

The Thermomajorization Polytope and Its Degeneracies

Frederik vom Ende ^{1,*} and Emanuel Malveti ^{2,3,†}¹ Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany² School of Natural Sciences, Technische Universität München, 85737 Garching, Germany; emanuel.malveti@tum.de³ Munich Center for Quantum Science and Technology (MCQST) & Munich Quantum Valley (MQV), 80799 München, Germany

* Correspondence: frederik.vom.ende@fu-berlin.de

† These authors contributed equally to this work.

Abstract: Drawing inspiration from transportation theory, in this work, we introduce the notions of “well-structured” and “stable” Gibbs states and we investigate their implications for quantum thermodynamics and its resource theory approach via thermal operations. It is found that, in the quasi-classical realm, global cyclic state transfers are impossible if and only if the Gibbs state is stable. Moreover, using a geometric approach by studying the so-called thermomajorization polytope, we prove that any subspace in equilibrium can be brought out of equilibrium via thermal operations. Interestingly, the case of some subsystem being in equilibrium can be witnessed via the degenerate extreme points of the thermomajorization polytope, assuming that the Gibbs state of the system is well structured. These physical considerations are complemented by simple new constructions for the polytope’s extreme points, as well as for an important class of extremal Gibbs-stochastic matrices.

Keywords: quantum thermodynamics; thermal operations; thermomajorization; thermal equilibrium; Gibbs-stochastic matrix; cyclic process



Citation: vom Ende, F.; Malveti, E. The Thermomajorization Polytope and Its Degeneracies. *Entropy* **2024**, *26*, 106. <https://doi.org/10.3390/e26020106>

Academic Editors: Lucas Chibebe Céleri and Norton G. De Almeida

Received: 20 December 2023

Revised: 19 January 2024

Accepted: 22 January 2024

Published: 24 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The core idea of quantum thermodynamics—a field that has gained significant traction within the last decade—is to apply thermodynamic ideas to (ensembles of) systems of finite size, instead of “thermodynamically large” setups [1]. While this discipline comes with a number of principal questions and concepts (such as fluctuation theorems, thermal machines, the fundamental laws, thermalization, and many more; cf. [2]), an all-round successful approach has been to take the open systems perspective [3]. There, one models changes to a system of interest as the reduced action of a larger, closed system (i.e., system plus some form of environment, such as a bath) through some total Hamiltonian $H_{\text{tot}} = H_S + H_B + H_I$ consisting of a system, an environment, and an interaction term. (To avoid any confusion, let us point out explicitly that the individual Hamiltonians H_S, H_B are to be understood as local terms $H_S \otimes \mathbf{1}_E, \mathbf{1}_S \otimes H_B$, in particular, when adding them). Typical thermodynamic constraints imposed on top are that the environment (bath) starts out in equilibrium, i.e., in a Gibbs state, or that the total energy is conserved, i.e., $[H_{\text{tot}}, H_S + H_B] = [H_I, H_S + H_B] = 0$. Note that some works have weakened energy conservation to average energy conservation in the past. This leads to a framework where the quantum free energy characterizes all possible state changes [4].

This perspective ties into the information-theoretic and, more specifically, into the resource theory approach to quantum thermodynamics. There, one attempts to formalize which operations can be carried out at no cost (with respect to some resource, e.g., work), and one of the main aspects of this approach is to characterize when state transfers under thermodynamic constraints are possible. This comes with a subset of quantum channels

called *thermal operations*, which can be carried out without having to expend any resource. They arise from the above open systems construction as follows [5–7]. Given an n -level system described by some system’s Hamiltonian $H_S \in iu(n)$, as well as some fixed background temperature $T > 0$ that every bath needs to adopt, the set $\text{TO}(H_S, T)$ of all thermal operations with respect to H_S, T is defined as

$$\left\{ \text{tr}_B \left(e^{iH_{\text{tot}}} \left((\cdot) \otimes \frac{e^{-H_B/T}}{\text{tr}(e^{-H_B/T})} \right) e^{-iH_{\text{tot}}} \right) : \begin{matrix} m \in \mathbb{N}, H_B \in iu(m), H_{\text{tot}} \in iu(mn) \\ [H_{\text{tot}}, H_S + H_B] = 0 \end{matrix} \right\}.$$

Here, $iu(m)$ is the unitary Lie algebra in m dimensions so $iu(m)$ is the collection of Hermitian $m \times m$ matrices. As hinted at before, one of the central questions of this framework is the following: given an initial and a target state of some quantum system, can the former be transformed into the latter by means of a thermal operation?

The resource theory approach to quantum thermodynamics leads to a number of structural insights, ranging from optimal protocols for work extraction [4] and cooling [8] to the so-called second laws of quantum thermodynamics [9,10]. The latter precisely relates to the above state interconversion problem via a generalization of classical majorization called “thermomajorization”. First described by Ruch et al. [11] in the 1970s, thermomajorization has gained widespread popularity over the last decade due to the influential works of Brandão et al. [10], Horodecki and Oppenheim [9], and Renes [12], as well as many more [8,13–17]; there, among others, it has been used to solve the state interconversion problem in the quasi-classical realm; more details on this are given below. However, one can tackle this problem from another, more geometric perspective: for this, one abstractly defines the collection of all states that can (approximately) be generated via thermal operations starting from some initial state ρ

$$M_{H_S, T}(\rho) := \{ \Phi(\rho) : \Phi \in \overline{\text{TO}(H_S, T)} \},$$

and then one studies the geometric properties of $M_{H_S, T}(\rho)$. This set has been called the (*future*) *thermal cone* (Note that the term “cone” originates from light cones in general relativity and should not be confused with (convex) cones from linear algebra) [18–20] or, in the case of quasi-classical states ρ , the *thermal polytope* [8], and the elements of $M_{H_S, T}(\rho)$ are precisely those states that are said to be *thermomajorized* by ρ . In the quasi-classical case $[\rho, H_S] = 0$, the structure of the thermal cone is known to simplify considerably. This is due to the following crucial facts:

- Every thermal operation leaves the set of quasi-classical states invariant [21]: if $[\rho, H_S] = 0$, then $[\Phi(\rho), H_S] = 0$ for all $\Phi \in \overline{\text{TO}(H_S, T)}$.
- Thermal operations and general Gibbs-preserving quantum channels (approximately) coincide on quasi-classical states when considering only the diagonal (Section 3.4 [22]).

Combining these facts shows that for any state ρ with $[\rho, H_S] = 0$ where H_S is non-degenerate (i.e., ρ is diagonal in “the” eigenbasis of H_S), $M_{H_S, T}(\rho)$ equals the set of all diagonal states $\Phi(\rho)$ where Φ is a quantum channel that preserves the Gibbs state $e^{-H_S/T}/Z$. However, for diagonal states—where we write $x, y \in \mathbb{R}^n$ for the vectors of their respective diagonal entries—the existence of such a channel is equivalent to the existence of a Gibbs-stochastic matrix A such that $Ay = x$ (Corollary 4.5 [23]). For this recall that given some Hamiltonian $H_S = \text{diag}(E_1, \dots, E_n)$, as well as $T > 0$, a matrix A is called Gibbs-stochastic if the entries of A are non-negative, every column of A sums up to one, and $Ad = d$. Here, $d := (e^{-E_j/T})_{j=1}^n$ is the (unnormalized, but can also be normalized) vector of Gibbs weights. Moreover, we write $s_d(n)$ for the collection of all Gibbs-stochastic matrices with regard to the Gibbs vector d . Thus, $M_{H_S, T}(\text{diag}(y))$ equals

$\{\text{diag}(Ay) : A \in \mathbb{R}^{n \times n} \text{ Gibbs-stochastic}\}$ where one, instead and equivalently, can focus solely on the diagonal by considering the so-called *thermomajorization polytope*

$$M_d(y) := \{Ay : A \in \mathbb{R}^{n \times n} \text{ Gibbs-stochastic}\}. \tag{1}$$

At first glance, focusing on the quasi-classical case may appear fruitless in understanding the behavior of quantum systems. However, we point out that optimal cooling protocols rely on two-level Gibbs-stochastic matrices [8], which can be realized within the Jaynes–Cummings model [15]. Moreover, taking the quasi-classical perspective allows one to identify thermal operations that are simple to implement experimentally [24], and tools from the quasi-classical case have been used in quantum control theory to find non-trivial upper bounds on reachable states for dissipative bath couplings [25,26].

In this work, we will combine this geometric approach to quantum thermodynamics with the established field of transportation theory. While this is not the first time that these fields are being linked—cf. Section 2.2 for a short review and Section 3.2 for some results that this perspective has lead to already—this paper’s main idea is not to take tools, but rather (not as obvious, yet key) concepts from transportation theory and to investigate their implications in quantum thermodynamics. Thus, this work is structured as follows. We begin by reviewing characterizations of thermomajorization in the quasi-classical realm in Section 2.1, followed by a review of the basics of transportation theory in Section 2.2. Therein, we also introduce (or rather, translate) the key concepts of “stable” and “well-structured” Gibbs states, which will emerge to be quite intuitive. The implications of these notions will be the topic of Section 3, which contains the main results. More precisely, stable Gibbs states are found to be in one-to-one correspondence with the impossibility of (global) cyclic state transfers—which will also lead us to the notion of a “subspace in equilibrium”—cf. Section 3.1. The latter notion is found to be closely connected to the geometry of $M_d(y)$ and is reflected in the (number of) extreme points of $M_d(y)$, assuming that the Gibbs state is well structured (Section 3.2). The point of view taken in this work also leads to simple, intuitive ways to construct extreme points, as well as Gibbs-stochastic matrices that realize the corresponding state transfer; this is what the example and visualization Section 4 will illustrate.

2. Preliminaries and Review

2.1. Thermomajorization for Quasi-Classical States

The object related to thermomajorization most commonly found in the literature is the following: let any $d \in \mathbb{R}_{++}^n$ (that is, $d \in \mathbb{R}^n$ with $d > 0$ as d plays the role of the vector of Gibbs weights) as well as $y \in \mathbb{R}^n$ be given. One defines the *thermomajorization curve of y (with respect to d)*, denoted by $\text{th}_{d,y}$, as the piecewise linear, continuous curve fully characterized by the elbow points $\{(\sum_{i=1}^j d_{\tau(i)}, \sum_{i=1}^j y_{\tau(i)})\}_{j=0}^n$, where $\tau \in S_n$ is any permutation such that $\frac{y_{\tau(1)}}{d_{\tau(1)}} \geq \dots \geq \frac{y_{\tau(n)}}{d_{\tau(n)}}$. Equivalently (Remark 7 [27]), $\text{th}_{d,y} : [0, \mathbf{e}^\top d] \rightarrow \mathbb{R}$ for all $c \in [0, \mathbf{e}^\top d]$ satisfies

$$\begin{aligned} \text{th}_{d,y}(c) &= \min_{\{i=1,\dots,n : d_i > 0\}} \left(\left(\sum_{j=1}^n \max \left\{ y_j - \frac{y_i}{d_i} d_j, 0 \right\} \right) + \frac{y_i}{d_i} c \right) \\ &= \frac{\mathbf{e}^\top y}{2} + \min_{\{i=1,\dots,n : d_i > 0\}} \left(\left\| y - \frac{y_i}{d_i} d \right\|_1 + \frac{y_i}{d_i} \left(c - \frac{\mathbf{e}^\top d}{2} \right) \right) \end{aligned} \tag{2}$$

where, here and henceforth, $\mathbf{e} := (1, \dots, 1)^\top$ and $\|\cdot\|_1$ is the usual vector 1-norm. Note that this curve is a generalization of the notion of Lorenz curves from majorization theory (p. 5 [28]). While the condition $d_i > 0$ from the minimum in (2) is redundant for now, it allows us to generalize the definition of $\text{th}_{d,y}$ to the zero-temperature case where some of the entries of d vanish; cf. Remark 2 below.

Remark 1. It is clear from the definition that the thermomajorization curve is invariant under permutations in the sense that $\text{th}_{\underline{\sigma}d, \underline{\sigma}y} \equiv \text{th}_{d,y}$ for all $\sigma \in S_n$. Here, given some permutation $\sigma \in S_n$, we write $\underline{\sigma}$ for the corresponding permutation matrix $\sum_{i=1}^n e_i e_{\sigma(i)}^\top$. In particular, the identities $\underline{\sigma} \circ \underline{\tau} = \underline{\tau} \cdot \underline{\sigma}$, $(\underline{\sigma}x)_j = x_{\sigma(j)}$, and $(\underline{\sigma}A\underline{\tau})_{jk} = A_{\sigma(j)\tau^{-1}(k)}$ hold for all $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$, $j, k = 1, \dots, n$, and all $\sigma, \tau \in S_n$.

Now, the precise connection between thermomajorization and (quasi-classical) state transfers is summarized in the following well-known result. Given any $x, y \in \mathbb{R}^n$, the following statements are equivalent (Proposition 1 [27]).

- There exists a Gibbs-stochastic matrix A such that $Ay = x$. We denote this by $x \prec_d y$.
- $\mathbf{e}^\top x = \mathbf{e}^\top y$ and $\text{th}_{d,x}(c) \leq \text{th}_{d,y}(c)$ for all $c \in [0, \mathbf{e}^\top d]$.
- $\mathbf{e}^\top x = \mathbf{e}^\top y$ and $\text{th}_{d,x}(\sum_{i=1}^j d_{\tau(i)}) \leq \text{th}_{d,y}(\sum_{i=1}^j d_{\tau(i)})$, i.e.,

$$\sum_{i=1}^j x_{\tau(i)} \leq \min_{\{i=1, \dots, n: d_i > 0\}} \left(\left(\sum_{j=1}^n \max \left\{ y_j - \frac{y_i}{d_i} d_j, 0 \right\} \right) + \frac{y_i}{d_i} \left(\sum_{k=1}^j d_{\tau(k)} \right) \right)$$

for all $j = 1, \dots, n - 1$, where $\tau \in S_n$ is any permutation such that $\frac{x_{\tau(1)}}{d_{\tau(1)}} \geq \dots \geq \frac{x_{\tau(n)}}{d_{\tau(n)}}$.

- $\|x - td\|_1 \leq \|y - td\|_1$ for all $t \in \mathbb{R}$.
- $\mathbf{e}^\top x = \mathbf{e}^\top y$ and $\|d_i x - y_i d\|_1 \leq \|d_i y - y_i d\|_1$ for all $i = 1, \dots, n$.

These criteria slightly simplify for probability vectors $x, y \in \mathbb{R}^n$ (e.g., containing the spectrum of any two quantum states), i.e., for vectors $x, y \geq 0$ such that $\mathbf{e}^\top x = \mathbf{e}^\top y = 1$: in this case, the “thermomajorization curve” criterion reduces to $\text{th}_{d,x}(c) \leq \text{th}_{d,y}(c)$ for all $c \in [0, \mathbf{e}^\top d]$ (resp. all $c \in \{\sum_{i=1}^j d_{\tau(i)} : j = 1, \dots, n - 1\}$, where τ sorts $\frac{x}{d}$ non-increasingly). Starting from the thermomajorization curves, there even exists an algorithm to find a Gibbs-stochastic matrix that implements the state transition in question [22]. Nonetheless, the reason that thermomajorization curves are equivalent to conditions using the 1-norm is their fundamental link by means of the Legendre transformation; recall that, given a convex function $f : D \rightarrow \mathbb{R}$ on a connected domain $D \subseteq \mathbb{R}$, its Legendre–Fenchel transformation $f^* : D^* \rightarrow \mathbb{R}$ is defined to be $f^*(p) := \sup_{x \in D} (px - f(x))$ for all $p \in D^* := \{p \in \mathbb{R} : \sup_{x \in D} (px - f(x)) < \infty\}$ (Chapter 3.3 [29]). Indeed, for all non-zero $d \geq 0$, all $y \in \mathbb{R}^n$, and all $t \in \mathbb{R}$, one has

$$2(-\text{th}_{d,y})^*(t) = \mathbf{e}^\top y + t(\mathbf{e}^\top d) + \|y + td\|_1 \tag{3}$$

This readily follows from the definition of the Legendre transformation together with the fact that $\text{th}_{d,y+td}(c) = \text{th}_{d,y}(c) + ct$ for all $c \in [0, \mathbf{e}^\top d]$, as well as the fact that the maximum of any thermomajorization curve equals the sum over all non-negative entries of the initial state (Lemma 13 (iii) [27]). With this, the equivalence of the corresponding characterizations of d -majorization is due to the fact that the Legendre transformation is an involution that respects (more precisely, reverses) the order (Theorem 4.2.1 [29]). While these conditions are mathematically equivalent, in practice, it is often easier to prove the results using the curves $\text{th}_{d,y}$. This empirical observation will also be supported by the main part of this paper.

Remark 2 (Thermomajorization for Zero Temperature). *The above conditions for thermomajorization remain well defined even if some entries of d vanish, assuming that $\text{th}_{d,y}$ is defined via (2). Physically, this relates to the temperature being zero, in which case all entries of d that do not correspond to the zero-point energy of the system vanish. If this lowest energy level with corresponding index j is non-degenerate, then $d = d(T) \rightarrow e_j$ as $T \rightarrow 0^+$. Interestingly, the above characterization of thermomajorization remains valid in this regime. For proof, we refer to Appendix A. This explains the known continuity problems of the associated polytope (cf. Example 3 in [27]). If some of the d_i vanish, so do the corresponding inequalities, meaning that there are less restrictions. Hence, the polytope can only become larger in this case.*

The importance of the 1-norm conditions is that they lead to a remarkable characterization of $M_d(y)$ from (1); cf. (Theorem 10 [27]):

$$M_d(y) = \{x \in \mathbb{R}^n : \mathbf{e}^\top x = \mathbf{e}^\top y \text{ and } \forall_{m \in \{0,1\}^n} m^\top x \leq \text{th}_{d,y}(m^\top d)\} \quad (4)$$

The geometric interpretation of each of these inequalities is that every binary vector $m \in \{0,1\}^n \setminus \{0, \mathbf{e}\}$ is the normal vector to a halfspace that limits $M_d(y)$. The location of said halfspace is determined by the value $\text{th}_{d,y}(m^\top d)$ and thus by the thermomajorization curve. Note that the orientation of these halfspaces is universal, i.e., they are independent of any of the system parameters; subsequently, y, d only influence the *location* of the faces. Another means of expressing this is to say that thermomajorization polytopes are obtained by shifting the faces of a classical majorization polytope. Note that this can lead to the situation where some of these halfspace conditions become redundant. This description of $M_d(y)$ has been used to prove continuity of the map $(d, P) \mapsto M_d(P)$, where $d > 0$ and P is from the collection of non-empty compact sets in \mathbb{R}^n equipped with the Hausdorff metric (Theorem 12 (ii) [27]). Alternatively, this result can be obtained from the continuity of the set of Gibbs-stochastic matrices in H_0 and $T > 0$ (Theorem 5.1 [30]).

Now, writing the bounded set $M_d(y)$ as the solution to finitely many inequalities shows that it is a convex polytope, i.e., it ultimately can be written as the convex hull of finitely many points [31]. Rather than merely being an abstract result, these extreme points have an analytic form. Moreover, halfspaces becoming redundant leads to the coalescence of extreme points; more details on this are given in Section 3.2.

2.2. Transportation Polytopes

It emerges that Gibbs-stochastic matrices are closely related to transportation matrices, which are at the core of the well-studied field of transportation theory [32–35]. It appears that, so far, this has been very much overlooked: this notion does not even appear in the standard work on majorization by Marshall et al. [28], and the only papers from the quantum thermodynamics literature (that we are aware of) that have used results from transportation theory are ones by Mazurek et al. [17,36].

Assuming finite domains, *transportation matrices* are non-negative matrices with fixed column and row sums. More precisely, these are matrices $A \in \mathbb{R}_+^{m \times n}$ such that $A\mathbf{e} = r$ and $c^\top = \mathbf{e}^\top A$ for some $r \in \mathbb{R}^m, c \in \mathbb{R}^n$ with $\mathbf{e}^\top r = \mathbf{e}^\top c$. The collection of all such matrices is called the *transportation polytope* and is often denoted by $T(r, c)$. As already observed by Hartfiel [30], the connection to our setting then is the following. For non-zero temperatures (i.e., $d > 0$), there is a one-to-one correspondence between (the extreme points of) the Gibbs-stochastic matrices and (the extreme points of) the symmetric transportation polytope $T(d, d)$ by means of the isomorphism $X \mapsto X \text{diag}(d)$. From our point of view, pursuing this approach may seem counter-intuitive because the geometry of the Gibbs-stochastic matrices is known to be more complicated than the thermomajorization polytope. Already, in three dimensions, the number of extreme points of the Gibbs-stochastic matrices depends on the temperature of the bath (Figure 1 [36]), i.e., on certain relations between the entries of d (Appendix A [27]). However, drawing this connection grants access to powerful tools from combinatorics and graph theory. Roughly speaking, there is a relation between the extreme points of $T(d, d)$ and the spanning trees of the associated bipartite graphs. We do not provide details of the underlying techniques (instead, we refer to [17]); rather, we adopt useful notions from this field and adapt them to thermomajorization as well as the associated polytope. For this, our starting point is a paper by Loewy et al. [37], where conditions on the vector d that classify certain features of (the polytope of) Gibbs-stochastic matrices are identified. We present these conditions—which Loewy et al. simply call “property (a)” and “property (b)” (To be precise, while Definition 1 (i) is the same as their “property (a)”, Definition 1 (ii) strengthens “property (b)”. However, (i) and (ii) together are equivalent to property (a) and property (b) combined.)—in the following definition.

Definition 1. Given $d \in \mathbb{R}_{++}^n$, define a map $\mathcal{D} : \mathcal{P}(\{1, \dots, n\}) \rightarrow [0, \infty)$ on the power set of $\{1, \dots, n\}$ via $\mathcal{D}(I) := \sum_{i \in I} d_i$.

- (i) We say that d is well structured if, for all $I, J \subseteq \{1, \dots, n\}$, $|I| < |J|$, one has $\mathcal{D}(I) < \mathcal{D}(J)$.
- (ii) We call d stable if \mathcal{D} is injective.

Some remarks on these notions are in order. By definition, d is stable if summing up two sets of entries of d only yields the same result if the entries coincide in the first place. In particular, stability implies that d is non-degenerate. Note that, for non-degenerate systems, stability is a generic property as only finitely many temperatures give rise to unstable Gibbs states. On the other hand, d is well structured if summing up $k - 1$ arbitrary entries of d always yields less than summing up any k entries of d . Interestingly, this notion is fully captured by the inequality

$$\sum_{i=1}^{\lceil \frac{n}{2} \rceil - 1} d_i^\downarrow < \sum_{i=n - \lceil \frac{n}{2} \rceil + 1}^n d_i^\downarrow \tag{5}$$

where $d_i^\downarrow := (d^\downarrow)_i$ is the i -th largest component of d , in the sense that d is well structured if and only if (5) holds. One way to observe this is to first reduce the well-structuredness of d to a family of inequalities $\sum_{i=1}^\alpha d_i^\downarrow < \sum_{i=n-\alpha+1}^n d_i^\downarrow$, $\alpha = 1, \dots, n - 1$ (where α plays the role of $|I|$ from Definition 1), and, in a second step, to realize that the inequality corresponding to $\alpha = \lceil \frac{n}{2} \rceil - 1$ implies all other ones. Nonetheless, from (5), one sees that well-structuredness is a high-temperature phenomenon. Given the energies of the system $E_1 \leq \dots \leq E_n$ (without loss of generality, $E_1 < E_n$ to avoid the trivial case of full degeneracy), there exists a unique critical temperature $T_c \geq 0$ such that $\sum_{i=1}^{\lceil \frac{n}{2} \rceil - 1} e^{-E_i/T_c} = \sum_{i=n - \lceil \frac{n}{2} \rceil + 1}^n e^{-E_i/T_c}$, and the corresponding Gibbs vector is well structured if and only if $T > T_c$. One way to prove this is to examine the auxiliary function $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined via

$$\phi(T) := \frac{\sum_{i=n - \lceil \frac{n}{2} \rceil + 1}^n e^{-E_i/T}}{\sum_{i=1}^{\lceil \frac{n}{2} \rceil - 1} e^{-E_i/T}}$$

and to see that $\lim_{T \rightarrow 0^+} \phi(T) \in [0, 1]$, $\lim_{T \rightarrow \infty} \phi(T) > 1$, and $\phi'(T) > 0$ for all $T > 0$; the latter inequality follows at once from the readily verified expression

$$\phi'(T) = \left(T \sum_{i=1}^{\lceil \frac{n}{2} \rceil - 1} e^{-E_i/T} \right)^{-2} \sum_{i=n - \lceil \frac{n}{2} \rceil + 1}^n \sum_{j=1}^{\lceil \frac{n}{2} \rceil - 1} (E_i - E_j) e^{-(E_i + E_j)/T}$$

together with the observation that $E_i - E_j \geq 0$, because $i > j$, and even $E_n - E_1 > 0$ by assumption. Thus, by the intermediate value theorem, there exists a unique $T_c \geq 0$ such that $\phi(T_c) = 1$, and $\phi(T) > 1$ (which is equivalent to (5)) holds if and only if $T > T_c$.

Regardless, the notion of stable Gibbs states, as well as the fact that the well-structuredness of the Gibbs state appears if (and only if) the temperature exceeds a critical value, can be easily visualized via the standard simplex and the ordered Weyl chamber; cf. Figure 1.

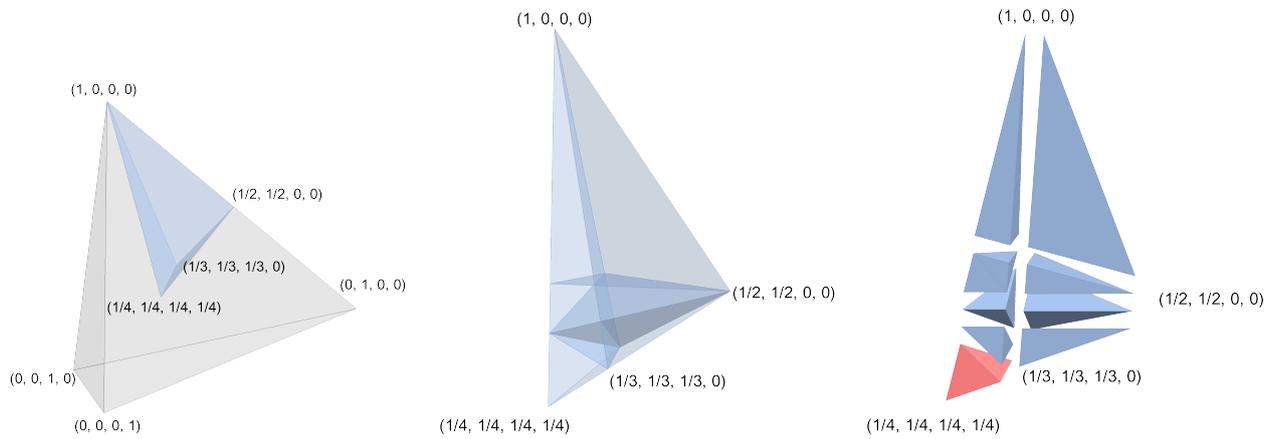


Figure 1. Illustration of stability and well-structuredness. Consider $d \in \mathbb{R}_{++}^4$, where, without loss of generality, $\mathbf{e}^\top d = 1$. Then, d lies in the relative interior of the standard simplex shown on the left. By reordering its entries in a non-increasing fashion—cf. Remark 1—we can assume that d lies in the ordered Weyl chamber shown in the middle. The unstable points are composed of the walls of the Weyl chamber as well as five planes intersecting the Weyl chamber. These planes cut the Weyl chamber into nine subchambers, and the one that includes the maximally mixed state $\mathbf{e}/4$ (highlighted in red on the right) contains exactly the well-structured Gibbs vectors.

Now, based on these notions, Loewy et al. were able to prove the following. Given any $d \in \mathbb{R}_{++}^n$, the well-structuredness of d is equivalent to every extreme point of $s_d(n)$ (i.e., the Gibbs-stochastic matrices) being invertible (Theorem 3.1 [37]); the number of extreme points of $s_d(n)$ is maximal (when d is taken as a parameter) if and only if d is well structured and stable (Theorems 5.1 and 6.1 [37]); and, as a lower and upper bound to the number of extreme points of $s_d(n)$, they found $(n - 1)!n^{n-2}$ and $n!n^{n-2}$, respectively. (Actually, they explained how to calculate an even better lower bound that grows asymptotically as $(n - 1)!n^{n-2} \log n$ but it cannot be written down as well as the bound $(n - 1)!n^{n-2}$.) The bound $(n - 1)!n^{n-2}$ is a substantial improvement over the lower bound $n!$ as first proven by Perfect and Mirsky [38].

3. Results

As we will see, the notions of well-structured and stable Gibbs states are key to answering fundamental questions in quantum thermodynamics. Not only do this definition and the associated, already known results (e.g., on extreme points of the Gibbs-stochastic matrices) carry over to our setting; this language is even suitable to solve seemingly unrelated problems in quantum thermodynamics and, subsequently, allows us to uncover new connections. Consequently, this section will feature two types of results: first, ones that are, at most, superficially concerned with the geometry of the thermomajorization polytope (Section 3.1), followed by some results on the geometric quantities (e.g., extreme points) of said polytope; cf. Section 3.2. The former can be seen as the general principles underlying quantum thermodynamics, while the latter are more state-dependent and more explicit in nature.

3.1. Cyclic State Transfers and Subspaces in Equilibrium

An overarching framework for this section is given by the notion of catalysis. There, in the most constrained form, one calls a state transition $\rho \mapsto \omega$ *strictly catalytic* if there exists an ancilla in state ω_C as well as an “allowed” operation Φ on the new overall system such that $\Phi(\rho \otimes \omega_C) = \omega \otimes \omega_C$ (although there are also “weaker” versions of catalytic transformations; cf. the review article [39]). For thermal operations in the quasi-classical realm, strict catalysis reduces to state transfers $x \otimes z \prec_d y \otimes z$.

The idea behind such processes is of course that the catalyst ω_C , resp. z , undergoes a cyclic process in order to be returned (uncorrelated and) unchanged. This raises fundamen-

tal questions such as, e.g., which cyclic processes are even possible in our thermodynamic framework, which properties such processes have, etc. This is what our first result concerns: in physical terms, it states that *global* cyclic thermodynamic processes are impossible for almost all temperatures (at least without access to external resources). While Theorem 1 below is concerned with two-step processes, as an immediate corollary, one obtains an analogous result for time-continuous cyclic processes. The precise statement here is that the impossibility of non-trivial cyclic processes is in one-to-one correspondence to the notion of stable Gibbs vectors.

Theorem 1. *Given $d \in \mathbb{R}_{++}^n$, the following statements are equivalent:*

- (i) d is stable.
- (ii) Given any $x, y \in \mathbb{R}^n$, if $x \prec_d y \prec_d x$, then $x = y$.

Proof. “(ii) \Rightarrow (i)”: We will prove this direction by contraposition, i.e., given d not stable, we construct x, y with $x \neq y$ such that $x \prec_d y \prec_d x$. Indeed, assume, to the contrary, that there exist $I, J \subseteq \{1, \dots, n\}$, $I \neq J$ such that $\sum_{i \in I} d_i = \sum_{j \in J} d_j$. Define $x := \sum_{i \in I} d_i e_i$, $y := \sum_{j \in J} d_j e_j$ and note that $I \neq J$ implies $x \neq y$. We claim that $x \prec_d y \prec_d x$. Recalling Section 2.1, we prove this, equivalently, by showing that $\|x - td\|_1 = \|y - td\|_1$ for all $t \in \mathbb{R}$:

$$\begin{aligned} \|x - td\|_1 &= \sum_{i \in I} |d_i - td_i| + \sum_{i \in \{1, \dots, n\} \setminus I} |0 - td_i| \\ &= \left(\sum_{i \in I} d_i \right) |1 - t| + \left(\mathbf{e}^\top d - \sum_{i \in I} d_i \right) |t| \\ &= \left(\sum_{j \in J} d_j \right) |1 - t| + \left(\mathbf{e}^\top d - \sum_{j \in J} d_j \right) |t| = \dots = \|y - td\|_1 \end{aligned}$$

“(i) \Rightarrow (ii)”: The idea of this part of the proof is that any vectors $x \prec_d y \prec_d x$ induce the same thermomajorization curve, and—because d is stable—applying this to the points where the curves change slope allows us to conclude $x = y$. More precisely, assume that d is stable and let $x, y \in \mathbb{R}^n$ be given such that $x \prec_d y \prec_d x$. This implies $\mathbf{e}^\top x = \mathbf{e}^\top y$ and, more importantly, $\text{th}_{d,x}(c) = \text{th}_{d,y}(c)$ for all $c \in [0, \mathbf{e}^\top d]$. However, these are piecewise linear functions characterized by elbow points, so, in particular, the points where $\text{th}_{d,x}$, $\text{th}_{d,y}$ have a (non-trivial) change in slope coincide (recall Section 2.1). To be more precise, there exist $k \in \{1, \dots, n\}$ and sets $I_1^x, \dots, I_{k-1}^x, I_1^y, \dots, I_{k-1}^y \in \mathcal{P}(\{1, \dots, n\})$ such that

- $\emptyset \neq I_1^x \subsetneq I_2^x \subsetneq \dots \subsetneq I_{k-1}^x \subsetneq \{1, \dots, n\}$ and $\emptyset \neq I_1^y \subsetneq I_2^y \subsetneq \dots \subsetneq I_{k-1}^y \subsetneq \{1, \dots, n\}$.
- $\text{th}_{d,x}$ changes slope precisely at the inputs $\{\sum_{i \in I_l^x} d_i : l = 1, \dots, k - 1\}$, and $\text{th}_{d,y}$ changes slope precisely at $\{\sum_{i \in I_l^y} d_i : l = 1, \dots, k - 1\}$.
- $\text{th}_{d,x}(\sum_{i \in I_l^x} d_i) = \sum_{i \in I_l^x} x_i$ and $\text{th}_{d,y}(\sum_{i \in I_l^y} d_i) = \sum_{i \in I_l^y} y_i$ for all $l = 1, \dots, k - 1$.

However, because the changes in slope coincide, this (due to $d > 0$) shows $\sum_{i \in I_l^x} d_i = \sum_{i \in I_l^y} d_i$. By assumption, d is stable, so we obtain $I_l^x = I_l^y$ for all $l = 1, \dots, k - 1$. In particular,

$$\sum_{i \in I_l^x} x_i = \text{th}_{d,x} \left(\sum_{i \in I_l^x} d_i \right) = \text{th}_{d,y} \left(\sum_{i \in I_l^x} d_i \right) = \text{th}_{d,y} \left(\sum_{i \in I_l^y} d_i \right) = \sum_{i \in I_l^y} y_i = \sum_{i \in I_l^x} y_i$$

for all $l = 0, \dots, k$ when defining $I_0^x := \emptyset =: I_0^y$ and $I_k^x := \{1, \dots, n\} =: I_k^y$. Now that we have taken care of the points where the curves change slope, all that remain are the points in between. Consider any $l = 1, \dots, k$. Because $\text{th}_{d,y}$ is affine linear on $[\sum_{i \in I_{l-1}^x} d_i, \sum_{i \in I_l^x} d_i]$ and the slope of $\text{th}_{d,y}$ at “the” increment d_i is given by $\frac{y_i}{d_i}$, there exists $c_{y,l} \in \mathbb{R}$ such that $y_i = c_{y,l} \cdot d_i$ for all $i \in I_l^x \setminus I_{l-1}^x$; one argues analogously for $\text{th}_{d,x}$ and obtains a constant $c_{x,l}$. Note that we use the stability of d here: the length $\sum_{i \in I_l^x \setminus I_{l-1}^x} d_i$ of the interval

$[\sum_{i \in I_{l-1}^x} d_i, \sum_{i \in I_l^x} d_i]$ can only come from adding up $\{d_i : i \in I_l^x \setminus I_{l-1}^x\}$. Now, because $\text{th}_{d,x}$ and $\text{th}_{d,y}$ coincide, by assumption, we find that $c_{x,l} = c_{y,l} (=: c_l)$ and thus

$$x = \sum_{i=1}^n x_i e_i = \sum_{l=1}^k \sum_{i \in I_l^x \setminus I_{l-1}^x} x_i e_i = \sum_{l=1}^k \sum_{i \in I_l^x \setminus I_{l-1}^x} \underbrace{c_l d_i}_{=y_i} e_i = y. \quad \square$$

Of course, this result does not prohibit *local* cyclic processes, i.e., thermodynamic processes where only a subsystem returns to its original state at the end (so, precisely, catalysis). What Theorem 1 does assert, however, is that, in general, a quasi-classical cyclic process (modeled by thermal operations) that is *not* local has to use up some external resource along the way.

For the remainder of this section, our focus lies on (states on) subspaces that are “in equilibrium”. The motivation behind this notion is that if a state restricted to some subspace is a multiple of the Gibbs state, then all thermal operations on that subspace act trivially on it. Indeed, given a state x and a subspace P such that $x|_P$ is a multiple of the Gibbs vector, then any Gibbs-stochastic matrix that acts non-trivially only on P —i.e., it is of the form $A = A_P \oplus \mathbf{1}_{P^\perp}$ —necessarily leaves x invariant. This of course generalizes to subsystems of coupled systems by choosing P appropriately. The precise definition reads as follows.

Definition 2. Let $d \in \mathbb{R}_{++}^n$, $y \in \mathbb{R}^n$. We say that a subset $P \subseteq \{1, \dots, n\}$, $|P| > 1$ of the system’s energy levels is in equilibrium if $\frac{y_i}{d_i} = \frac{y_j}{d_j}$ for all $i, j \in P$. On the other hand, if no such subset P satisfies this condition, we say that the system is in total non-equilibrium.

In this language, our next result states that, regardless of whether there is a subspace in equilibrium (as long as the full system is not), every such subspace can be brought out of equilibrium by means of d -stochastic matrices. This, in particular, applies to catalytic state transfers: if, for example, a system is in equilibrium, then any catalyst (which itself is not in the Gibbs state) allows for bringing arbitrary energy levels of the original system out of equilibrium. The precise statement is derived via the dimension of the thermomajorization polytope and reads as follows.

Theorem 2. Let $d \in \mathbb{R}_{++}^n$, $y \in \mathbb{R}^n$. The following statements are equivalent.

- (i) $M_d(y)$ is not singular, i.e., $M_d(y)$ consists of more than merely y .
- (ii) y is not a multiple of d .
- (iii) The dimension of the convex polytope $M_d(y)$ is maximal, i.e., its dimension is $n - 1$, which is equal to the dimension of the standard simplex Δ^{n-1} of all n -dimensional probability vectors.

In particular, if there exists a subset $P \subsetneq \{1, \dots, n\}$, $|P| > 1$ in equilibrium (i.e., $y_i/d_i = y_j/d_j$ for all $i, j \in P$), then there exists $z \in M_d(y)$ such that z is in total non-equilibrium.

Proof. “(iii) \Rightarrow (i)”: Trivial. “(i) \Rightarrow (ii)”: Obvious via contrapositive. “(ii) \Rightarrow (iii)”: The dimension of $M_d(y)$ is trivially upper bounded by $n - 1$ as it is a subset of the $n - 1$ -dimensional standard simplex (Corollary 17 [27]). For the converse, we argue by contraposition: if the dimension is strictly less than $n - 1$, then there must exist a condition in (4) that is an equality (Chapter 8.2 [40]). More precisely, there must exist $m \in \{0, 1\}^n$, $0 < \mathbf{e}^\top m < n$ and $c \in \mathbb{R}$ such that $m^\top x = c$ for all $x \in M_d(y)$. First, we determine c . Let $\sigma \in S_n$ be any permutation such that σm is sorted non-increasingly, i.e., $\sigma m = (1, \dots, 1, 0, \dots, 0)^\top$. We know that $E_{d,y}(\sigma) \in M_d(y)$ so

$$c = m^\top E_{d,y}(\sigma) = \sum_{j=1}^{\mathbf{e}^\top m} (E_{d,y}(\sigma))_{\sigma(j)} = \text{th}_{d,y} \left(\sum_{j=1}^{\mathbf{e}^\top m} d_{\sigma(j)} \right) = \text{th}_{d,y}(m^\top d).$$

The final step is to evaluate the condition $m^\top x = \text{th}_{d,y}(m^\top d)$ at a multiple of d : because $\frac{de^\top}{e^\top d} \in s_d(n)$, one has $\frac{de^\top}{e^\top d}y = \frac{e^\top y}{e^\top d}d \in M_d(y)$. Therefore,

$$\begin{aligned} \text{th}_{d,y}(m^\top d) = c &= m^\top \left(\frac{e^\top y}{e^\top d} d \right) = \frac{e^\top d - m^\top d}{e^\top d} \cdot 0 + \frac{m^\top d}{e^\top d} e^\top y \\ &= \frac{e^\top d - m^\top d}{e^\top d - 0} \text{th}_{d,y}(0) + \frac{m^\top d - 0}{e^\top d - 0} \text{th}_{d,y}(e^\top d). \end{aligned} \tag{6}$$

Because $\text{th}_{d,y}$ is concave and because $m^\top d \in (0, e^\top d)$ by the assumptions on m and d , Lemma A3 (iv) (Appendix B) shows that (6) can only hold if $\text{th}_{d,y}$ is linear. However, the latter is equivalent to y being a multiple of d as, by definition, the slopes of $\text{th}_{d,y}$ are given by y_j/d_j .

Finally, the additional statement follows at once from the following two facts: (a) the collection of all vectors in total non-equilibrium is dense in Δ^{n-1} , and (b) $M_d(y)$ contains an interior point with regard to the hyperplane $W_e := \{z \in \mathbb{R}^n : e^\top z = 1\}$. While (b) is due to (iii), for (a), note that, given any $P \subsetneq \{1, \dots, n\}$, $|P| > 1$ the set of vectors $y \in W_e$ in equilibrium (with regard to P) forms a lower-dimensional subspace of W_e . In particular, this set is nowhere dense (i.e., its closure has an empty interior [41]), which continues to hold when taking the union over all (finitely many) such P . However, this, in particular, implies that the complement of this set—i.e., the collection of all vectors in total non-equilibrium—is dense in W_e . This concludes the proof. \square

Returning to the language of subspaces in equilibrium, note that this result is an existence result and does not indicate anything about the potential “amplitude” of such transfers. Mathematically, Theorem 2 complements the old result of Hartfiel that the dimension of $s_d(n)$ for all $d > 0$ is $(n - 1)^2$ (Theorem 3.1 [30]).

Remark 3. The assumption $d > 0$ in Theorem 2 is necessary as hinted at by the fact that $\dim(s_d(n))$ for a general $d \in \mathbb{R}_+^n$ depends on the number of zeros in d (Theorem 3.2 [30]). This is well illustrated by the simple example $d = (1, 1, 0)^\top$, $y = (1, 0, 0)^\top$ because then $M_d(y) = \{(c, 1 - c, 0)^\top : c \in [0, 1]\}$, which is not two- but only one-dimensional.

3.2. Extreme Points of the Thermomajorization Polytope

Let us stress that—until now—the concept of subspaces in equilibrium (and hence Theorem 2) has been logically independent from the notions of stability and well-structuredness. This missing connection will be established below, where well-structured states and subspaces in equilibrium will be linked via the geometric properties of the thermomajorization polytope—in particular, the number of its extreme points. (An analytic form of the extreme points of $M_d(y)$ has appeared independently in the physics [8,15] and the mathematics [27] literature.) The key mathematical object for doing so is the extreme point map $E_{d,y}$, which is defined as follows.

Definition 3. Let $d \in \mathbb{R}_{++}^n$, $y \in \mathbb{R}^n$. Define the extreme point map $E_{d,y}$ on the symmetric group S_n via

$$E_{d,y} : S_n \rightarrow \mathbb{R}^n, \quad \sigma \mapsto \left(\text{th}_{d,y} \left(\sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)} \right) - \text{th}_{d,y} \left(\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)} \right) \right)_{j=1}^n.$$

Equivalently, $(\underline{\sigma} E_{d,y}(\sigma))_j = (E_{d,y}(\sigma))_{\sigma(j)} = \text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) - \text{th}_{d,y}(\sum_{i=1}^{j-1} d_{\sigma(i)})$, where $\underline{\sigma}$ is the permutation matrix corresponding to σ as defined above (Remark 1).

As the name suggests, for all $d \in \mathbb{R}_{++}^n$ and $y \in \mathbb{R}^n$, the image of $E_{d,y}$ equals the set $\text{ext}(M_d(y))$ of extreme points of $M_d(y)$, and thus $M_d(y) = \text{conv}\{E_{d,y}(\sigma) : \sigma \in S_n\}$ (Theorem 16 [27]). It should be noted that the extreme point property also manifests in

the thermomajorization curves, relating to the concept of *tight thermomajorization*. Given any $y, z \in \mathbb{R}_+^n$, the point z is an extreme point of $M_d(y)$ if and only if all elbow points of $\text{th}_{d,z}$ lie on the curve $\text{th}_{d,y}$ (Theorem 2 [17]). Moreover, there is a strong connection between the extreme points of $M_d(y)$ and a special class of extreme points of the Gibbs-stochastic matrices, which we will elaborate on at the end of this section. Moreover, Section 4.1 below presents a step-by-step calculation of the extreme point map and explains how it, equivalently, can be computed graphically using the thermomajorization curve.

Our focus for now, however, lies on the properties of the map $E_{d,y}$ and geometric aspects of $M_d(y)$. From Definition 3, it is clear that the maximal number of extreme points of $M_d(y)$ is $n! = |S_n|$; if there are strictly fewer than $n!$ extreme points, we say that the polytope is *degenerate*. The goal of this section is to prove the following result, which states that the degeneracy of the polytope is a “witness” for subspaces in equilibrium, assuming that d is well structured (equivalently, assuming large enough temperatures; cf. Section 2.2).

Theorem 3. *Let $d \in \mathbb{R}_{++}^n, y \in \mathbb{R}^n$. If $E_{d,y}$ is not injective, i.e., $|\text{ext}(M_d(y))| < n!$, then at least one of the following holds:*

- (i) y has a subspace which is in equilibrium with respect to d .
- (ii) d is not well structured, i.e., the temperature is below the critical value $T \leq T_c$; cf. Section 2.2.

Note that in the example of Section 4.1, degenerate extreme points occur, and the degeneracy stems from the fact that there exists a subspace in equilibrium. This can be shown generally as specifying the preimage of y under the extreme point map $E_{d,y}$ is found to be straightforward.

Lemma 1. *Let $d \in \mathbb{R}_{++}^n, y \in \mathbb{R}^n$, and $\sigma \in S_n$ be given. One has $E_{d,y}(\sigma) = y$ if and only if $\frac{y_{\sigma(1)}}{d_{\sigma(1)}} \geq \dots \geq \frac{y_{\sigma(n)}}{d_{\sigma(n)}}$.*

Proof. Assume that $\sigma \in S_n$ satisfies $E_{d,y}(\sigma) = y$ so $(E_{d,y}(\sigma))_{\sigma(j)} = y_{\sigma(j)}$ for all $j = 1, \dots, n$. By definition of $E_{d,y}$, this means $\text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) = \text{th}_{d,y}(\sum_{i=1}^{j-1} d_{\sigma(i)}) + y_{\sigma(j)}$, which, by induction, is equivalent to $\text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) = \sum_{i=1}^j y_{\sigma(i)}$ for all $j = 1, \dots, n - 1$. However, by Lemma A2 (Appendix B), this holds if and only if $y_{\sigma(1)}/d_{\sigma(1)} \geq \dots \geq y_{\sigma(n)}/d_{\sigma(n)}$. \square

Clearly, then, the extreme point y is degenerate (in the sense that it has multiple preimages under $E_{d,y}$) if and only if there exists a subspace that is in equilibrium; cf. Definition 2. The converse, however, is not true. Indeed, the example in Section 4.2 shows that degenerate extreme points can occur even if the system is in total non-equilibrium. Note, however, that, in this example, the vector d is not well structured, indicating a low temperature, as required by Theorem 3.

Note that Lemma 1 relates to the concept of virtual temperatures [6,42–44] as multiple permutations are mapped to y under $E_{d,y}$ if and only if there exist $i, j \in \{1, \dots, n\}$ such that the background temperature equals $T = \frac{E_j - E_i}{\ln(y_i) - \ln(y_j)} =: T_{ij}$ (which is equivalent to $\frac{y_i}{d_i} = \frac{y_j}{d_j}$). In other words, virtual temperatures characterize when another corner of the polytope “crosses” the initial state. This also relates to the notion of passivity: the degeneracy of $M_d(y)$ at temperature T_{ij} corresponding to the transition between E_i and E_j is physical (i.e., $T_{ij} > 0$) only if the initial state $\text{diag}(y)$ is passive, meaning that no work can be extracted via unitary transformations (Section III [43]).

Example 1. *It is worth addressing the case of classical majorization, i.e., $d = \mathbf{e}$ (up to a factor that is of no consequence). We find that $\text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) = \sum_{i=1}^j y_i^\downarrow$ for any $j = 1, \dots, n, \sigma \in S_n$; this recovers $E_{d,y}(\sigma) = \underline{\sigma}^{-1} y^\downarrow$ (Chapter 4, Proposition C.1 [28]). Therefore, degeneracies of $M_{\mathbf{e}}(y)$ correspond to some entries of the initial state coinciding: given any $y \in \mathbb{R}^n$, the number of extreme points of $M_{\mathbf{e}}(y)$ equals $n! / (p_1! \dots p_m!)$, where $(p_i)_{i=1}^m$ denote the multiplicities of y , i.e., $y^\downarrow = \sum_{i=1}^m y_i^\downarrow \sum_{j=1}^{p_i} e_{p_1+\dots+p_{i-1}+j}$ where $y_1^\downarrow > \dots > y_m^\downarrow$.*

This, in fact, implies the “uniformity” of the classical majorization polytope’s degeneracy in the sense that the preimage of each extreme point under $E_{d,y}$ has the same size, which is certainly false for general $d \in \mathbb{R}_{++}^n$; cf. Section 4.1 and, in particular, Table 1. This uniformity is related to the fact that $s_e(n)$ contains all permutation matrices, which yields a group action of S_n on $s_e(n)$. This group action is transitive on the vertices, and hence all vertices are equivalent. This does not hold for general $d \in \mathbb{R}_{++}^n$: the inverse of some invertible element of $s_d(n)$ is again in $s_d(n)$ if and only if it is a permutation matrix (Remark 4.1 [45]).

Table 1. Image of $E_{d,y}$ for $y = (4, 0, 1)^\top$ and $d = (4, 2, 1)^\top$ with intermediate steps.

σ	$\text{th}_{d,y}(d_{\sigma(1)})$	$\text{th}_{d,y}(d_{\sigma(1)} + d_{\sigma(2)})$	$E_{d,y}(\sigma)$
(1 2 3)	4	$\min\{6, 5\} = 5$	$\begin{pmatrix} 4 \\ 5-4 \\ 5-5 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}$
(3 1 2)	1	$\min\{5, 5\} = 5$	$\begin{pmatrix} 5-1 \\ 5-5 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$
(2 3 1)	2	$\min\{3, 5\} = 3$	$\begin{pmatrix} 5-3 \\ 2 \\ 3-2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$
(2 1 3)	2	$\min\{6, 5\} = 5$	$\begin{pmatrix} 5-2 \\ 2 \\ 5-5 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$
(1 3 2)	4	$\min\{5, 5\} = 5$	$\begin{pmatrix} 4 \\ 5-5 \\ 5-4 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$
(3 2 1)	1	$\min\{3, 5\} = 3$	$\begin{pmatrix} 5-3 \\ 3-1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$

Reviewing the construction of extreme points in Section 4.1, and especially the graphical approach, one notices quickly how degeneracies can occur; cf. in particular Remark 5. Indeed, given $d \in \mathbb{R}_{++}^n, y \in \mathbb{R}^n$, there exists $m \in \{0, \dots, n - 1\}$ as well as real numbers $\Delta_0, \dots, \Delta_{m+1}$ such that $0 = \Delta_0 < \Delta_1 < \dots < \Delta_m < \Delta_{m+1} = \mathbf{e}^\top d$ and $\text{th}_{d,y}$ changes slope at an input $x \in (0, \mathbf{e}^\top d)$ if and only if $x = \Delta_k$ for some $k = 1, \dots, m$. The example in Section 4.1 shows that degeneracies occur when two or more intervals of length $d_i, i = 1, \dots, m$ (ordered according to σ) are contained in the same interval (Δ_{k-1}, Δ_k) .

This motivates the following definition. For $\sigma \in S_n, k \in \{1, \dots, m + 1\}$, define

$$I_k^\sigma := \left\{ j \in \{1, \dots, n\} : \left[\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)}, \sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)} \right] \cap (\Delta_{k-1}, \Delta_k) \neq \emptyset \right\}. \tag{7}$$

The elements of I_k^σ are those $j \in \{1, \dots, n\}$ for which the interval of length d_j corresponding to the partition induced by σ intersects (Δ_{k-1}, Δ_k) . In particular, $\bigcup_{k=1}^{m+1} I_k^\sigma = \{1, \dots, n\}$ for all $\sigma \in S_n$. An illustrative example is given in Section 4.4.

The importance of this definition comes from its ability to characterize the image of the extreme point map, as the following result shows.

Proposition 1. *Let $d \in \mathbb{R}_{++}^n, y \in \mathbb{R}^n$ and let $m \in \{1, \dots, n - 1\}$ be the number of changes in the slope of $\text{th}_{d,y}$. Given any $\sigma, \tau \in S_n$, one has $E_{d,y}(\sigma) = E_{d,y}(\tau)$ if and only if $I_k^\sigma = I_k^\tau$ for all $k = 1, \dots, m + 1$.*

Proof. “ \Rightarrow ”: Assume $E_{d,y}(\sigma) = E_{d,y}(\tau)$. Then,

$$\text{th}_{d,y} \left(\sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)} \right) - \text{th}_{d,y} \left(\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)} \right) = \text{th}_{d,y} \left(\sum_{i=1}^{\tau^{-1}(j)} d_{\tau(i)} \right) - \text{th}_{d,y} \left(\sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)} \right) \tag{8}$$

for all $j = 1, \dots, n$; see Definition 3. Now, given $j \in \{1, \dots, n\}$, there are two (non-exclusive) possibilities: either $[\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)}, \sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)}] = [\sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)}, \sum_{i=1}^{\tau^{-1}(j)} d_{\tau(i)}]$, which implies $j \in I_k^\sigma$ if and only if $j \in I_k^\tau$, or the two intervals do not coincide. The latter implies that $\text{th}_{d,y}$ is affine linear on $\text{conv}([\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)}, \sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)}] \cup [\sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)}, \sum_{i=1}^{\tau^{-1}(j)} d_{\tau(i)}])$ by our argument from above. However, this means that both these intervals have to be contained in the same interval $[\Delta_{k-1}, \Delta_k]$; hence, $j \in I_k^\sigma$ and $j \in I_k^\tau$.

“ \Leftarrow ”: Assume by contraposition that $E_{d,y}(\sigma) \neq E_{d,y}(\tau)$ so there exists $j = 1, \dots, n$ such that (8) does not hold. Thus, $J_\sigma := [\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)}, \sum_{i=1}^{\sigma^{-1}(j)} d_{\sigma(i)}]$ has to differ from $J_\tau := [\sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)}, \sum_{i=1}^{\tau^{-1}(j)} d_{\tau(i)}]$ (which is only possible if $\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)} \neq \sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)}$ because the intervals have the same length). Moreover, $\text{th}_{d,y}$ cannot be affine linear on $\text{conv}(J_\sigma \cup J_\tau)$ so there exists $k \in \{1, \dots, m\}$ such that the change in slope Δ_k lies in the interior of $\text{conv}(J_\sigma \cup J_\tau)$.

Now, there are two possible cases (cf. Figure 2): if $\Delta_k \notin J_\sigma \cap J_\tau$, then $\Delta_k \notin J_\sigma$ (or $\Delta_k \notin J_\tau$) so $\text{int}(J_\sigma)$ ($\text{int}(J_\tau)$) is to the left or to the right of Δ_k . Either way, there exists $\tilde{k} \in \{1, \dots, m+1\}$ such that $j \in I_{\tilde{k}}^\tau$ but $j \notin I_{\tilde{k}}^\sigma$ so we are finished. For the second case, assume that $\Delta_k \in J_\sigma \cap J_\tau$ and, without loss of generality, that $\min J_\sigma < \min J_\tau$, i.e., $\sum_{i=1}^{\sigma^{-1}(j)-1} d_{\sigma(i)} < \sum_{i=1}^{\tau^{-1}(j)-1} d_{\tau(i)}$. Therefore, one can find $\alpha \in \{1, \dots, \tau^{-1}(j) - 1\}$ such that $\tau(\alpha) \notin \{\sigma(1), \dots, \sigma(\sigma^{-1}(j) - 1), j\}$. This in turn implies $\sigma^{-1}(\tau(\alpha)) > \sigma^{-1}(j)$, i.e., the interval corresponding to $d_{\tau(\alpha)}$ is to the left of $J_\tau \supseteq J_\tau \cap J_\sigma$ but to the right of $J_\sigma \supseteq J_\sigma \cap J_\tau$. Because $\Delta_k \in J_\sigma \cap J_\tau$, this yields $\tilde{k} \in \{1, \dots, k\}$ such that $\tau(\alpha) \in I_{\tilde{k}}^\tau$ but $\tau(\alpha) = \sigma(\sigma^{-1}(\tau(\alpha))) \notin I_{\tilde{k}}^\sigma$. This concludes the proof. \square

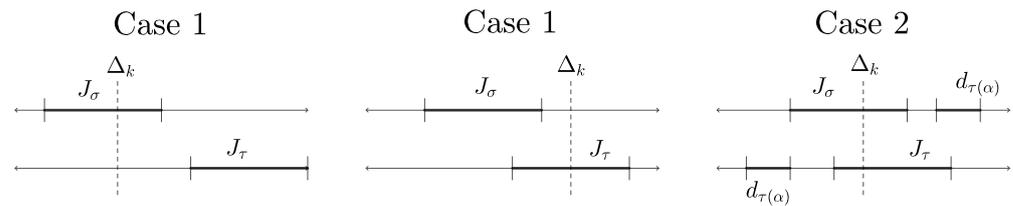


Figure 2. Possible combinations of whether J_σ and J_τ intersect, and where Δ_k lies relative to J_σ, J_τ .

This means that σ and τ yield the same extreme point if and only if they differ exactly by a permutation of the intervals of length $d_i(j)$ such that each interval remains in the same region where $\text{th}_{d,y}$ is affine linear. Note that the case of classical majorization follows from this since, in that case, such permutations differ exactly by the elements of the stabilizer of y . Similarly, the result about preimages of y (Lemma 1) follows from this.

Nonetheless, Proposition 1 gives a simple criterion to check whether the images of different permutations under the extreme point map $E_{d,y}$ coincide or not. In particular, a bound on the degeneracy of any extreme point can be given by how many of the d_i intervals fit into the same $[\Delta_{k-1}, \Delta_k]$ interval. This is related to the bin packing problem in computer science, which (is strongly NP-complete but) admits some reasonable approximations. Another way to look at the above results is that $|\text{ext}(M_d(y))| \leq n! + 1 - |(S_n)_{y/d}|$ by Lemma 1, where $(S_n)_{y/d}$ is the stabilizer of the vector $y/d = (y_i/d_i)_{i=1}^n$ in S_n , and the examples in Sections 4.1 and 4.2 show that this bound is not tight. These examples suggest that improving this bound via an analytic expression is a non-trivial task. Finally, Proposition 1 allows us to prove Theorem 3.

Proof of Theorem 3. Assume that there exist distinct $\sigma, \tau \in S_n$ such that $E_{d,y}(\sigma) = E_{d,y}(\tau)$. Then, by Proposition 1, σ and τ differ by a permutation such that each d -interval remains in the same affine linear region of the thermomajorization curve. If $i \mapsto \frac{y_i}{d_i}$ is injective, this means that there have to exist pairwise distinct $i, j, k \in \{1, \dots, n\}$ such that both the intervals corresponding to d_j, d_k “fit inside” d_i . Therefore, $d_i \geq d_j + d_k$, meaning that d cannot be well structured. \square

Note that the proof of Theorem 3 actually shows that $d_i \geq d_j + d_k$ for some distinct i, j, k , which is a property stronger than the lack of well-structuredness. We wish to stress that while condition (i) of Theorem 3 ensures the degeneracy of $M_d(y)$ (Lemma 1), the lack of well-structuredness of d is not sufficient for $M_d(y)$ to be degenerate. This phenomenon is easy to understand in the graphical representation: even if it holds that $d_i \geq d_j + d_k$ for some distinct i, j, k —and hence d is not well-structured—it might happen that there is no permutation of the intervals that achieves the degeneracy. An example of this is given in Section 4.3.

Let us conclude this section by examining the operator lift. More precisely, due to $M_d(y) = \text{conv}\{Ay : A \in \text{ext}(s_d(n))\}$ (Chapter 14, Section C, C.2. Observations (iii) [28]), Minkowski’s theorem (Theorem 5.10 [46]) shows that, given any extreme point z of $M_d(y)$, there exists an extreme point A of $s_d(n)$ such that $z = Ay$. Now, the obvious question is whether, given some extreme point of $M_d(y)$, there is an easy way to recover one (or every) process that maps the initial state to the point in question. While, given any initial and any final state, there already exists an algorithm to construct a Gibbs-stochastic matrix mapping the former to the latter [22], it emerges that if the final state is an extreme point, then this procedure simplifies considerably.

Definition 4. Given $d \in \mathbb{R}_{++}^n$ and permutations $\sigma, \tau \in S_n$, there exists, for all $j = 1, \dots, n - 1$, a unique $\alpha_j \in \{1, \dots, n\}$ such that $\sum_{i=1}^j d_{\sigma(i)} \in (\sum_{i=1}^{\alpha_j-1} d_{\tau(i)}, \sum_{i=1}^{\alpha_j} d_{\tau(i)})$. Moreover, set $\alpha_0 := 1$ and $\alpha_n := n$. Based on this, define a matrix $A_{\sigma\tau} \in \mathbb{R}_+^{n \times n}$ via

$$(A_{\sigma\tau})_{\sigma(j)\tau(k)} := \begin{cases} (\sum_{i=1}^{\alpha_{j-1}-1} d_{\tau(i)} - \sum_{i=1}^{j-1} d_{\sigma(i)}) \cdot d_{\tau(\alpha_{j-1})}^{-1} & \text{if } k = \alpha_{j-1} < \alpha_j \\ 1 & \text{if } \alpha_{j-1} < k < \alpha_j \\ (\sum_{i=1}^j d_{\sigma(i)} - \sum_{i=1}^{\alpha_j-1} d_{\tau(i)}) \cdot d_{\tau(\alpha_j)}^{-1} & \text{if } \alpha_{j-1} < \alpha_j = k \\ \frac{d_{\sigma(j)}}{d_{\tau(\alpha_j)}} & \text{if } \alpha_{j-1} = \alpha_j = k \\ 0 & \text{else} \end{cases} \quad (9)$$

for all $j, k = 1, \dots, n$.

This object has already appeared in the literature as “ β -permutation” [8] and it coincides with the concept of a “bipolar extremal transportation matrix” [17] (up to the isomorphism $X \mapsto X \text{diag}(d)$ from Section 2.2). The latter name, rightly, suggests that the matrix $A_{\sigma\tau}$ for all $y \in \mathbb{R}^n$, $d \in \mathbb{R}_{++}^n$ and all $\sigma, \tau \in S_n$ is an extreme point of the Gibbs-stochastic matrices (Theorem 1 ff. [17]). However—due to the lower bound $(n - 1)!n^{n-2}$ on the number of extreme points of $s_d(n)$ from Section 2.2—for $n \geq 4$, there must exist extreme points of $s_d(n)$ that are not of the form $A_{\sigma\tau}$ (i.e., they are not a β -permutation). Actually, the lower bound in question shows that, for large n , “almost no” extreme point of $s_d(n)$ is a β -permutation as $(n!)^2 / ((n - 1)!n^{n-2}) \rightarrow 0$ as $n \rightarrow \infty$ (Section IV.B [17]).

Moreover, and more importantly, if τ is chosen such that $\frac{y_{\tau(1)}}{d_{\tau(1)}} \geq \dots \geq \frac{y_{\tau(n)}}{d_{\tau(n)}}$, then $A_{\sigma\tau}$ maps the initial state y to the extreme point $E_{d,y}(\sigma)$, and if y/d is non-degenerate, then $A_{\sigma\tau}$ is the unique Gibbs-stochastic matrix that maps y to $E_{d,y}(\sigma)$; cf. [8], (Lemma 3 [17]). Not only does this yield a simple way to reverse-engineer a process that generates an extreme point in question; it also constitutes an alternative means to evaluate the extreme point map $E_{d,y}$ from Definition 3.

Remark 4. Given a permutation σ , a matrix $A_{\sigma\tau}$ (stored in a sparse matrix format) can be constructed algorithmically in, at most, $O(n \log n)$ steps using Definition 4, as the limiting step is to find an appropriate permutation τ . Any algorithm that computes a process matrix A for an arbitrary state transfer, including the one given in [22], must have the worst time complexity of at least $\Omega(n^2)$, since A is dense in general. Hence, the structure of the $A_{\sigma\tau}$ leads to an improved runtime for the special case where the final state is extremal.

Now, our main contribution to this concept reads as follows. While Definition 4 (which matches the definition given by [8]) as well as Mazurek’s construction for biplanar extremal transportation matrices appear rather convoluted, we will present a very simple construction of this matrix in Section 4.5 below. The only required step is to compare the sets $\{\sum_{i=1}^j d_{\sigma(i)} : j = 1, \dots, n\}$ and $\{\sum_{i=1}^j d_{\tau(i)} : j = 1, \dots, n\}$, which, en passant, reaffirms the observation made by Alhambra et al. that Definition 4 is independent of the initial state y . Note that these ideas are closely related to the calculation of extreme points given in Section 4.1 and to the index sets I_k^σ defined in (7), which are the main concepts in the proof of Proposition 1.

4. Detailed Examples

The objects introduced in Section 3.2 can be computed explicitly and they have simple graphical interpretations, e.g., via thermomajorization curves. The following examples show this in detail.

4.1. Extreme Point Map

Definition 3 contains a simple algorithm to evaluate the extreme point map $E_{d,y}(\sigma)$. Given some permutation $\sigma \in S_n$, find the value of the thermomajorization curve at $d_{\sigma(1)}, d_{\sigma(1)} + d_{\sigma(2)}, \dots, \sum_{i=1}^{n-1} d_{\sigma(i)}$, take the difference in consecutive values, and arrange them into a vector that is ordered according to σ . Let us provide a detailed example.

Let $y = (4, 0, 1)^\top$ and $d = (4, 2, 1)^\top$. One verifies $\text{th}_{d,y}(c) = \min\{c, 5\}$ for all $c \in [0, 7]$ by direct computation; cf. Figure 3 below. Now, let $\sigma \in S_3$ be the permutation $\sigma(1) = 2, \sigma(2) = 3$, and $\sigma(3) = 1$; in two-line notation, this reads $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ (henceforth, $\sigma = (2\ 3\ 1)$ for short). Our goal is to compute the extreme point $E_{d,y}(\sigma)$ of $M_d(y)$ that “corresponds” to σ . We will use the second formulation provided in Definition 3. First,

$$(E_{d,y}(\sigma))_2 = (E_{d,y}(\sigma))_{\sigma(1)} = \text{th}_{d,y}(d_{\sigma(1)}) - \text{th}_{d,y}(0) = \text{th}_{d,y}(d_2) = \min\{2, 5\} = 2,$$

followed by $(E_{d,y}(\sigma))_3 = (E_{d,y}(\sigma))_{\sigma(2)} = \text{th}_{d,y}(d_{\sigma(1)} + d_{\sigma(2)}) - \text{th}_{d,y}(d_{\sigma(1)}) = \min\{3, 5\} - \min\{2, 5\} = 1$ and analogously for $(E_{d,y}(\sigma))_1$. Thus, $E_{d,y}(\sigma) = (2, 2, 1)^\top$. This procedure can be easily visualized; cf. Figure 3.

The full image of $E_{d,y}$ is computed analogously: one finds (cf. Table 1)

$$\text{ext}(M_d(y)) = \left\{ \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \right\}.$$

With this in mind, let us reformulate the definition of the map $E_{d,y}(\sigma)$ to obtain an even better understanding.

Remark 5. Given $d \in \mathbb{R}_{++}^n$ and $\sigma \in S_n$, the permutation σ indicates how to order the segments of length d_i . These can be visualized as lying head to tail on the x -axis, i.e., as a tiling of the interval $[0, \mathbf{e}^\top d]$. The contact points between the intervals are then $d_{\sigma(1)}, d_{\sigma(1)} + d_{\sigma(2)},$ and so on. Now, we evaluate $\text{th}_{d,y}$ at these points and look at the corresponding increments $(E_{d,y}(\sigma))_{\sigma(j)} = \text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) - \text{th}_{d,y}(\sum_{i=1}^{j-1} d_{\sigma(i)})$ for $j = 1, \dots, n$, as visualized in Figure 3. This construction relates to the previously mentioned notion of tight thermomajorization, as these contact points, in turn, are the elbow points of the thermomajorization curve of $E_{d,y}(\sigma)$.

Note that, by Definition 3, $(E_{d,y}(\sigma))_{\sigma(j)}$ is the increment in $\text{th}_{d,y}$ over a distance of length $d_{\sigma(j)}$. In particular, for any $j = 1, \dots, n$, the entry $(E_{d,y}(\sigma))_j$ corresponds to the increment over the interval d_j , regardless of where it is in our tiling. Hence, two permutations σ, τ give the same extreme point $E_{d,y}(\sigma) = E_{d,y}(\tau)$ if and only if each interval d_j yields the same increment in $\text{th}_{d,y}$ for both permutations. In the example above, this happens because the vector y/d is degenerate. The following example shows, however, that this is not necessary.

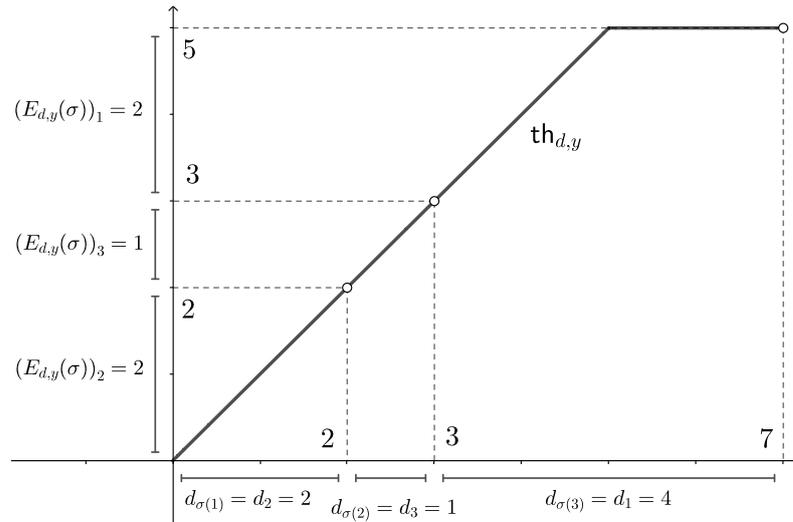


Figure 3. Visualization of the detailed calculation of $E_{d,y}(\sigma)$ for $\sigma = (2\ 3\ 1)$. First, we extract the value of $\text{th}_{d,y}$ at $d_{\sigma(1)} = d_2 = 2$, which, because we considered the second entry of d , becomes the second entry of $E_{d,y}(\sigma)$. Next, we add $d_{\sigma(2)} = d_3 = 1$ to the previous x -axis value; then, the third entry of $E_{d,y}(\sigma)$ is the corresponding increment $\text{th}_{d,y}(3) - \text{th}_{d,y}(2) = 1$. Finally, we add $d_{\sigma(3)} = d_1$ to our position on the x -axis to arrive at $\mathbf{e}^\top d = 7$, so the first entry of $E_{d,y}(\sigma)$ is given by the final increment $\text{th}_{d,y}(7) - \text{th}_{d,y}(3) = 2$.

4.2. Degeneracy

Now, we present a different example: consider $d = (1, 2, 10)^\top$, $y = (1, 4, 5)^\top$. The key insight is that even though $\frac{y}{d} = (1, 2, 0.5)^\top$ is non-degenerate, the polytope $M_d(y)$ appears to be degenerate. The reason that this is allowed to happen is that d is not well structured. Indeed, as in Section 4.1, one computes

$$\text{th}_{d,y}(c) = \begin{cases} 2c & c \in [0, 2] \\ c + 2 & c \in [2, 3] \\ 0.5c + 3.5 & c \in [3, 13] \end{cases} \quad (10)$$

for all $c \in [0, 13]$ and hence one finds

$$E_{d,y}(3\ 1\ 2) = \begin{pmatrix} \text{th}_{d,y}(11) - \text{th}_{d,y}(10) \\ \text{th}_{d,y}(13) - \text{th}_{d,y}(11) \\ \text{th}_{d,y}(10) \end{pmatrix} = \begin{pmatrix} 9 - 8.5 \\ 10 - 9 \\ 8.5 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \\ 8.5 \end{pmatrix}$$

and

$$E_{d,y}(3\ 2\ 1) = \begin{pmatrix} \text{th}_{d,y}(13) - \text{th}_{d,y}(12) \\ \text{th}_{d,y}(12) - \text{th}_{d,y}(10) \\ \text{th}_{d,y}(10) \end{pmatrix} = \begin{pmatrix} 10 - 9.5 \\ 9.5 - 8.5 \\ 8.5 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \\ 8.5 \end{pmatrix}.$$

Despite the two permutations differing, their image under $E_{d,y}$ coincides. This comes from the fact that $\text{th}_{d,y}(10) + \text{th}_{d,y}(13) = \text{th}_{d,y}(11) + \text{th}_{d,y}(12)$ because $\text{th}_{d,y}|_{[3,13]}$ is affine linear; cf. (10).

4.3. Non-Degeneracy

The following example shows that even if d fails to be well structured, this does not guarantee that $M_d(y)$ is degenerate. Indeed, choosing $d = (4, 2, 1)^\top$, $y = \frac{1}{5}(3, 1, 1)^\top$, one computes

$$\text{ext}(M_d(y)) = \frac{1}{20} \left\{ \begin{pmatrix} 12 \\ 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 13 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 13 \\ 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 11 \\ 7 \\ 2 \end{pmatrix}, \begin{pmatrix} 10 \\ 7 \\ 3 \end{pmatrix}, \begin{pmatrix} 10 \\ 6 \\ 4 \end{pmatrix} \right\}$$

(recall Section 4.1 for details on how to evaluate $E_{d,y}$). In particular, $|\text{ext}(M_d(y))| = 6 = 3!$, although d is not well structured ($1 + 2 < 4$). The reason for this is that fact that when partitioning the interval $[0, 7]$ into subintervals of length $(1, 4, 2)$ (which is the re-ordering of d such that it matches $\frac{y}{d}$ being non-increasing), there is no way to permute these subintervals such that the two small intervals are contained in the large one.

4.4. Index Sets I_k^σ

In (7), we defined the index sets I_k^σ . Here, we wish to demonstrate how to easily compute them. Using the same example as in Section 4.1, i.e., $y = (4, 0, 1)^\top$ and $d = (4, 2, 1)^\top$, we again have the thermomajorization curve in Figure 3. This curve changes slope only once (i.e., $m = 1$) and it does so at the input 5; thus, $\Delta_0 = 0$, $\Delta_1 = 5$, and $\Delta_2 = \mathbf{e}^\top d = 7$. Let us specify the sets $I_{\sigma_1}^\sigma, I_{\sigma_2}^\sigma$ for the permutation $\sigma = (2\ 3\ 1)$ from Figure 4.

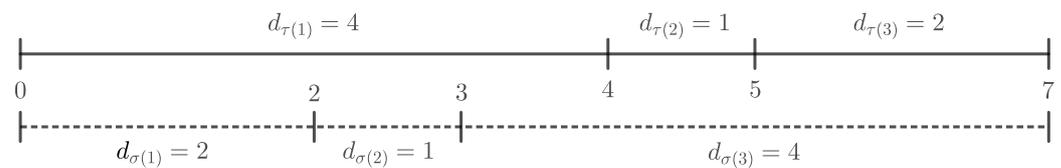


Figure 4. Comparison of the sets $\{\sum_{i=1}^j d_{\tau(i)} : j = 1, \dots, n\}$ (top) and $\{\sum_{i=1}^j d_{\sigma(i)} : j = 1, \dots, n\}$ (bottom) as subsets of the real number line.

One finds $I_1^\sigma = \{\sigma(1), \sigma(2), \sigma(3)\} = \{2, 3, 1\}$ because all three subintervals intersect the interval $(\Delta_0, \Delta_1) = (0, 5)$ corresponding to the first affine linear segment. Moreover, $I_2^\sigma = \{\sigma(3)\} = \{1\}$ because $(\Delta_1, \Delta_2) = (5, 7)$ only intersects the third subinterval (with regard to σ).

4.5. Extremal d -Stochastic Matrices

We already saw that for $y = (4, 0, 1)^\top$ and $d = (4, 2, 1)^\top$, the extreme point corresponding to $\sigma = (2\ 3\ 1)$ is $E_{d,y}(\sigma) = (2, 2, 1)^\top$. Finding a Gibbs-stochastic matrix that maps y to this extreme point via Definition 4 amounts to specifying a permutation that orders $\frac{y}{d}$ non-increasingly, e.g., choose $\tau = (1\ 3\ 2)$ because then $\underline{\tau} \frac{y}{d} = (\frac{4}{4}, \frac{1}{1}, \frac{0}{2})^\top = (1, 1, 0)^\top$. Now, all the information needed for the definition in (9) is contained in Figure 4.

This figure allows us to easily build $\underline{\sigma} A_{\sigma\tau} \underline{\tau}^{-1}$ because the non-zero entries of this matrix correspond to how much of the interval $(\sum_{i=1}^{j-1} d_{\tau(i)}, \sum_{i=1}^j d_{\tau(i)})$ is overlapped by a given element of the partition $\{d_{\sigma(i)} : i = 1, \dots, n\}$ (by slight abuse of notation, we identify $d_{\sigma(i)}$ with the interval of corresponding length). For example, $(0, d_{\sigma(1)})$ overlaps with $d_{\tau(1)}$ (covering $\frac{2}{4} = \text{half}$ of it) but not with $d_{\tau(2)}, d_{\tau(3)}$. This means that the first row of $\underline{\sigma} A_{\sigma\tau} \underline{\tau}^{-1}$ is given by $(\frac{1}{2}, 0, 0)^\top$. Similarly, the second row reads $(\frac{1}{4}, 0, 0)^\top$. On the other hand, $d_{\sigma(3)}$ in Figure 4 overlaps with all three sections $d_{\tau(1)}, d_{\tau(2)}, d_{\tau(3)}$, and the corresponding overlap ratios are given by $\frac{1}{4}, \frac{1}{1}$, and $\frac{2}{2}$. Hence, the third row of $\underline{\sigma} A_{\sigma\tau} \underline{\tau}^{-1}$ is given by $(\frac{1}{4}, 1, 1)^\top$, which altogether yields

$$\underline{\sigma} A_{\sigma\tau} \underline{\tau}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 1 & 1 \end{pmatrix} \Leftrightarrow A_{\sigma\tau} = \underline{\sigma}^{-1} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 1 & 1 \end{pmatrix} \underline{\tau} = \begin{pmatrix} \frac{1}{4} & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \end{pmatrix}.$$

One readily verifies that $A_{\sigma\tau} \in s_d(3)$ maps y to $E_{d,y}(\sigma) = (2, 2, 1)^\top$. Another observation to make here is that $(A_{\sigma\tau})_{ij}$ is always given by the portion of $d_j = d_{\tau^{-1}(j)}$ that is covered by $d_i = d_{\sigma^{-1}(i)}$. In the above example, $(A_{\sigma\tau})_{21} = \frac{1}{2}$ because half of $d_1 = d_{\tau(1)}$ is covered by $d_2 = d_{\sigma(1)}$ in Figure 4.

There is more to uncover here: on the one hand, the τ that we chose is not the only permutation that orders $\frac{y}{d}$ non-increasingly in this example, and, on the other hand, we saw in Section 4.1 that there exists a permutation $\sigma' \in S_3$ other than the σ that we chose

that is mapped to $(2, 2, 1)^\top$ under $E_{d,y}$. Thus, we can apply the above procedure to $2 \cdot 2 = 4$ combinations of permutations (σ, τ) , which all yield extremal Gibbs-stochastic matrices mapping y to $(2, 2, 1)^\top$. These are given in Table 2 below.

Table 2. Different combinations of $\sigma, \tau \in S_3$ and the corresponding matrix $\sigma A_{\sigma\tau} \tau^{-1}$ (left), as well as $A_{\sigma\tau}$ (right).

$\sigma \downarrow / \tau \rightarrow$	(1 3 2)	(3 1 2)	$\sigma \downarrow / \tau \rightarrow$	(1 3 2)	(3 1 2)
(2 3 1)	$\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$	(2 3 1)	$\begin{pmatrix} \frac{1}{4} & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & 0 & 1 \\ \frac{1}{4} & 0 & 0 \end{pmatrix}$
(3 2 1)	$\begin{pmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$	(3 2 1)	$\begin{pmatrix} \frac{1}{4} & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

5. Conclusions

In this work, building upon [8,15,17–20], we further explored thermomajorization in the quasi-classical realm, as well as the rich geometry of the associated polytope. The former notion comes from the resource theory approach to quantum thermodynamics, and, in particular, the corresponding set of allowed operations, known as thermal operations.

Inspired by transportation theory, the core notions of this work were “stable” and “well-structured” Gibbs vectors, which are simple conditions on the spectrum of a Gibbs state. We found that these concepts relate to and give rise to conceptual insights and unexpected results regarding system properties and state transfers. On the one hand, quasi-classical cyclic state transfers with regard to thermomajorization are impossible if and only if the Gibbs vector is stable (which is the generic case). Put differently, within the model of (quasi-classical) thermal operations, performing cyclic state transfers in general comes with a non-zero work cost. On the other hand, we uncovered two connections to the notion of subspaces in equilibrium. (1). Thermal operations can bring any subspace in equilibrium out of equilibrium without having to expend work. This generalizes the intuitive fact that a system in a Gibbs state can be brought out of equilibrium by coupling it to a non-equilibrium system. Note that for the latter scenario—while any out-of-equilibrium system is necessarily a resource—our result shows that this is the only price that one has to pay, i.e., once the systems are coupled, there is some process on the composite system that takes the original system out of the Gibbs state and that can be implemented at no work cost. (2). The existence of subspaces in equilibrium is reflected in the geometry of the thermomajorization polytope. Indeed, assuming the well-structuredness of the Gibbs state—which is equivalent to the system’s temperature exceeding a critical value—the existence of a subspace in equilibrium corresponds to the polytope having degenerate corners.

While we explored the case of quasi-classical states in detail, as usual in quantum thermodynamics, the general case is vastly more difficult. Most notably, the number of extreme points of the thermal operations, as well as the number of extreme points of the future thermal cone, is uncountable for non-classical initial states [7,21]. Hence, one loses access to tools from the theory of convex polytopes. One of the few notions that pertain to general systems is the notion of well-structured and stable Gibbs states; investigating whether these notions encode any properties of general thermal cones (such as, e.g., the impossibility of cyclic processes) could be an interesting topic for future research.

Author Contributions: Conceptualization, F.v.E. and E.M.; methodology, F.v.E. and E.M.; software, F.v.E. and E.M.; investigation, F.v.E. and E.M.; writing—original draft preparation, F.v.E. and E.M.; writing—review and editing, F.v.E. and E.M.; visualization, F.v.E. and E.M.; supervision, F.v.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We are grateful to Gunther Dirr and Thomas Schulte-Herbrüggen for their valuable and constructive comments during the preparation of this manuscript. Moreover, we would like to thank the the anonymous referees for their helpful comments. F.v.E. has been supported by the Einstein Foundation (Einstein Research Unit on Quantum Devices) and the MATH+ Cluster of Excellence. E.M. is part of the Bavarian excellence network ENB via the International PhD Programme of Excellence *Exploring Quantum Matter* (EXQM), as well as the *Munich Quantum Valley* of the Bavarian State Government with funds from *Hightech Agenda Bayern Plus*.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviation

The following abbreviation is used in this manuscript:

w.l.o.g. without loss of generality

Appendix A. Thermomajorization in the Zero-Temperature Limit

Before proving the result in question, we first need the following lemma.

Lemma A1. *Let $x, y, d \in \mathbb{R}^n$ with $d > 0$ as well as numbers $x_+, x_-, y_+, y_- \geq 0$ be given. If $\mathbf{e}^\top x + x_+ - x_- = \mathbf{e}^\top y + y_+ - y_-$ and*

$$\|d_i x - y_i d\|_1 + d_i(x_+ + x_-) \leq \|d_i y - y_i d\|_1 + d_i(y_+ + y_-) \tag{A1}$$

for all $i = 1, \dots, n$, then $x_- \leq y_-$ and $x_+ \leq y_+$.

Proof. Combining (A1) with the trace condition yields

$$\|d_i x - y_i d\|_1 + d_i(x_+ + x_-) \leq \|d_i y - y_i d\|_1 + d_i(y_+ + (\mathbf{e}^\top y + y_+ - \mathbf{e}^\top x - x_+ + x_-))$$

which is equivalent to $\|d_i x - y_i d\|_1 \leq \|d_i y - y_i d\|_1 + d_i \mathbf{e}^\top (y - x) + 2d_i(y_+ - x_+)$ and thus

$$y_+ - x_+ \geq \frac{\|d_i x - y_i d\|_1 - \|d_i y - y_i d\|_1 - d_i \mathbf{e}^\top (y - x)}{2d_i}. \tag{A2}$$

Hence, all we have to find is $i \in \{1, \dots, n\}$ such that the right-hand side of (A2) is non-negative. For this, let $i = 1, \dots, n$ be the index that satisfies $\frac{y_i}{d_i} = \max_{j=1, \dots, n} \frac{y_j}{d_j}$. Then, one finds $d_i y_j - y_i d_j = d_i d_j (\frac{y_j}{d_j} - \frac{y_i}{d_i}) \leq 0$ for all $j = 1, \dots, n$; hence, $\|d_i y - y_i d\|_1 = -\mathbf{e}^\top (d_i y - y_i d)$. Therefore, the numerator of the right-hand side of (A2) for this index i is equal to $\|d_i x - y_i d\|_1 + \mathbf{e}^\top (d_i y - y_i d) - \mathbf{e}^\top (d_i y) + \mathbf{e}^\top (d_i x) = \|d_i x - y_i d\|_1 + \mathbf{e}^\top (d_i x - y_i d) \geq 0$. The inequality $x_- \leq y_-$ is shown analogously by replacing y_+ (instead of y_-) in (A1) via the trace condition. One then shows $\|d_i x - y_i d\|_1 - \|d_i y - y_i d\|_1 + d_i \mathbf{e}^\top (y - x) \geq 0$, similar to our argument above, by choosing i to be the index such that $\frac{y_i}{d_i} = \min_{j=1, \dots, n} \frac{y_j}{d_j}$, because then $\|d_i y - y_i d\|_1 = \mathbf{e}^\top (d_i y - y_i d)$. \square

With this, we can show that the characterization of thermomajorization for quasi-classical states (Section 2.1) extends to the case of temperature zero, assuming that the lowest energy level of the system is non-degenerate. What this means mathematically is that the (normalized) Gibbs vector becomes some standard basis vector e_j in the limit $T \rightarrow 0^+$. Without loss of generality, $j = 1$, because, given any permutation matrix $\sigma \in \mathbb{R}^{n \times n}$, $x \prec_d y$ is equivalent $\sigma x \prec_{\sigma d} \sigma y$. We claim that the following statements are equivalent for all $x, y \in \mathbb{R}^n$:

- (i) There exists $A \in \mathbb{R}_+^{n \times n}$ such that $\mathbf{e}^\top A = \mathbf{e}^\top$, $Ae_1 = e_1$, and $Ay = x$.
- (ii) $\mathbf{e}^\top x = \mathbf{e}^\top y$, and for all $t \in \mathbb{R}$, one has $\|x - te_1\|_1 \leq \|y - te_1\|_1$.
- (iii) $\mathbf{e}^\top x = \mathbf{e}^\top y$ and $\|x - y_1 e_1\|_1 \leq \|y - y_1 e_1\|_1$.
- (iv) $\mathbf{e}^\top x = \mathbf{e}^\top y$, and for all $c \in [0, \mathbf{e}^\top d]$, one has $\text{th}_{d,x}(c) \leq \text{th}_{d,y}(c)$.

Proof. “(ii) ⇔ (iv)”: Equation (3). “(i) ⇒ (ii)”: By assumption, $\mathbf{e}^\top x = (\mathbf{e}^\top A)y = \mathbf{e}^\top y$ and $x - y_j e_j = Ay - y_j A e_j = A(y - y_j e_j)$. This, together with the fact that every column-stochastic matrix is 1-norm contractive (Remark 1 (i) [27]), implies (ii). “(ii) ⇒ (iii)”: Trivial. “(iii) ⇒ (i)”: The idea is as follows: partition $y = \begin{pmatrix} y_1 \\ y' \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x' \end{pmatrix}$, where $x', y' \in \mathbb{R}^{n-1}$, and perform the following chain of transformations:

$$y = \begin{pmatrix} y_1 \\ y' \end{pmatrix} \xrightarrow{(1)} \begin{pmatrix} y_1 \\ y_+ \\ -y_- \\ 0 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} y_1 + (y_+ - x_+) - (y_- - x_-) \\ x_+ \\ -x_- \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_+ \\ -x_- \\ 0 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} x_1 \\ x' \end{pmatrix} = x$$

Before explaining the undefined objects, let us make the following key observation. If some entries of d are zero, then any d -stochastic matrix has to be block triangular. More precisely, given $A \in \mathbb{R}_+^{n \times n}$ and $d = \begin{pmatrix} \tilde{d} \\ 0_{n-m} \end{pmatrix}$, where $\tilde{d} \in \mathbb{R}^m$, $\tilde{d} > 0$, the matrix A is d -stochastic if and only if there exist a \tilde{d} -stochastic matrix $\tilde{A} \in \mathbb{R}_+^{m \times m}$ and probability vectors $v_1, \dots, v_k \in \mathbb{R}_+^n$ (i.e., $\mathbf{e}^\top v_1 = \dots = \mathbf{e}^\top v_k = 1$) such that

$$A = \begin{pmatrix} \tilde{A} & & & \\ 0_{(n-m) \times m} & v_1 & \dots & v_k \end{pmatrix}.$$

Now, let us detail the above instructions step by step.

- (1) Define row vectors $w_+, w_- \in \mathbb{R}^{n-1}$ as follows: given $i = 1, \dots, n - 1$ set $(w_+)_i := 1$ if $y'_i \geq 0$, else $(w_+)_i := 0$. Then, set $w_- := \mathbf{e}^\top - w_+$. Thus, w_+ indicates the location of the non-negative entries of y' , and w_- indicates where the negative entries of y' are. Now, the column-stochastic matrix

$$A_1 := \begin{pmatrix} 1 & 0 \\ 0 & w_+ \\ 0 & w_- \\ 0 & 0 \end{pmatrix}$$

collects the positive entries of y' in the second position, and the negative entries of y' in the third position (We may, without loss of generality, assume $n \geq 3$ because we can append as many zeros to the original vectors as we wish without altering the problem: one readily verifies that $x \prec_d y$ is equivalent to $\begin{pmatrix} x \\ 0_k \end{pmatrix} \prec_{\begin{pmatrix} d \\ 0_k \end{pmatrix}} \begin{pmatrix} y \\ 0_k \end{pmatrix}$ for all $x, y, d \in \mathbb{R}^n$,

$d \geq 0$, and $k \in \mathbb{N}$). We write $y_+ := w_+ y'$ and $y_- := -(w_- y')$.

- (2) Similarly, define $x_+ := \sum_{i=1}^k \max\{x'_i, 0\}$, $x_- := -\sum_{i=1}^k \min\{x'_i, 0\}$. Note $x_\pm, y_\pm \geq 0$, so evaluating the 1-norm inequality from (iii) readily implies the assumptions of Lemma A1 (when setting $n = 1$, $x := x_1$, $y := y_1$, $d := 1$). Hence, Lemma A1 shows $x_- \leq y_-, x_+ \leq y_+$. This implies that

$$A_2 := \begin{pmatrix} 1 & 1 - \frac{x_+}{y_+} & 1 - \frac{x_-}{y_-} \\ 0 & \frac{x_+}{y_+} & 0 \\ 0 & 0 & \frac{x_-}{y_-} \\ 0 & 0 & 0 \end{pmatrix}$$

has non-negative entries so A_2 is column-stochastic, has e_j as a fixed point, and carries out the desired transformation. Note that if y_+ (y_-) is zero, as is x_+ (x_-), meaning that we can set $\frac{x_+}{y_+} := 1$ ($\frac{x_-}{y_-} := 1$). Concluding this step, $\mathbf{e}^\top x = \mathbf{e}^\top y$ implies $y_1 + (y_+ - x_+) - (y_- - x_-) = x_1$.

- (3) Similarly to (1), define A_3 such that the bottom right $(n - 1) \times (n - 1)$ block of A_3 redistributes the accumulated positive and negative entries of x' back to its original place. □

Based on this result, it seems reasonable to conjecture that, for all $d \in \mathbb{R}_+^n$, $x \prec_d y$ is equivalent to $\mathbf{e}^\top x = \mathbf{e}^\top y$ and $\|d_i x - y_i d\|_1 \leq \|d_i y - y_i d\|_1, i = 1, \dots, n$. However, proving this seems to not be as straightforward as the above non-degenerate temperature-zero case, and it would be beyond the scope of this article in any case.

Appendix B. Basic Properties of Concave Functions

We start with the following simple observation. Given a compact interval $I \subseteq \mathbb{R}$, a concave function $f : I \rightarrow \mathbb{R}$, and $r, s, t \in I$ with $r < s < t$, one finds

$$f\left(\frac{s-r}{t-r}t + \frac{t-s}{t-r}r\right) \geq \frac{s-r}{t-r}f(t) + \frac{t-s}{t-r}f(r)$$

which in turn is equivalent to

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(s) - f(r)}{s - r}. \tag{A3}$$

The following now is a direct consequence of this.

Lemma A2. *Let $d \in \mathbb{R}_{++}^n$, $y \in \mathbb{R}^n$, $\sigma \in S_n$. One has $\text{th}_{d,y}(\sum_{i=1}^j d_{\sigma(i)}) = \sum_{i=1}^j y_{\sigma(i)}$ for all $j = 1, \dots, n - 1$ if and only if $\frac{y_{\sigma(1)}}{d_{\sigma(1)}} \geq \dots \geq \frac{y_{\sigma(n)}}{d_{\sigma(n)}}$.*

Proof. “ \Leftarrow ”: Shown in (Lemma 15 (iii) [27]). “ \Rightarrow ”: Because $\text{th}_{d,y}$ is concave (Lemma 15 (i) [27]), we may insert $r = \sum_{i=1}^{j-1} d_{\sigma(i)}$, $s = \sum_{i=1}^j d_{\sigma(i)}$, $t = \sum_{i=1}^{j+1} d_{\sigma(i)}$ for $j = 2, \dots, n - 1$ in (A3). This yields the desired inequality. \square

Moreover, the following basic property of concave functions will be essential for our analysis of thermomajorization curves.

Lemma A3. *Let $I \subset \mathbb{R}$ be a compact interval, i.e., $I = [a, b]$ for some $a, b \in \mathbb{R}$, $a < b$. Given $f : I \rightarrow \mathbb{R}$ concave or convex, the following statements are equivalent.*

(i) f is affine linear, i.e., for all $x \in [a, b]$

$$f(x) = \frac{x - a}{b - a}f(b) + \frac{b - x}{b - a}f(a).$$

- (ii) For all $y_1, y_2 \in (a, b)$ that satisfy $y_1 + y_2 = a + b$, one has $f(y_1) + f(y_2) = f(a) + f(b)$.
- (iii) There exist $y_1, y_2 \in (a, b)$ such that $y_1 + y_2 = a + b$ and $f(y_1) + f(y_2) = f(a) + f(b)$.
- (iv) There exists $y \in (a, b)$ such that $(y, f(y))$ lies on the line that connects $(a, f(a))$ and $(b, f(b))$.
More precisely,

$$f(y) = \frac{y - a}{b - a}f(b) + \frac{b - y}{b - a}f(a).$$

Proof. “(i) \Rightarrow (ii)”: Straightforward calculation. “(ii) \Rightarrow (iii)”: Trivial. “(iii) \Rightarrow (iv)”: Assume without loss of generality that f is concave—the convex case is examined analogously. First, note that concavity implies

$$f(x) = f\left(\frac{x - a}{b - a}b + \frac{b - x}{b - a}a\right) \geq \frac{x - a}{b - a}f(b) + \frac{b - x}{b - a}f(a) \tag{A4}$$

for all $x \in [a, b]$. This leads to

$$f(a) + f(b) = f(y_1) + f(y_2) \geq f(y_1) + \frac{y_2 - a}{b - a}f(b) + \frac{b - y_2}{b - a}f(a) \tag{A5}$$

and thus

$$\begin{aligned} \frac{y_1 - a}{b - a} f(b) + \frac{b - y_1}{b - a} f(a) &\stackrel{(A4)}{\leq} f(y_1) \\ &\stackrel{(A5)}{\leq} \left(1 - \frac{y_2 - a}{b - a}\right) f(b) + \left(1 - \frac{b - y_2}{b - a}\right) f(a) \\ &= \frac{b - a - y_2 + a}{b - a} f(b) + \frac{b - a - b + y_2}{b - a} f(a) \\ &= \frac{y_1 - a}{b - a} f(b) + \frac{b - y_1}{b - a} f(a). \end{aligned}$$

which is the essence of the demonstration. In the last step, we use $y_1 = a + b - y_2$ (because $y_1 + y_2 = a + b$ by assumption).

“(iv) \Rightarrow (i)”: Again, assume without loss of generality that f is concave—the convex case is demonstrated similarly. Let $x \in (a, y)$ be given. Using the assumption on $f(y)$, as well as concavity, we compute

$$\frac{y - a}{b - a} f(b) + \frac{b - y}{b - a} f(a) = f(y) = f\left(\frac{y - x}{b - x} b + \frac{b - y}{b - x} x\right) \geq \frac{y - x}{b - x} f(b) + \frac{b - y}{b - x} f(x)$$

and thus

$$\begin{aligned} f(x) &\leq \frac{b - x}{b - y} \left(\frac{y - a}{b - a} f(b) + \frac{b - y}{b - a} f(a) - \frac{y - x}{b - x} f(b)\right) \\ &= \frac{(b - y)(x - a)}{(b - y)(b - a)} f(b) + \frac{b - x}{b - a} f(a) = \frac{x - a}{b - a} f(b) + \frac{b - x}{b - a} f(a) \end{aligned}$$

by means of straightforward computation. The converse inequality comes from (A4); hence, the equality holds. The case $x \in (y, b)$ is examined analogously, which concludes the proof. \square

References

1. Vinjanampathy, S.; Anders, J. Quantum Thermodynamics. *Contemp. Phys.* **2016**, *57*, 545–579. [[CrossRef](#)]
2. Binder, F.; Correa, L.; Gogolin, C.; Anders, J.; Adesso, G. (Eds.) *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*; Springer International: Cham, Switzerland, 2018.
3. Kosloff, R. Quantum Thermodynamics: A Dynamical Viewpoint. *Entropy* **2013**, *15*, 2100–2128. [[CrossRef](#)]
4. Skrzypczyk, P.; Short, A.; Popescu, S. Work Extraction and Thermodynamics for Individual Quantum Systems. *Nat. Commun.* **2014**, *5*, 4185. [[CrossRef](#)] [[PubMed](#)]
5. Lostaglio, M. An Introductory Review of the Resource Theory Approach to Thermodynamics. *Rep. Prog. Phys.* **2019**, *82*, 114001. [[CrossRef](#)]
6. Janzing, D.; Wocjan, P.; Zeier, R.; Geiss, R.; Beth, T. Thermodynamic Cost of Reliability and Low Temperatures: Tightening Landauer’s Principle and the Second Law. *Int. J. Theor. Phys.* **2000**, *39*, 2717–2753. [[CrossRef](#)]
7. vom Ende, F. Which Bath-Hamiltonians Matter for Thermal Operations? *J. Math. Phys.* **2022**, *63*, 112202. [[CrossRef](#)]
8. Alhambra, Á.; Lostaglio, M.; Perry, C. Heat-Bath Algorithmic Cooling with Optimal Thermalization Strategies. *Quantum* **2019**, *3*, 188. [[CrossRef](#)]
9. Horodecki, M.; Oppenheim, J. Fundamental Limitations for Quantum and Nanoscale Thermodynamics. *Nat. Commun.* **2013**, *4*, 2059. [[CrossRef](#)]
10. Brandão, F.; Horodecki, M.; Ng, N.; Oppenheim, J.; Wehner, S. The Second Laws of Quantum Thermodynamics. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3275–3279. [[CrossRef](#)]
11. Ruch, E.; Schranner, R.; Seligman, T. The Mixing Distance. *J. Chem. Phys.* **1978**, *69*, 386–392. [[CrossRef](#)]
12. Renes, J. Work Cost of Thermal Operations in Quantum Thermodynamics. *Eur. Phys. J. Plus* **2014**, *129*, 153. [[CrossRef](#)]
13. Faist, P.; Oppenheim, J.; Renner, R. Gibbs—Preserving Maps Outperform Thermal Operations in the Quantum Regime. *New J. Phys.* **2015**, *17*, 043003. [[CrossRef](#)]
14. Gour, G.; Müller, M.; Narasimhachar, V.; Spekkens, R.; Halpern, N. The Resource Theory of Informational Nonequilibrium in Thermodynamics. *Phys. Rep.* **2015**, *583*, 1–58. [[CrossRef](#)]
15. Lostaglio, M.; Alhambra, Á.; Perry, C. Elementary Thermal Operations. *Quantum* **2018**, *2*, 1–52. [[CrossRef](#)]
16. Sagawa, T.; Faist, P.; Kato, K.; Matsumoto, K.; Nagaoka, H.; Brandão, F. Asymptotic Reversibility of Thermal Operations for Interacting Quantum Spin Systems via Generalized Quantum Stein’s Lemma. *J. Phys. A* **2021**, *54*, 495303. [[CrossRef](#)]

17. Mazurek, P. Thermal Processes and State Achievability. *Phys. Rev. A* **2019**, *99*, 042110. [[CrossRef](#)]
18. Lostaglio, M.; Jennings, D.; Rudolph, T. Description of Quantum Coherence in Thermodynamic Processes Requires Constraints Beyond Free Energy. *Nat. Commun.* **2015**, *6*, 6383. [[CrossRef](#)]
19. Korzekwa, K. Structure of the Thermodynamic Arrow of Time in Classical and Quantum Theories. *Phys. Rev. A* **2017**, *95*, 052318. [[CrossRef](#)]
20. de Oliveira, A., Jr.; Czartowski, J.; Życzkowski, K.; Korzekwa, K. Geometric Structure of Thermal Cones. *Phys. Rev. E* **2022**, *106*, 064109.
21. Lostaglio, M.; Korzekwa, K.; Jennings, D.; Rudolph, T. Quantum Coherence, Time-Translation Symmetry, and Thermodynamics. *Phys. Rev. X* **2015**, *5*, 021001. [[CrossRef](#)]
22. Shiraishi, N. Two Constructive Proofs on d -Majorization and Thermo-Majorization. *J. Phys. A* **2020**, *53*, 425301. [[CrossRef](#)]
23. vom Ende, F. Strict Positivity and D -Majorization. *Linear Multilinear Algebra* **2022**, *70*, 4023–4048. [[CrossRef](#)]
24. Perry, C.; Ćwikliński, P.; Anders, J.; Horodecki, M.; Oppenheim, J. A Sufficient Set of Experimentally Implementable Thermal Operations. *Phys. Rev. X* **2018**, *8*, 041049.
25. Dirr, G.; vom Ende, F.; Schulte-Herbrüggen, T. Reachable Sets from Toy Models to Controlled Markovian Quantum Systems. In Proceedings of the 2019 IEEE 58th Conference on Decision and Control (IEEE-CDC), Nice, France, 11–13 December 2019; Volume 58, p. 2322.
26. vom Ende, F.; Malveti, E.; Dirr, G.; Schulte-Herbrüggen, T. Exploring the Limits of Controlled Markovian Quantum Dynamics with Thermal Resources. *Open Syst. Inf. Dyn.* **2023**, *30*, 2350005. [[CrossRef](#)]
27. vom Ende, F.; Dirr, G. The d -Majorization Polytope. *Linear Algebra Appl.* **2022**, *649*, 152–185. [[CrossRef](#)]
28. Marshall, A.; Olkin, I.; Arnold, B. *Inequalities: Theory of Majorization and Its Applications*, 2nd ed.; Springer: New York, NY, USA, 2011.
29. Borwein, J.; Lewis, A. *Convex Analysis and Nonlinear Optimization*, 2nd ed.; Springer: New York, NY, USA, 2006.
30. Hartfiel, D. A Study of Convex Sets of Stochastic Matrices Induced by Probability Vectors. *Pac. J. Math.* **1974**, *52*, 405–418. [[CrossRef](#)]
31. Grünbaum, B. *Convex Polytopes*, 2nd ed.; Graduate Texts in Mathematics; Springer: New York, NY, USA, 2003.
32. Klee, V.; Witzgall, C. Facets and Vertices of Transportation Polytopes. In *Mathematics of the Decision Sciences: Part I*; Boeing Scientific Research Laboratories: Melville, NY, USA, 1968; pp. 257–282.
33. Bolker, E. Transportation Polytopes. *J. Comb. Theory. Ser. B* **1972**, *13*, 251–262. [[CrossRef](#)]
34. Dubois, J. Polytopes de Transport Symétriques. *Discrete Math.* **1973**, *4*, 1–27. [[CrossRef](#)]
35. Brualdi, R. *Combinatorial Matrix Classes*; Cambridge University Press: Cambridge, UK, 2006.
36. Mazurek, P.; Horodecki, M. Decomposability and Convex Structure of Thermal Processes. *New J. Phys.* **2018**, *20*, 053040. [[CrossRef](#)]
37. Loewy, R.; Shier, D.; Johnson, C. Perron Eigenvectors and the Symmetric Transportation Polytope. *Linear Algebra Appl.* **1991**, *150*, 139–155. [[CrossRef](#)]
38. Perfect, H.; Mirsky, L. Extreme Points of Certain Convex Polytopes. *Mon. Hefte Math.* **1964**, *68*, 143–149. [[CrossRef](#)]
39. Lipka-Bartosik, P.; Wilming, H.; Ng, N. Catalysis in Quantum Information Theory. 2023. Available online: <http://xxx.lanl.gov/abs/2306.00798> (accessed on 15 December 2023).
40. Schrijver, A. *Theory of Linear and Integer Programming*; Wiley & Sons: New York, NY, USA, 1986.
41. Willard, S. *General Topology*; Addison Wesley: Reading, MA, USA, 1970.
42. Singh, U.; Das, S.; Cerf, N. Partial Order on Passive States and Hoffman Majorization in Quantum Thermodynamics. *Phys. Rev. Res.* **2021**, *3*, 033091. [[CrossRef](#)]
43. Skrzypczyk, P.; Silva, R.; Brunner, N. Passivity, Complete Passivity, and Virtual Temperatures. *Phys. Rev. E* **2015**, *91*, 052133. [[CrossRef](#)] [[PubMed](#)]
44. Brunner, N.; Linden, N.; Popescu, S.; Skrzypczyk, P. Virtual Qubits, Virtual Temperatures, and the Foundations of Thermodynamics. *Phys. Rev. E* **2012**, *85*, 051117. [[CrossRef](#)]
45. Joe, H. Majorization and Divergence. *J. Math. Anal. Appl.* **1990**, *148*, 287–305. [[CrossRef](#)]
46. Brøndsted, A. *An Introduction to Convex Polytopes*; Graduate Texts in Mathematics; Springer: New York, NY, USA, 1983; Volume 90.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.