P_i' \omega, \tag{3}$$

easily verified to be positive and base-norm-preserving.

In fact, the overall process that separates particles into the cylinders C_i could just be represented as the positive map:

$$T : x \otimes \omega \mapsto \sum_i (x + i - 1) \otimes P_i \omega \tag{4}$$

where the first register is classical and takes values $1, ..., N$, where N is the maximal number of distinguishable states. Again this is positive and trace preserving on the overall tensor of the classical N-state system with the GPT system (which is equivalent in structure to the direct sum of N copies of the GPT system). The possibility of such transformations is due to the projectivity of the state space (which implies such properties as the neutrality of filters). Whether or not it is reasonable to consider them thermodynamically costless is less clear, especially because the overall transformation on the GPT-classical composite is not in general an automorphism of the normalized state space (not "reversible"). Ultimately, the reasonableness of this assumption probably requires that the "measurement record" be kept in a system for which the overall measurement dynamics on the composite with the original system, can be reversible, a property which we are investigating.

Obviously, if $\omega = \sum_i q_i \omega_i$ where $\omega_i \in F_i$, then we have $T(1 \otimes \omega) = \sum_i q_i i \otimes \omega_i$.

At this point (or after the next step, it does not matter), we "adiabatically" transform the internal state ω_i of the particle in each cylinder C_i, to some fixed i-independent pure state, ω_0.

Assumption ("Adiabatic assumption"): This can be achieved without doing any work on the system, or exchanging any heat with the bath. (Thus we could do it while the system is isolated.)

If the reversible transformation group of the GPT system (the group of permitted transformations that are in the automorphism group of Ω) acts transitively on the pure states, that would motivate this assumption. This would follow, for example, from the much stronger property of Strong Symmetry.

Next, we isothermally (in contact with the heat bath at temperature T) compress the contents of each cylinder C_i to a volume $V_i := q_i V$. The work done on cylinder i, *if* it contains a particle, is $W_i = -\int_V^{qV} P dV$. By the ideal gas law, $PV = nkT$; with $n = 1$, $P = kT/V$ so $W_i = -kT \int_V^{qV} (1/V) dV = -kT(\ln qV - \ln V) = -kT \ln q \ln V / \ln V = -kT \ln q$. Since the probability the cylinder contains the particle is q_i, and these events, for the various cylinders i, are mutually exclusive and exhaustive, the expected work done in this step is $\sum_i q_i W_i = -kT \sum_i q_i \ln q_i = S(\omega)$.

Then we put the cylinders of compressed volume $q_i V$ next to each other and remove partitions separating them, obtaining a particle whose internal GPT degree of freedom is in the pure state ω_0, in equilibrium at temperature T and in the volume $\sum_i q_i V = V$. Since this process was reversible, we see that we may go from a particle in volume V and state ω, whose motional degrees of freedom are in equilibrium at temperature T, to one in motional equilibrium at temperature T and volume V but with state σ, by doing expected work $S(\omega) - S(\sigma)$ on the particle. ■

Just because our assumptions imply we can do this process, does not imply that we have a consistent thermodynamics (i.e. one without work-extracting cycles). It's possible that further properties of a GPT beyond projectivity and spectrality might be necessary for this. A notion of *energy*, such as Postulate 4 (Energy Observability) in [4] provides, would be needed for a thermodynamics that resembles our current thermodynamic theories, if we wish to discuss work done by or on GPT systems. We already mentioned the principle that if there exists a reversible transformation (automorphism of the normalized state space Ω), it can be applied with the system isolated or in contact with a heat bath, at no cost in work and with no heat exchange, and used it along with a further property, that the automorphism group of Ω acts transitively on pure states, to motivate assuming zero thermodynamic cost for a step in the von Neumann protocol. Perhaps we can find a similar motivation for the Strong Symmetry axiom of [4]. In [4] it was shown that the absence of higher-order interference was equivalent, given Weak Spectrality and Strong Symmetry, to the postulate that filters take pure states to multiples of pure states. This purity-preservation property greatly constrained state spaces, giving irreducible or classical Jordan systems, and at first blush it seems it might be important for thermodynamics. We suspect that it is not, and that a robust thermodynamics may be developed for GPT systems that share many of the remarkable properties of quantum theory, but are distinctly non-quantum in their interference behavior.

References

[1] F. G.S.L. Brandao, M. Horodecki, N. H. Y. Ng, J. Oppenheim, S. Wehner, *The second laws of quantum thermodynamics*, Proc. Natl. Acad. Sci. **112**, 3275 (2015). `arXiv:1305.5278`.

[2] M. Lostaglio, D. Jennings and T. Rudolph, *Description of coherence in thermodynamic processes requires constraints beyond free energy*, Nature Communications **6**, 6383 (2015), `arXiv:1405.2188`.

[3] H. Barnum, J. Barrett, M. Krumm, and M. P. Müller, in preparation.

[4] H. Barnum, M. Müller and C. Ududec, *Higher-order interference and single-system postulates characterizing quantum theory*, New J. Phys. **16**, 123029 (2014). `arXiv:1403.4147`.

[5] M. Krumm, *Thermodynamics and the Structure of Quantum Theory as a Generalized Probabilistic Theory*, Master Thesis, Heidelberg University, April 2015.

[6] G. Chiribella and C. M. Scandolo, *Entanglement and thermodynamics in general probabilistic theories*, `arXiv:1504.07045`.

[7] G. Chiribella and C. M. Scandolo, *Operational axioms for state diagonalization*, `arXiv:1506.00380`.

[8] M. P. Müller, O. C. O. Dahlsten, and V. Vedral, *Unifying typical entanglement and coin tossing: on randomization in probabilistic theories*, Commun. Math. Phys. **316**(2), 441–487 (2012). `arXiv:1107.6029`.

[9] M. P. Müller, J. Oppenheim, and O. C. O. Dahlsten, *The black hole information problem beyond quantum theory*, JHEP **09**, 116 (2012). `arXiv:1206.5030`.

[10] A. W. Olkin, I. Marshall and B. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd edition, Springer (2011).

[11] H. Barnum, J. Barrett, L. Clark, M. Leifer, R. W. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, *Entropy and information causality in general probabilistic theories*, New J. Phys. **12**, 033024 (2010). `arXiv:0909.5075`. Addendum, New J. Phys. **14** 129401 (2012).

[12] A. J. Short and S. Wehner, *Entropy in general physical theories*, New J. Phys. **12**, 033023 (2010). `arXiv:0909.4801`.

[13] E. M. Alfsen and F. W. Shultz, *Geometry of State Spaces of Operator Algebras*, Birkhäuser, Boston, 2003.

[14] B. Iochum, *Cônes Autopolaires dans les espaces de Hilbert*, Thèse de 3$^{\text{ème}}$ cycle, Marseille, 1975.

[15] B. Iochum, *Cônes Autopolaires et Algèbres de Jordan*, Springer (1984).

[16] H. Araki, *On a Characterization of the State Space of Quantum Mechanics*, Comm. Math. Phys. **75**, 1-24 (1980).

A   Proof of Proposition 5.1

In order to prove Proposition 5.1 we state (with very minor notational changes, and the incorporation of some definitions that Alfsen and Shultz place in surrounding text) Alfsen and Shultz' theorem of which it is a corollary (note that it is *V* that corresponds to what we've been calling *A*; *A* corresponds to what we call A^*):

Theorem A.1 ([13], Theorem 8.64)**.** *Assume A and V are in spectral duality, and let $a \in A$. Then there is a unique family $\{e_\lambda\}_{\lambda \in \mathbb{R}}$ of projective units with associated compressions P_λ such that*

(i) e_λ is compatible with a for each $\lambda \in \mathbb{R}$,

(ii) $P_\lambda a \leq \lambda e_\lambda$ and $P'_\lambda \geq \lambda e'_\lambda$ for each $\lambda \in \mathbb{R}$,

(iii) $e_\lambda = 0$ for $\lambda < -||a||$, and $e_\lambda = 1$ for $\lambda > ||a||$,

(iv) $e_\lambda \leq e_\mu$ for $\lambda < \mu$,

(v) $\bigvee_{\mu > \lambda} e_\mu = e_\lambda$ for each $\lambda \in \mathbb{R}$.

The family $\{e_\lambda\}$ is given by $e_\lambda = r((a - \lambda u)^+)$, each e_λ is in the $\mathscr{P}$-bicommutant of a, and the Riemann sums

$$s_\gamma = \sum_{i=1}^{n} \lambda_i (e_{\lambda_i} - e_{\lambda_{i-1}}) \tag{5}$$

converge in norm to a when $||\gamma|| \to 0$. Here the γ are finite increasing sequences $\lambda_0, \lambda_1, ..., \lambda_n$ of elements of $\mathbb{R}$, satisfying $\lambda_0 < -||a||$ and $\lambda_n > ||a||$ and the norm $||\gamma||$ of such a sequence γ is taken to be $\max_{i \in \{1,...,n\}} \lambda_i - \lambda_{i-1}$ (Note that n depends on γ, indeed $n \to \infty$ is necessary to secure $||\gamma|| \to 0$ given the bounds on λ_0 and λ_n.)

Proof of Proposition 5.1: As already mentioned, in finite dimension spectral duality just means that the system is projective. In finite dimension [and probably in infinite dimension too], $q < p$ for projective units q and p implies that $\lim F_q$ is a proper subspace of $\lim F_p$ (here F_p is the face of the cone A generated by p). Hence any chain $0 = e_0 < e_1 < e_2 < e_3 < \cdots e_n = 1$ of projective units has finite length no greater than one plus the dimension d of A. Thus, the family $\{e_\lambda\}$ contains only a finite number $n \leq d + 1$ of distinct projective units, which we index in increasing order as $e_0 = 0 < e_1 < \cdots e_n = 1$. (Whenever we write e with a roman index, the index indexes this set; the expression does *not* (except perhaps accidentally) refer to e_λ for the real number $\lambda = i$.)

Consider the sets $S_i := \{\lambda : \forall \mu \geq \lambda, e_\mu \geq e_i\}$. Each of these is an up-set in the ordering $\leq$ of $\mathbb{R}$, so it is either a closed or open half-line $[\mu_i, \infty)$ or (μ_i, ∞) unless $i = 0$ in which case it is $\mathbb{R}$. It is in fact closed: if it were open, $S_i = (\mu_i, \infty) \equiv \{\lambda : \lambda > \mu_i\}$, then by (v) of Theorem A.1, $e_{\mu_i} = \bigvee_{\lambda \in S_i} e_\lambda$, so by the definition of S_i, $\mu_i \in S_i$, contradicting $S_i = (\mu_i, \infty)$.

Consequently the function $\mathbb{R} \to \{e_0, ..., e_n\}$ which maps μ to e_μ is a sort of step-function; there are n distinct real numbers $\mu_1, ..., \mu_n$ such that the preimage of e_0 is $(-\infty, \mu_1)$, the preimage of e_j for $1 < j < n$ is $[\mu_j, \mu_{j+1})$, and the preimage of e_n is $[\mu_n, \infty)$. Let θ be the length of the shortest of these intervals; we have $0 < \theta < 2||a||/n$ [the only important thing about this seems to be that $\theta > 0$].

Since γ is an increasing sequence, by (iv) we have $e_{\lambda_i} - e_{\lambda_{i-1}} \geq 0$. There are n such differences in the Riemann sum; at most $n \leq d$ of them are nonzero; we call them $p_1, ..., p_n$. Since $\sum_i p_i = 1$, by Proposition 8.8 of [13], the nonzero ones are mutually orthogonal.

All sequences γ with $||\gamma|| < \theta$ have the same finite set of nonzero differences $p_i := e_{\lambda_i} - e_{\lambda_{i-1}}$, of cardinality $n \leq d$. For such γ, the Riemann sums s_γ lie in the finite-dimensional subspace of A spanned by the p_i. Like all subspaces of a finite-dimensional vector space, it is closed. Hence $\lim_{||\gamma|| \to 0}$ lies in this subspace, so it, too, is a finite linear combination of mutually orthogonal projective units. Since the family $\{e_\lambda\}_{\lambda \in \mathbb{R}}$ was unique, so is this linear combination. ∎

B   Proof of Theorem 6.1

There exists a basis $\{w_i\}$ for A^* consisting of atoms; there is a unique linear map $\varphi : A^* \to A$ that agrees with the map $x \mapsto \widehat{x}$ on this basis. We need to show that this agrees with $x \mapsto \widehat{x}$ more generally, or what is the same thing, that it is independent of the choice of atomic basis; and also that it is positive.

Let $x_1,...,x_n \in A^*$ be atoms, and $\lambda_i \in \mathbb{R}$ such that $\sum_{i=1}^{n} \lambda_i x_i \geq 0$. By STP, for each atom y

$$\langle y, \sum_{i=1}^{n} \lambda_i \widehat{x_i} \rangle = \langle \sum_{i=1}^{n} \lambda_i x_i, \widehat{y} \rangle. \tag{6}$$

So if $\sum_i \lambda_i x_i = 0$, then $\sum_i \lambda_i \widehat{x_i} = 0$, where x_i are any atoms. Now let x be an arbitrary unit, and expand it in the basis of units w_i: $x = \sum_i \alpha_i w_i$. So $x - \sum_i \alpha_i w_i = 0$, so $\widehat{x} = \sum_i \alpha_i \widehat{w_i} \equiv \sum_i \alpha_i \varphi(w_i) = \varphi(x)$. Since φ takes the set of all atomic effects (which generate the cone A_+^*) to the set of all extreme points of Ω (which generate the cone A_+), it takes A_+^* onto A_+, so it is an order-isomorphism. By φ's linearity the form $(.,.) := \langle ., \varphi(.) \rangle$ is bilinear. Since an order-isomorphism is in particular an isomorphism of linear spaces, the form $(.,.)$ is nondegenerate. That it is symmetric is easy to see from STP: for arbitrary $a = \sum_i a_i w_i$, $b = \sum_j b_j w_j$,

$$(a,b) = \sum_{ij} a_i b_j \langle w_i, \varphi(w_j) \rangle \tag{7}$$

but since w_i, w_j are atoms, $\varphi(w_j) = \widehat{w}_j$, and by STP $\langle w_i, \widehat{w}_j \rangle = \langle w_j, \widehat{w}_i \rangle = \langle w_j, \varphi(w_i) \rangle$, we have

$$(a,b) = \sum_{ij} a_i b_j \langle w_j, \varphi(w_i) \rangle = \langle \sum_j b_j w_j, \sum_i a_i \varphi(w_i) \rangle = (b,a). \tag{8}$$

To establish that $(.,.)$ is an inner product, it remains to be shown that $(x,x) \geq 0$ for all $x \in A^*$. To see this, use the spectral expansion $x = \sum_i \lambda_i p_i$, $\lambda_i \in \mathbb{R}$, p_i mutually orthogonal atoms, afforded by Proposition 5.1. Then

$$(x,x) = \sum_{ij} \lambda_i \lambda_j \langle p_i, \widehat{p}_j \rangle = \sum_{ij} \lambda_i \lambda_j \delta_{ij} = \sum_i \lambda_i^2 \geq 0. \tag{9}$$

Since φ is an order-isomorphism between A^* and A, and we have just established that the corresponding bilinear form is an inner product, we have shown that A^* (equivalently, A) is self-dual.

To show symmetry of P with respect to the form, we first establish

$$((I - P)x, Py) = 0 \tag{10}$$

where x, y are atoms. Write the spectral expansion of Py, $Py = \sum_i \lambda_i y_i$, with y_i mutually orthogonal atoms in $\text{im}_+ P$ and $\lambda_i \in \mathbb{R}$. Note that $w_i \in \text{im}P \implies \widehat{w}_i \in \text{im}P^*$. [Need details on why this is so, possibly an earlier proposition. It's Eq. 9.10 in [13]]. Then

$$((I - P)x, Py) = \sum_i \lambda_i ((I - P)x, y_i) \tag{11}$$

$$= \sum_i \lambda_i \langle (I - P)x, \widehat{y}_i \rangle = \sum_i \lambda_i \langle x, (I - P)^* \widehat{y}_i \rangle = 0. \tag{12}$$

The last equality is because $\widehat{y}_i \in \text{im}P^*$.

Now $(x, Py) \equiv ((I - P + P)x, Py) = ((I - P)x, Py) + (Px, Py)$, but by (10), this is equal to (Px, Py). Interchanging x and y, we have $(y, Px) = (Py, Px)$. But using symmetry of $(.,.)$ twice: $(Px, y) = (y, Px)$ and $(Py, Px) = (Px, Py)$ we obtain, as claimed, $(x, Py) = (Px, y)$.

Since a symmetric projection in a real inner product space is an orthogonal projection, we see that the compressions on A^*, equivalently (when A is identified with A^* via φ) the filters on A, are orthogonal projections with respect to the self-dualizing inner product $(.,.)$. So our cone is perfect by a result of Iochum [14, 15]: that a self-dual cone is perfect if and only if the orthogonal (in the self-dualizing inner product) projection P_F onto the the linear span of F is positive for each face F.[1]

∎

---

[1]A proof may be found in Appendix A of [4].